

537

1195

К. В. ШАЛИМОВА

ФИЗИКА

ПОЛУПРОВОДНИКОВ



6Ф0.32

Ш18

УДК 621.315.592(075.8)

КЛАВДИЯ ВАСИЛЬЕВНА ШАЛИМОВА
ФИЗИКА ПОЛУПРОВОДНИКОВ

Редактор Л. П. Павлов
Редактор издательства А. А. Цитленко
Переплет художника А. А. Иванова
Технический редактор Т. А. Маслова
Корректор А. Д. Халанская

▽

Сдано в набор 16/II 1976 г. Подписано к печати 31/VIII 1976 г. Т-15761.
Формат 60X90^{1/16}. Бумага типографская № 3. Усл. печ. л. 26,0.
Уч.-изд. л. 27,41. Тираж 25 000 экз. Зак. № 604. Цена 1 р. 12 к.

▽

Издательство «Энергия», Москва, М-114, Шлюзовая наб., 10.

Ленинградская типография № 4 Союзполиграфпрома при Государственном комитете Совета Министров СССР по делам издательства, полиграфии и книжной торговли, 196126, Ленинград, Ф-126, Социалистическая ул., 14.

К. В. Шалимова

Ш18 **Физика полупроводников. Учебник для студентов вузов, обучающихся по специальности «Полупроводниковые и микроэлектронные приборы».**
Изд. 2-е, перераб. и доп. М., «Энергия», 1976.

416 с. с ил.

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

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

Ш - 30407-439 135-76
051(01)-76

6Ф0.32

© Издательство «Энергия», 1976.

ОГЛАВЛЕНИЕ

Предисловие	7
Список основных обозначений	8
Глава первая	
ПОЛУПРОВОДНИКИ. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ЭЛЕКТРОПРОВОДНОСТИ	11
1-1. Классификация веществ по величине удельной проводимости. Полупроводники	11
1-2. Модельные представления о механизме электропроводности собственных полупроводников	18
1-3. Модельные представления о механизме электропроводности примесных полупроводников	24
1-4. Элементарная теория электропроводности полупроводников	26
Глава вторая	
ОСНОВЫ ЗОННОЙ ТЕОРИИ ПОЛУПРОВОДНИКОВ	29
2-1. Уравнение Шредингера для кристалла	29
2-2. Адиабатическое приближение и валентная аппроксимация	31
2-3. Одноэлектронное приближение	32
2-4. Приближение сильно связанных электронов	36
2-5. Число состояний электронов в энергетической зоне	42
2-6. Квазимпульс	45
2-7. Зоны Бриллюэна	46
2-8. Возможное заполнение электронных состояний валентной зоны	48
2-9. Зависимость энергии электрона от волнового вектора у дна и потолка энергетической зоны	50
2-10. Движение электронов в кристалле под действием внешнего электрического поля	54
2-11. Эффективная масса носителей заряда	60
2-12. Циклотронный резонанс	67
2-13. Зонная структура некоторых полупроводников	68
2-14. Метод эффективной массы	74
2-15. Элементарная теория примесных состояний	76
Глава третья	
КОЛЕБАНИЯ АТОМОВ КРИСТАЛЛИЧЕСКОЙ РЕШЕТКИ	79
3-1. Одномерные колебания однородной струны	79
3-2. Колебания одноатомной линейной цепочки	81
3-3. Кинетическая и потенциальная энергия одномерной решетки. Нормальные координаты	85
3-4. Колебания двухатомной линейной цепочки	87
3-5. Колебания атомов трехмерной решетки	91
3-6. Статистика фононов	94
3-7. Теплоемкость кристаллической решетки	96
3-8. Тепловое расширение и тепловое сопротивление твердого тела	102

Глава четвертая

СТАТИСТИКА ЭЛЕКТРОНОВ И ДЫРОК В ПОЛУПРОВОДНИКАХ 105

4-1. Плотность квантовых состояний	105
4-2. Функция распределения Ферми—Дирака	109
4-3. Степень заполнения примесных уровней	111
4-4. Концентрация электронов и дырок	113
4-5. Примесный полупроводник	116
4-6. Собственный полупроводник	124
4-7. Зависимость положения уровня Ферми от концентрации примеси и температуры для невырожденного полупроводника	128
4-8. Зависимость положения уровня Ферми от температуры для невырожденного полупроводника с частично компенсированной примесью	136
4-9. Примесные зоны	140

Глава пятая

МЕХАНИЗМЫ РАССЕЯНИЯ НОСИТЕЛЕЙ ЗАРЯДА

В ПОЛУПРОВОДНИКАХ 144

5-1. Кинетическое уравнение Больцмана	144
5-2. Равновесное состояние	150
5-3. Время релаксации	150
5-4. Эффективное сечение рассеяния. Типы центров рассеяния	154
5-5. Рассеяние на ионах примеси	156
5-6. Рассеяние на атомах примеси и дислокациях	161
5-7. Рассеяние на тепловых колебаниях решетки	162

Глава шестая

КИНЕТИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ 170

6-1. Неравновесная функция распределения	170
6-2. Удельная проводимость полупроводников	173
6-3. Зависимость подвижности носителей заряда от температуры	176
6-4. Эффект Холла	181
6-5. Эффект Холла в полупроводниках с двумя типами носителей заряда	185
6-6. Магниторезистивный эффект	190
6-7. Термоэлектрические явления	195
6-8. Теплопроводность полупроводников	203
6-9. Электропроводность полупроводников в сильном электрическом поле	205
6-10. Эффект Ганна	211
6-11. Ударная ионизация	215
6-12. Туннельный эффект и электростатическая ионизация	218

Глава седьмая

ГЕНЕРАЦИЯ И РЕКОМБИНАЦИЯ НЕРАВНОВЕСНЫХ

НОСИТЕЛЕЙ ЗАРЯДА. 221

7-1. Равновесные и неравновесные носители заряда	221
7-2. Биполярная световая генерация носителей заряда	224
7-3. Монополярная световая генерация. Максвелловское время релаксации	226
7-4. Виды рекомбинации	228
7-5. Межзонная излучательная рекомбинация	229
7-6. Межзонная ударная рекомбинация	234
7-7. Рекомбинация носителей заряда через рекомбинационные ловушки.	236

7-8. Температурная зависимость времени жизни носителей заряда при рекомбинации через рекомбинационные ловушки	243
7-9. Ловушки захвата и рекомбинационные ловушки	246

Глава восьмая

ДИФФУЗИЯ И ДРЕЙФ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА 249

8-1. Уравнение непрерывности	249
8-2. Диффузионный и дрейфовый токи	251
8-3. Соотношение Эйнштейна	253
8-4. Диффузия и дрейф неравновесных носителей заряда в случае монополярной проводимости	254
8-5. Диффузия и дрейф неосновных избыточных носителей заряда в примесном полупроводнике	257
8-6. Диффузия и дрейф неравновесных носителей заряда в полупроводнике с проводимостью, близкой к собственной	262

Глава девятая

КОНТАКТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ 266

9-1. Полупроводник во внешнем электрическом поле	266
9-2. Работа выхода	270
9-3. Контакт металл—металл. Контактная разность потенциалов	273
9-4. Контакт металл—полупроводник	276
9-5. Выпрямление на контакте металл—полупроводник	280
9-6. Двухдиагональная теория выпрямления	283
9-7. Диффузионная теория выпрямления	286
9-8. Контакт электронного и дырочного полупроводников	288
9-9. Выпрямление на <i>p-n</i> переходе	293
9-10. Теория тонкого <i>p-n</i> перехода	295

Глава десятая

ПОВЕРХНОСТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ 301

10-1. Природа поверхностных уровней	301
10-2. Теория слоя пространственного заряда	304
10-3. Эффект поля	309
10-4. Поверхностная рекомбинация	317
10-5. Влияние поверхностной рекомбинации на время жизни носителей заряда	320

Глава одиннадцатая

ПОГЛОЩЕНИЕ СВЕТА ПОЛУПРОВОДНИКАМИ 323

11-1. Спектр отражения и спектр поглощения	323
11-2. Собственное поглощение при прямых переходах	325
11-3. Собственное поглощение при непрямых переходах	330
11-4. Собственное поглощение сильно легированного полупроводника	335
11-5. Влияние внешних воздействий на собственное поглощение полупроводников	337
11-6. Экситонное поглощение	345
11-7. Поглощение свободными носителями заряда	349
11-8. Примесное поглощение	356
11-9. Решеточное поглощение	357

Глава двенадцатая

ЛЮМИНЕСЦЕНЦИЯ ПОЛУПРОВОДНИКОВ 359

12-1. Типы люминесценции	359
12-2. Мономолекулярное свечение твердых тел	360

12-3. Рекомбинационное излучение полупроводников при фундаментальных переходах	361
12-4. Рекомбинационное излучение при переходах между зоной и примесными уровнями	366
12-5. Релаксация люминесценции полупроводников	370
12-6. Температурное тушение люминесценции полупроводников	371
12-7. Спонтанное и вынужденное излучение атома	372
12-8. Стимулированное излучение твердых тел	377

Глава тринадцатая

ФОТОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ . . . 383

13-1. Внутренний фотоэффект	383
13-2. Фотопроводимость	387
13-3. Релаксация фотопроводимости	390
13-4. Фотопроводимость при наличии поверхностной рекомбинации и диффузии носителей заряда	393
13-5. Эффект Дембера	395
13-6. Фотоэлектромагнитный эффект	396
13-7. Фотовольтаический эффект в <i>p-n</i> переходах	399
13-8. Фотовольтаический эффект на барьере Шоттки	404
13-9. Внешний фотоэффект	405

Приложения 408

I. Физические константы	408
II. Свойства Ge, Si и GaAs (при 300 К)	409
III. Свойства полупроводников	410
Предметный указатель	412

ПРЕДИСЛОВИЕ

Данный учебник написан на основе учебного пособия «Физика полупроводников», изданного в 1971 г. При написании учебника, так же как и при издании учебного пособия, автор считал, что книга по физике полупроводников должна отвечать задаче подготовки студентов, специализирующихся в области полупроводниковой техники. Поэтому наряду с достаточно высоким теоретическим уровнем изложения, обуславливающим высокую математическую насыщенность материала, уделено большое внимание физической интерпретации явлений и результатам экспериментальных исследований на конкретных полупроводниковых материалах. При этом кроме рассмотрения фундаментальных проблем физики полупроводников проведено обсуждение ряда явлений прикладного характера, имеющих большое значение для понимания процессов, происходящих в полупроводниковых приборах.

Этими положениями и определялся отбор материала для настоящего учебника, написанного на основе курса лекций по физике полупроводников, читаемого автором для студентов специальности «Полупроводниковые и микроэлектронные приборы» в Московском ордена Ленина энергетическом институте.

Многолетний педагогический опыт автора показывает, что перегруженность материала сложными математическими выкладками отрицательно сказывается на прочном усвоении студентами физической стороны явлений.

Отнюдь не ратуя за излишнюю упрощенность при изложении вопросов, где строгость рассмотрения не страдает от отсутствия сложных математических выкладок, автор старался их избежать, приводя ссылки на оригинальные работы.

Автор выражает благодарность профессорам А. Н. Горбань и Ю. К. Шалабутову за ценные советы и замечания, сделанные ими при рецензировании рукописи. Автор признателен доценту Л. П. Павлову, взявшему на себя труд редактирования рукописи.

Все замечания и предложения читателей просьба направлять по адресу: 113114, Москва, М-114, Шлюзовая наб., 10, издательство «Энергия».

СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ

- a — постоянная решетки; боровский радиус; ускорение;
 a, b, c — полуоси эллипсоида;
 A, B, C — постоянные величины;
 B, B — магнитная индукция;
 C — емкость; теплопроводность;
 c — скорость света в вакууме;
 D — биполярный коэффициент диффузии; прозрачность потенциального барьера;
 D_n, D_p — коэффициент диффузии электронов, дырок;
 \mathcal{E}, \mathcal{E} — напряженность электрического поля;
 $\mathcal{E}_{кр}$ — критическая напряженность электрического поля;
 \mathcal{E}_m — максимальная напряженность электрического поля;
 E — энергия;
 E_c, E_v — энергии, соответствующие дну зоны проводимости и верху валентной зоны;
 E_a, E_d, E_{xc} — энергии ионизации акцептора, донора, экситона;
 E_g — ширина запрещенной зоны;
 $E_{и}$ — энергия ионизации;
 E_p — энергия фонона;
 E_T — порог внешнего фотоэффекта;
 E_t — энергия уровня захвата;
 e — заряд электрона; величина деформации;
 \mathbf{H} — оператор Гамильтона;
 \mathbf{H}, H — напряженность магнитного поля;
 h — постоянная Планка; $\hbar = h/2\pi$;
 F, F — сила;
 F — уровень Ферми (энергия Ферми);
 F_n, F_p — квазиуровень Ферми для электронов и дырок;
 f — функция Ферми—Дирака; неравновесная функция распределения;
 f_0, f_{0p} — равновесная функция распределения для электронов, дырок;
 G — скорость генерации;
 G_s — поверхностная проводимость;
 g — фактор спинового вырождения; вероятность поглощения;
 l — ток; интенсивность света;
 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ — единичные векторы;
 \mathbf{j}, j — плотность тока;
 j_s — плотность тока насыщения;
 \mathbf{k}, k — волновой вектор носителя заряда; модуль волнового вектора носителя заряда;
 k — постоянная Больцмана;
 k_n, k_p — вероятность захвата электрона, дырки;
 \mathcal{L} — длина; толщина слоя объемного заряда;
 L — длина; диффузионная длина; длина затягивания;
 L_n, L_p — длина диффузии электронов, дырок;
 l — длина свободного пробега; длина;

- l_3 — длина экранирования;
 M — масса ядра; масса атома (иона);
 m_0 — масса свободного носителя заряда;
 m^* — эффективная масса носителя заряда;
 m_n^* , m_p^* — эффективная масса электрона, дырки;
 m_{dn}^* , m_{dp}^* — эффективная масса плотности состояний электронов, дырок;
 N — плотность квантовых состояний; концентрация;
 N_a , N_d — концентрация акцепторов, доноров;
 N_c , N_v — эффективная плотность состояний в зоне проводимости, валентной зоне;
 N_t — концентрация ловушек захвата;
 N_p — плотность фононов;
 N_s — плотность поверхностных состояний;
 n — неравновесная концентрация электронов; квантовое число;
 n_a , n_d — концентрация электронов на акцепторной, донорной примеси;
 n_0 — равновесная концентрация электронов;
 n_i — собственная концентрация носителей заряда;
 \bar{n}^* — комплексный показатель преломления;
 \bar{n} — показатель преломления среды;
 P — импульс электрона; давление; поляризуемость среды; вероятность перехода;
 p , p — квазиимпульс носителя заряда;
 p — неравновесная концентрация дырок; показатель степени;
 p_a , p_d — концентрация ионизованной акцепторной, донорной примеси;
 p_0 — равновесная концентрация дырок;
 Π — коэффициент Пельтье;
 Q — квазиимпульс решетки; объемный заряд; полный заряд; количество тепла;
 Q_{ss} — поверхностный заряд;
 Q_{sp} — объемный заряд в приповерхностной области полупроводника;
 q , q — волновой вектор фонона; модуль волнового вектора фонона;
 R — радиус-вектор;
 R — сопротивление; коэффициент Холла; скорость рекомбинации; коэффициент отражения;
 R_{II}^0 , R_{II}^0 — плотность равновесного излучения, поглощения;
 r — радиус-вектор электрона;
 s — площадь; скорость поверхностной рекомбинации;
 T — температура; время; кинетическая энергия; период; пропускание;
 t — время;
 U — потенциальная энергия;
 u — смещение;
 u_v — плотность излучения;
 V — потенциал кристаллической решетки; объем кристалла; разность потенциалов;
 V_B — объем зоны Бриллюэна;
 v , v — дрейфовая скорость;
 v_0 — средняя скорость теплового движения;
 v_D — диффузионная скорость;
 v_d — скорость домена;
 $v_{зв}$ — скорость звука;
 $v_{гр}$ — групповая скорость;
 $v_{ф}$ — фазовая скорость;
 W — поток тепловой энергии;
 W — вероятность;
 ω — энергия активации;

- U — безразмерный поверхностный потенциал;
 α — термический коэффициент сопротивления; коэффициент ударной ионизации; удельная термо-э.д.с.; коэффициент поглощения;
 β — температурный коэффициент; коэффициент квазипругой силы; квантовый выход;
 γ — коэффициент рекомбинации; коэффициент пропорциональности;
- $\Delta n, \Delta p$ — концентрация избыточных электронов, дырок;
 δ — относительное изменение энергии; дельта-функция; толщина туннельного барьера;
- ϵ, ϵ_0 — диэлектрическая проницаемость материала, вакуума;
 η — приведенный уровень Ферми; импульс фотона;
 θ — температура Дебая; угол;
 ϑ — угол;
- κ_0 — магнитная проницаемость вакуума;
 κ — магнитная проницаемость вещества; теплопроводность;
- κ_e, κ_L — электронная, решеточная теплопроводность;
 λ — длина волны;
- μ, μ_d — дрейфовая подвижность носителей заряда;
 μ_D — биполярная диффузионная подвижность;
- μ_n, μ_p — подвижность электронов, дырок;
 μ_ϵ — биполярная дрейфовая подвижность;
- μ_H — Холлова подвижность носителей заряда;
 ν — частота; постоянная спада;
 ξ — уровень Ферми, отсчитанный от E_c ; химический потенциал;
 ρ — удельное сопротивление; линейная плотность; объемный заряд;
- σ — удельная проводимость; эффективное сечение;
 τ — среднее время свободного пробега; время релаксации; постоянная времени; время жизни;
- τ_n, τ_p — время жизни электронов, дырок;
 τ_μ — максвелловское время релаксации;
- τ_T — коэффициент Томсона;
- Φ — функция; угол; угол Холла; потенциал; контактная разность потенциалов;
 Φ — функция; работа выхода;
 χ — энергия сродства к электрону; работа выхода; показатель поглощения;
- Ψ — функция; поверхностный потенциал;
 ω — угловая частота;
 ω_c — циклотронная частота.

$10^{-6} \div 0,1 \text{ м} + \dots$

Глава первая
ПОЛУПРОВОДНИКИ.
ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ЭЛЕКТРОПРОВОДНОСТИ

1-1. КЛАССИФИКАЦИЯ ВЕЩЕСТВ
ПО ВЕЛИЧИНЕ УДЕЛЬНОЙ ПРОВОДИМОСТИ.
ПОЛУПРОВОДНИКИ

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

Из приведенных примеров видно, что при переходе от одного класса вещества к другому значения удельного сопротивления перекрываются. Поэтому удельное сопротивление не может служить в качестве однозначного критерия для классификации веществ. Однако при снятии температурных зависимостей удельного сопротивления различие между металлами и полупроводниками часто проявляется достаточно четко. Графики температурной зависимости удельного сопротивления некоторых металлов и полупроводников представлены на рис. 1-1. У химически чистых металлов удельное сопротивление с ростом температуры увеличивается пропорционально абсолютной температуре T , т. е.

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

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

Характерным для температурной зависимости удельной проводимости металлов, как следует из рис. 1-2, а, 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 .

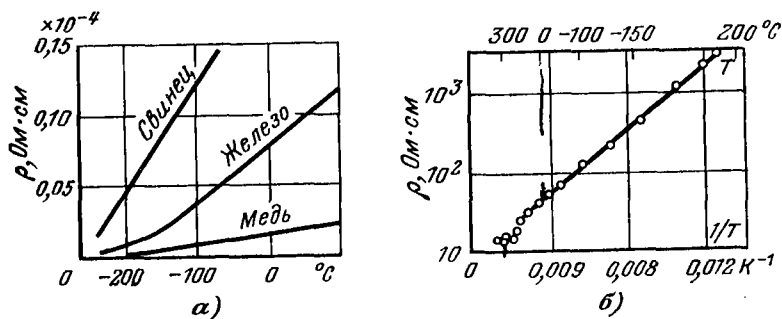


Рис. 1-1. Изменение удельного сопротивления чистых металлов (а) и кремния (б) в зависимости от температуры.

Для полупроводников характер температурной зависимости удельного сопротивления и проводимости иной. Для некоторого интервала температур эти зависимости имеют вид:

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

$$\sigma = \sigma_0 e^{-\beta/T}, \quad (1-3)$$

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

Такие зависимости удельного сопротивления и проводимости от температуры имеют так называемые невырожденные полупроводники. Для них, как видно из графика температурной зависимости удельной проводимости, изображенного на рис. 1-2, б, характерно наличие положительного температурного коэффициента удельной проводимости, т. е.

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

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

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

Ответить на этот вопрос можно, если проследить, как изменяется проводимость вещества при понижении температуры. Как следует из рис. 1-2, а, с понижением температуры удельная проводимость металлов растет. При температуре абсолютного нуля металлы имеют конечное значение удельной проводимости, а у ряда металлов и их сплавов наступает сверхпроводящее состояние, характеризующееся сильным возрастанием удельной проводимости.

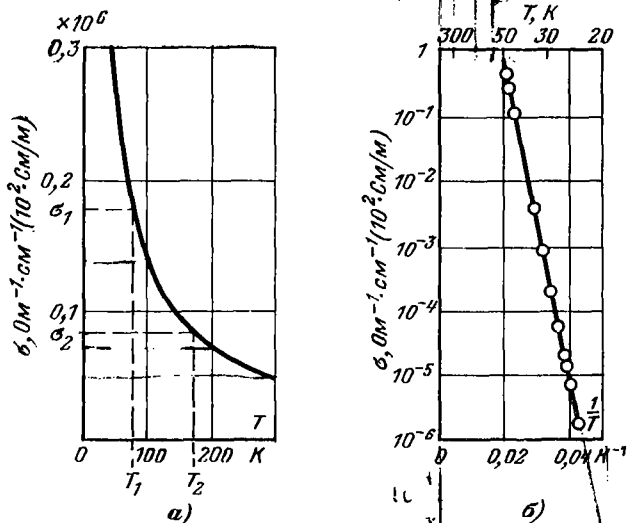


Рис. 1-2. Изменение удельной проводимости свинца (а) и кремния (б) в зависимости от температуры.

Такие изменения удельной проводимости металлов с понижением температуры возможны лишь потому, что независимо от температуры в металле всегда имеются свободные носители заряда — электроны. У полупроводников, наоборот, удельная проводимость уменьшается при понижении температуры (рис. 1-2, б), а по мере приближения температуры к абсолютному нулю полупроводники по своим свойствам приближаются к диэлектрикам. Из этого следует, что в полупроводнике свободные носители заряда возникают при подведении к нему тепловой энергии. Эти носители заряда называются **тепловыми** или **равновесными**. Опыт показывает, что появление свободных носителей заряда в полупроводнике имеет место также при освещении, облучении ядерными частицами, при наложении на полупроводник электрического поля, при изменении внешнего давления. Возникающие в этих случаях носители заряда называются **неравновесными**. Процесс образования как равновесных, так и неравновесных носителей

заряда очень сильно зависит от структуры полупроводникового вещества и наличия в нем примесей.

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

Согласно этому определению между полупроводниками и диэлектриками не существует принципиального качественного различия, ибо они обладают проводимостью только вследствие теплового возбуждения носителей заряда. Более различны по своей природе металлы и полупроводники. У металлов проводимость слабо зависит от присутствия примеси, внешних условий и при любой температуре концентрация свободных электронов остается постоянной и составляет величину порядка 10^{22} см⁻³. Различие между металлами и полупроводниками проявляется и в том, что при прохождении тока через цепь, состоящую из двух полупроводников, сила тока нелинейно зависит от приложенной к цепи разности потенциалов.

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

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

К электронным полупроводникам относится огромное количество самых различных веществ. Полупроводниками являются как простые вещества: бор В, углерод С, кремний Si, фосфор Р, сера α -S, германий Ge, мышьяк As, серое олово α -Sn, сурьма β -Sb, селен Se (красный), теллур Те, йод J, так и многие сложные химические соединения. В качестве примера в табл. 1-1 приведены некоторые соединения элементов периодической системы Менделеева (рис. 1-3), обладающие полупроводниковыми свойствами. В таблице указаны кристаллическая структура, постоянная решетки и другие характеристики этих важнейших полупроводников.

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

Особенности электрофизических свойств полупроводников обусловили их широкое применение для создания самых различных

Кристаллическая структура важнейших полупроводников

Полупроводники	Элемент или соединения	Название	Кристаллическая структура	Постоянная решетки при 300 К, Å	Ширина запрещенной зоны при 0 К, эВ	Диэлектрическая проницаемость
Элементарные полупроводники	C	Алмаз	Алмаза	3,56679	5,51	5,5
	Ge	Германий		5,65748	0,89	16
	Si	Кремний		5,43086	1,16	11,8
	Sn	Серое олово		6,4892	> 0,08	
IV—IV	SiC	Карбид кремния	Цинковой обманки (кубическая)	4,358	3,1	10
III—V	AlSb	Антимонид алюминия	Цинковой обманки	6,1355	1,75	11
	GaSb	Антимонид галлия	» »	6,0955	0,80	15
	GaAs	Арсенид галлия	» »	5,6534	1,52	10,9
	GaP	Фосфид галлия	» »	5,4505	2,40	10
	InSb	Антимонид индия	» »	6,4788	0,26	17
	InAs	Арсенид индия	» »	6,0585	0,46	14,5
	InP	Фосфид индия	» »	5,8688	1,34	14
II—VI	CdS	Сульфид кадмия	Цинковой обманки Вюрцита (гексагональная)	5,832	2,487	10
	CdS	»		a - 4,16 c - 6,756	2,56	
	CdSe	Селенид кадмия	Цинковой обманки Вюрцита	6,05	1,85	10
	ZnO	Оксид цинка		a - 3,21 c - 5,19	3,2	9
	ZnS	Сульфид цинка	Цинковой обманки Вюрцита	5,42	3,7	8
	ZnS	»		a - 3,82 c = 6,26	3,91	
IV—VI	PbS	Сульфид свинца	Кубическая	5,935	0,34	17
	PbTe	Теллурид свинца		»	6,460	0,24

приборов. Полупроводниковые приборы используются для выпрямления тока (диоды), усиления и генерации колебаний (транзисторы), преобразования тепловой энергии и энергии излучения в электрическую (термоэлементы и фотоэлементы), для преобразования электрической энергии в световую (фосфоры, светодиоды, лазеры). Полупроводниковые приборы применяются также в качестве «датчиков» для измерения температуры (термисторы), регистрации светового и корпускулярного излучения (фоторезисторы и дозиметры), для измерения давления (тензодатчики) и магнитных полей (датчики Холла).

В основе работы каждого прибора лежат определенные физические процессы и явления. Другими словами, физика полупроводников представляет собой теоретическую основу полупроводниковой электроники. Разработка полупроводниковых материалов, создание на их основе новых видов приборов, научные исследова-

ПЕРИОДИЧЕСКАЯ СИСТЕМА

ПЕРИОДЫ	РЯДЫ	ГРУППЫ				
		I	II	III	IV	V
1	I	H ¹ ВОДОРОД 1,00797 ₁				
2	II	Li ³ ЛИТИЙ 6,9341 ₁	Be ⁴ БЕРИЛЛИЙ 9,01218 ₂	B ⁵ БОР 10,81 ₂	C ⁶ УГЛЕРОД 12,011 ₂	N ⁷ АЗОТ 14,0067 ₅
3	III	Na ¹¹ НАТРИЙ 22,98977 ₂	Mg ¹² МАГНИЙ 24,305 ₂	Al ¹³ АЛЮМИНИЙ 26,98154 ₂	Si ¹⁴ КРЕМНИЙ 28,086 ₂	P ¹⁵ ФОСФОР 30,97376 ₅
4	IV	K ¹⁹ КАЛИЙ 39,09 ₂	Ca ²⁰ КАЛЬЦИЙ 40,08 ₂	Sc ²¹ СКАНДИЙ 44,9559 ₂	Ti ²² ТИТАН 47,90 ₂	V ²³ ВАНАДИЙ 50,941 ₂
	V	Cu ²⁹ МЕДЬ 63,546 ₂	Zn ³⁰ ЦИНК 65,38 ₂	Ga ³¹ ГАЛЛИЙ 69,72 ₂	Ge ³² германий 72,59 ₂	As ³³ мышьяк 74,9216 ₂
5	VI	Rb ³⁷ РУБИДИЙ 85,4678 ₂	Sr ³⁸ СТРОНЦИЙ 87,62 ₂	Y ³⁹ ИТРИЙ 88,9059 ₂	Zr ⁴⁰ ЦИРКОНИЙ 91,22 ₂	Nb ⁴¹ НИОБИЙ 92,9064 ₂
	VII	Ag ⁴⁷ СЕРЕБРО 107,868 ₂	Cd ⁴⁸ КАДМИЙ 112,40 ₂	In ⁴⁹ ИНДИЙ 114,82 ₂	Sn ⁵⁰ ОЛОВО 118,69 ₂	Sb ⁵¹ СУРЬМА 121,75 ₂
6	VIII	Cs ⁵⁵ ЦЕЗИЙ 132,9054 ₂	Ba ⁵⁶ БАРИЙ 137,34 ₂	La * ЛАНТАН 138,9055 ₂	Hf ГАФНИЙ 178,49 ₂	Ta ТАНТАЛ 180,9477 ₂
	IX	Au ⁷⁹ ЗОЛОТО 196,9665 ₂	Hg ⁸⁰ РУТУТЬ 200,59 ₂	Tl ⁸¹ ТАЛЛИЙ 204,37 ₂	Pb ⁸² СВИНЕЦ 207,2 ₂	Bi ⁸³ ВИСМУТ 208,9804 ₂
7	X	Fr ⁸⁷ франций [223]	Ra ⁸⁸ РАДИЙ 226,0254 ₂	Ac ** АКТИНИЙ [227]	Ku КУРЧАТОВИЙ [261]	105

* Л А Н Т А Н

58 ₂ Ce ЦЕРИЙ 140,12	59 ₂ Pr ПРАЗЕОДИМ 140,9077	60 ₂ Nd НЕОДИМ 144,24	61 ₂ Pm ПРОМЕТИЙ [147]	62 ₂ Sm САМАРИЙ 150,4	63 ₂ Eu ЕВРОПИЙ 151,96	64 ₂ Gd ГАДОЛИНИЙ 157,25
--	--	---	--	---	--	--

** А К Т И

90 ₂ Th ТОРИЙ 232,038	91 ₂ Pa ПРОТАКТИНИЙ 231,0359	92 ₂ U УРАН 238,029	93 ₂ Np НЕПТУНИЙ 237,0482	94 ₂ Pu ПЛУТОНИЙ [244]	95 ₂ Am АМЕРИЦИЙ [243]	96 ₂ Cm КЮРИЙ [247]
---	--	---	---	--	--	---

Рис. 1-3. Периодическая

ЭЛЕМЕНТОВ Д.И. МЕНДЕЛЕЕВА

Э Л Е М Е Н Т О В										ЭЛЕКТРОННЫЕ СЛОИ			
VI		VII		VIII						0			
		(H)								He ² ГЕЛИЙ 40026		1	K
8 O 6 КИСЛОРОД 2 15,9994		9 F 7 ФТОР 2 18,9984								Ne ¹⁰ НЕОН 20,179		8	L K
16 S 6 СЕРА 8 32,06		17 Cl 7 ХЛОР 8 35,453								Ar ¹⁸ АРГОН 39,948		8	M L K
Cr ²⁴ 1 ХРОМ 8 51,996		Mn ²⁵ 2 МАРГАНЕЦ 13 54,938		Fe ²⁶ 2 ЖЕЛЕЗО 14 55,847		Co ²⁷ 2 КОБАЛЬТ 15 58,9332		Ni ²⁸ 2 НИКЕЛЬ 16 58,7				8	N M L K
34 Se 6 СЕЛЕН 18 78,96		35 Br 7 БРОМ 18 79,904								Kr ³⁶ 8 КРИПТОН 18 83,80		8	N M L K
Mo ⁴² 13 МОЛИБДЕН 18 95,94		Tc ⁴³ 2 ТЕХНЕЦИЙ 13 98,9069		Ru ⁴⁴ 1 РУТЕНИЙ 15 101,07		Rh ⁴⁵ 1 РОДИЙ 16 102,9055		Pd ⁴⁶ 0 ПАЛЛАДИЙ 18 106,4				8	O N M L K
52 Te 6 ТЕЛЛУР 18 127,60		53 I 7 ИОД 18 126,9045								Xe ⁵⁴ 8 КСЕНОН 18 131,30		8	O N M L K
74 W 12 ВОЛЬФРАМ 18 183,85		75 Re 13 РЕНИЙ 32 186,207		76 Os 14 ОСМИЙ 32 190,2		77 Ir 15 ИРИДИЙ 32 192,22		78 Pt 1 ПЛАТИНА 17 195,09				8	P O N M L K
84 Po 6 ПОЛОНИЙ 18 [210]		85 At 7 АСТАТ 18 [210]								86 Rn 8 РАДОН 18 [222]		8	P O N M L K
												8	P O N M L K

О И Д Ы 58-71

65 Tь 2 ТЕРБИЙ 8 158,9254		66 Dy 2 ДИСПРОЗИЙ 8 162,50		67 Ho 2 ГОЛЬМИЙ 8 164,9304		68 Er 2 ЭРБИЙ 8 167,26		69 Tu 2 ТУЛИЙ 8 168,9342		70 Yь 2 ИТТЕРБИЙ 8 173,04		71 Lu 2 ЛЮТЕЦИЙ 8 174,97		9	P O N M L K
---------------------------------	--	----------------------------------	--	----------------------------------	--	------------------------------	--	--------------------------------	--	---------------------------------	--	--------------------------------	--	---	----------------------------

Н О И Д Ы 90-103

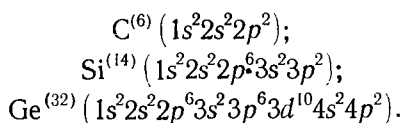
97 Bк 2 БЕРКЛИЙ 8 [247]		98 Cf 2 КАЛИФОРНИЙ 8 [252]		99 Es 2 ЭЙНШТЕЙНИЙ 8 [254]		100 Fm 2 ФЕРМИЙ 8 [257]		101 Md 2 МЕНДЕЛЕВИЙ 8 [258]		(No) 2 (НОБЕЛИЙ) 8 [255]		(Lr) 2 (ЛОУРЕНСИЙ) 8 [256]		9	P O N M L K
-------------------------------	--	----------------------------------	--	----------------------------------	--	-------------------------------	--	-----------------------------------	--	--------------------------------	--	----------------------------------	--	---	----------------------------

система элементов Менделеева.

ния в этой области, постановка высококачественного производства полупроводниковых приборов и интегральных схем, разработка новых устройств для разнообразных условий применения — все это возможно при условии серьезных знаний в области физики полупроводников.

1-2. МОДЕЛЬНЫЕ ПРЕДСТАВЛЕНИЯ О МЕХАНИЗМЕ ЭЛЕКТРОПРОВОДНОСТИ СОБСТВЕННЫХ ПОЛУПРОВОДНИКОВ

Рассмотрим механизм электропроводности полупроводниковых веществ на примере элементарных полупроводников. Кремний и германий находятся в одной подгруппе периодической системы Менделеева с углеродом. Электроны распределены у них по состояниям следующим образом:



Внешняя электронная оболочка у этих атомов заполнена частично, она содержит четыре электрона.

При образовании кристалла, например кремния, четыре валентных электрона каждого атома из состояния $3s^2 3p^2$ переходят

в гибридное sp^3 -состояние с неспаренными спинами и образуют четыре пространственно-эквивалентные связи. В результате каждый атом окружен четырьмя ближайшими соседями и находится в центре тетраэдра. Возникает так называемый алмазоподобный тип кристаллической решетки (рис. 1-4), которая является кубической. В ней четыре внешних электрона каждого атома участвуют в образовании ковалентных связей (по два электрона в каждой). Эти связи на

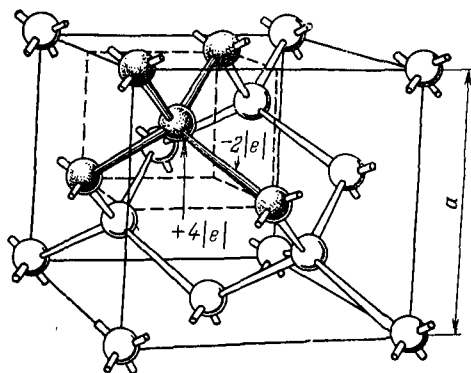


Рис. 1-4. Кристаллическая решетка типа алмаза
 a — постоянная решетки.

рис. 1-4 представлены в виде двух направляющих, соединяющих два ближайших атома. На рис. 1-5, a дано двумерное изображение связей в решетке кремния. Здесь в узле решетки находится ион кремния с зарядом $+4$, которому принадлежат четыре валентных электрона. Валентные электроны, обеспечивающие

ковалентную связь, на рис. 1-5, *а* представлены в виде черных точек.

В идеальном полупроводнике, изображенном на рис. 1-5, *а*, все электроны связаны. Если поместить такой полупроводник в электрическое поле, то электрический ток не может возникнуть, так как все связи в решетке заполнены и свободных носителей заряда нет.

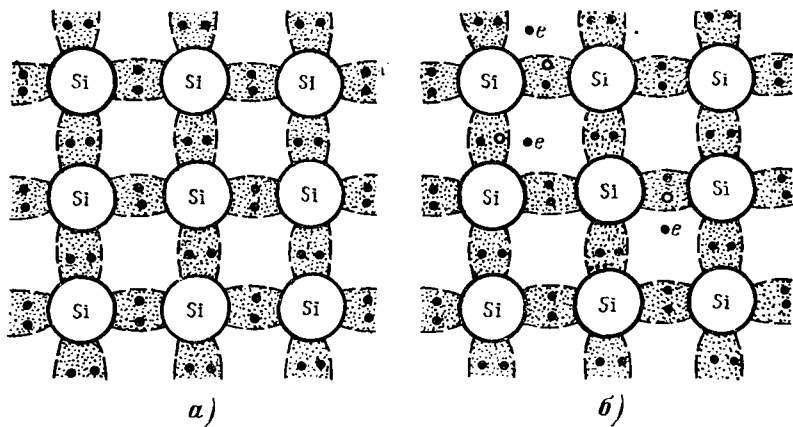


Рис. 1-5. Двумерное представление расположения связей в решетке кремния (собственный полупроводник).

Допустим, что под воздействием каких-либо возмущений, например тепловой энергии, произошел разрыв валентной связи и электрон стал свободным. Процесс превращения связанного электрона в свободный электрон носит название генерации. На месте ушедшего электрона образуется незавершенная связь, которая будет иметь избыточный положительный заряд, поскольку он теперь не скомпенсирован зарядом электрона. *Вакантное место в валентной связи получило название дырки.* В целом кристалл остается электронейтральным, так как каждому образовавшемуся положительному заряду в связи — дырке соответствует свободный электрон. На рис. 1-5, *б* свободные электроны и дырки изображены соответственно черными и светлыми кружочками. Если свободный электрон подойдет к тому атому, от которого он был оторван, то он может соединиться с атомом. Процесс превращения свободного электрона в связанный электрон носит название рекомбинации.

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

з следовательно, к росту концентрации свободных электронов и дырок в полупроводнике.

Свободные электроны за счет тепловой энергии перемещаются по кристаллу полупроводника. Но в реальном веществе идеальность кристаллической структуры всегда нарушена присутствием в нем разных дефектов. Такими дефектами являются тепловые колебания атомов кристалла, разные примеси, дислокации. Поэтому свободный электрон, перемещаясь по кристаллу, будет сталкиваться с дефектами кристаллической решетки, в результате чего меняется направление его движения. В силу этого тепловое движение свободного электрона является беспорядочным, как это изображено на рис. 1-6, а.

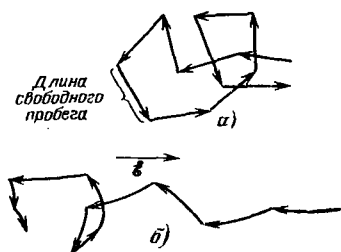


Рис. 1-6. Схема движения свободного электрона за счет тепловой энергии (а) и во внешнем электрическом поле (б).

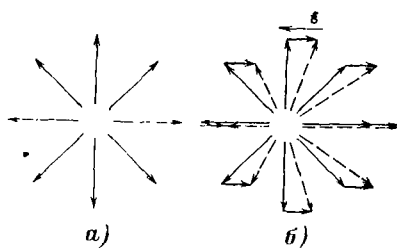


Рис. 1-7. Схематическое изображение скоростей электронов проводимости при отсутствии (а) и наличии (б) электрического поля.

Вакантное место в валентной связи—дырка может быть заполнена электроном, перешедшим за счет тепловой энергии с соседней насыщенной связи. При таком переходе от атома к атому дырка также будет совершать хаотическое движение.

Таким образом, тепловое движение свободных носителей заряда является беспорядочным. При этом каждый из носителей заряда описывает сложную траекторию движения (рис. 1-6, а). *Расстояние, проходимое свободным носителем заряда 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 называются временем свободного пробега и средним временем свободного пробега.*

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

$$l = v_0 \tau, \quad (1-4)$$

где v_0 — средняя скорость теплового движения свободного носителя заряда (среднее значение модуля скорости). В полупроводниках при комнатной температуре скорость теплового движения электронов составляет величину порядка 10^7 см/с.

Поскольку тепловое движение свободных носителей заряда — электронов и дырок — имеет хаотический характер, то их средняя скорость, рассматриваемая как векторная величина, равна нулю. Это означает, что каждому свободному носителю заряда можно сопоставить соответствующий по знаку носитель заряда со скоростью, вектор которой направлен в обратную сторону. Схематически это представлено на рис. 1-7, а. Следовательно, вследствие беспорядочного теплового движения количество свободных электронов и дырок, движущихся в любом направлении, в среднем равно числу электронов и дырок, движущихся в противоположном направлении. Поэтому в отсутствие внешнего электрического поля суммарный заряд, переносимый свободными электронами и дырками в любом направлении, равен нулю и беспорядочное тепловое движение носителей заряда не дает тока.

Поместим собственный полупроводник в электрическое поле. Под воздействием поля свободные электроны полупроводника будут ускоряться и приобретут скорость, направленную против поля. Благодаря этому у электронов, движение которых за счет тепловой энергии происходило против направления поля, скорость увеличится, а у электронов, движущихся по полю, уменьшится. В результате вся совокупность свободных электронов получает некоторую скорость направленного движения (рис. 1-7, б). Изменение скорости движения электронов скажется на их кинетической энергии. При столкновении с дефектами кристаллической решетки электроны полностью передают решетке приобретенную в поле скорость и энергию, вследствие чего они придут в тепловое равновесие с решеткой. После столкновения электроны, совершая хаотическое тепловое движение, опять будут иметь направленное движение во внешнем электрическом поле.

Таким образом, фактическое движение электрона в кристалле складывается из беспорядочного теплового и упорядоченного движения, вызванного действием внешнего электрического поля (рис. 1-6, б). В результате этого происходит медленное перемещение всей совокупности свободных электронов с некоторой средней скоростью в направлении, противоположном направлению внешнего электрического поля. Направленное движение совокупности свободных носителей заряда в электрическом поле носит название дрейфа, а скорость их направленного движения называется дрейфовой скоростью. На рис. 1-7, б горизонтальными линиями изображены векторы дрейфовой скорости, а пунктиром — векторы суммарной скорости теплового движения и дрейфа. В этом случае средняя скорость движения свободных электронов уже не равна нулю и через полупроводник в направлении электрического поля потечет ток, обусловленный свободными электронами.

Электроны насыщенных связей при переходе на вакантное место в связи под действием внешнего электрического поля будут перемещаться против направления поля. Тем самым вакантное место в валентной связи — дырка будет также перемещаться, но

по направлению внешнего электрического поля, что равносильно перемещению по полю положительного заряда. Механизм проводимости, обусловленный движением связанных электронов по вакантным связям, получил название дырочной проводимости.

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

Электропроводность собственного полупроводника можно объяснить, если исходить из энергетических представлений. Электроны

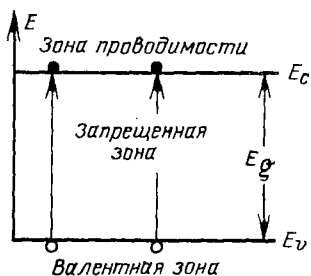


Рис. 1-8. Схематическое изображение энергетических зон собственного полупроводника.

в изолированном атоме, как показывает квантовая теория, обладают дискретными значениями энергии. При этом на каждом энергетическом уровне согласно принципу Паули может находиться не более двух электронов, которые должны иметь противоположно направленные спины. При образовании полупроводникового вещества, т. е. при сближении атомов на расстоянии примерно 10^{-8} см, валентные электроны будут двигаться в сильном электрическом поле соседних атомов. В результате действия этого поля энергетический уровень валентных электронов расщепляется в зону. Эта зона носит название валентной зоны. Из вышележащего уровня возбуждения атома образуется зона, которая называется свободной зоной, или зоной проводимости. Каждая из этих зон занимает определенную область энергии. Зоны дозванных энергий отделены друг от друга интервалом, называемым запрещенной зоной, или энергетической щелью. На рис. 1-8, на котором изображена энергетическая диаграмма собственного полупроводника, E_c соответствует минимальной энергии, которую могут иметь свободные электроны в кристалле, а величина E_v — максимальная энергия электронов валентной зоны.

Пусть в валентной зоне собственного полупроводника все возможные квантовые состояния заняты электронами. Но мгновенная плотность электрического тока \mathbf{j} , связанного с движением одного электрона, пропорциональна его скорости \mathbf{v} и не совпадает с ней по направлению; для полупроводника, имеющего объем Ω ,

$$\mathbf{j} = -\frac{e\mathbf{v}}{\Omega}. \quad (1-5)$$

Поскольку в валентной зоне все возможные квантовые состояния заняты электронами, то результирующее мгновенное

значение плотности электрического тока для всей системы электронов

$$j = \frac{1}{\Omega} \sum_s v_s$$

$$j = -\frac{e}{\Omega} \sum_s v_s = 0, \quad (1-6)$$

(по всем состояниям в зоне)

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

В собственном полупроводнике при разрыве парноэлектронной связи, например за счет тепловой энергии, появляются свободный электрон и вакантное место в валентной связи. В схеме энергетических зон, изображенной на рис. 1-8, это равнозначно переходу электрона из валентной зоны в зону проводимости.

В этом случае валентная зона полностью заполнена электронами, за исключением одного единственного квантового состояния. Соответствующую скорость вакантного места в связи обозначим через v_i . Тогда суммарный ток всех электронов валентной зоны запишется в виде

$$j = -\frac{e}{\Omega} \sum_s v_s = -\frac{e}{\Omega} \sum_s v_s + \frac{e}{\Omega} v_i. \quad (1-7)$$

Но сумма токов в правой части (1-7) по всем состояниям в зоне равна нулю. Поэтому

$$j = \frac{e}{\Omega} v_i. \quad (1-8)$$

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

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

Ширина запрещенной зоны E_g определяется природой химической связи, а также зависит от сорта атомов, образующих кристаллическую решетку. В кристаллах, состоящих из элементов IV группы периодической системы Менделеева, как видно из данных табл. 1-1, она убывает по мере роста постоянной решетки. Решетка алмаза более плотная, а следовательно, более прочная, и на разрыв парноэлектронной связи требуется больше энергии, чем для более рыхлой решетки.

1-3. МОДЕЛЬНЫЕ ПРЕДСТАВЛЕНИЯ О МЕХАНИЗМЕ ЭЛЕКТРОПРОВОДНОСТИ ПРИМЕСНЫХ ПОЛУПРОВОДНИКОВ

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

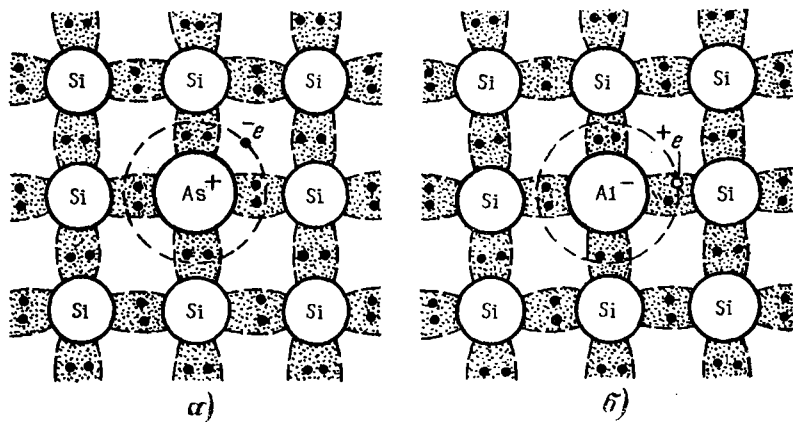


Рис. 1-9. Схематическое изображение кристаллической решетки допурного (а) и акцепторного (б) полупроводников.

У атома мышьяка пять валентных электронов расположены в $4s$ - и $4p$ -состояниях: $As^{(33)}(1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^3)$. В решетке кремния четыре валентных электрона атома мышьяка вместе с четырьмя электронами ближайших атомов кремния участвуют в образовании ковалентной связи, как это схематически представлено на рис. 1-9, а. Пятый электрон мышьяка не может принять участие в образовании связи, поскольку все связи завершены. Он слабо связан с атомом мышьяка, так как он испытывает воздействие со стороны окружающих атомов кремния. При низких температурах пятый электрон локализован около атомов мышьяка, но при повышенных температурах он будет оторван от примеси и может свободно перемещаться по кристаллу. В этом случае также соблюдается электронейтральность кристалла, так как атом мышьяка, отдавший пятый электрон, будет теперь являться положительным ионом.

Наряду с ионизацией примеси может происходить и ионизация атомов основного вещества. Но в области температур ниже той, при которой имеет место значительная собственная электропроводность, количество электронов, оторванных от примеси, будет значительно больше количества электронов и дырок, образовавшихся в результате разрыва валентных связей. В силу этого *доминирующую роль*

в проводимости полупроводника будут играть электроны, поэтому они называются основными носителями заряда, а дырки — неосновными носителями заряда. Такой полупроводник называется электронным или *n*-типа, а примесь, дающая 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 возникновения требуется значительно меньшая энергия, чем для разрыва валентных связей кремния, энергетический уровень донорной примеси E_d должен располагаться в запрещенной зоне на небольшой глубине под дном зоны проводимости (рис. 1-10, *а*).

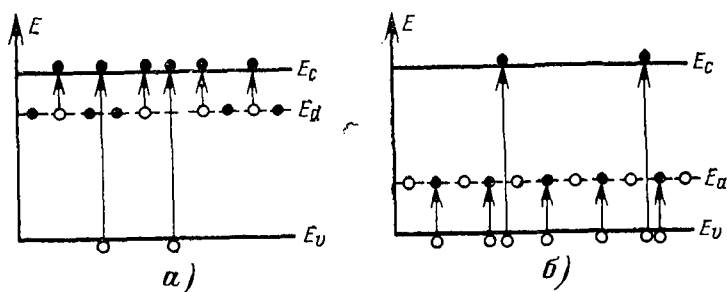


Рис. 1-10. Энергетическая диаграмма донорного (*а*) и акцепторного (*б*) полупроводников.

Пусть в качестве примеси в кристаллическую решетку полупроводника с ковалентной связью внесены атомы элемента третьей группы периодической системы Менделеева, например алюминий в решетке кремния. Поскольку высшая валентность алюминия равна трем: $Al^{(13)} (1s^2 2s^2 2p^6 3s^2 3p)$, 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 из атома кремния. При этом образуются отрицательный ион алюминия и свободная дырка, перемещающаяся по связям кремния и, следовательно, принимающая участие в проводимости полупроводника. *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 электропроводность кристалла будет дырочной. В таком полупроводнике основными носителями заряда будут дырки, а электроны — неосновными носителями заряда. Полупроводник с акцепторной примесью носит название дырочного или *p*-типа.

На энергетической диаграмме, изображенной на рис. 1-10, б, акцепторная примесь имеет в запрещенной зоне энергетический уровень E_a , расположенный на небольшом расстоянии над потолком валентной зоны. При ионизации акцепторной примеси 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.

1-4. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ЭЛЕКТРОПРОВОДНОСТИ ПОЛУПРОВОДНИКОВ

Проведем подсчет плотности тока для донорного полупроводника, электроны проводимости которого будем рассматривать как идеальные частицы, не имеющие собственного объема и не взаимодействующие друг с другом. Так как по классической теории радиус электрона $r_0 \approx 10^{-13}$ см, то при концентрации их $n \approx 10^{22}$ см⁻³ объем электронов составляет $\frac{4}{3} \pi r_0^3 n \approx 10^{-17}$ от объема вещества, и первое предположение вполне оправдано. Законы квантовой механики, как это будет показано во второй главе, в известном приближении позволяют рассматривать электроны как не взаимодействующие частицы.

Пусть концентрация электронов проводимости n (количество свободных электронов в 1 см³ полупроводника), а скорость их дрейфового движения v . Поскольку плотность тока есть заряд, проходящий в единицу времени через единичное сечение, то

$$\mathbf{j} = -env. \quad (1-9)$$

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

Вероятность столкновений n частиц за время dt соответственно будет равна $n dt/\tau$. Следовательно, за время dt число носителей заряда, движущихся в данном направлении, уменьшается в результате рассеяния на величину

$$-dn = n \frac{dt}{\tau}. \quad (1-10)$$

Решая уравнение (1-10) относительно n , получаем количество электронов, не испытавших к моменту времени t соударения:

$$n(t) = n_0 e^{-t/\tau}, \quad (1-11)$$

где $n = n_0$ при $t = 0$.

Из соотношения (1-11) следует, что количество электронов, движущихся в данном направлении, в результате столкновений с неоднородностями кристаллической решетки уменьшается по экспоненциальному закону с постоянной времени τ .

Для определения τ поступим следующим образом. Пусть ось x совпадает по направлению с полем. Предположим, что все электроны в момент времени $t = 0$ испытали столкновения. После этого в направлении x до следующего столкновения в результате дрейфа каждый электрон пройдет расстояние $x_1, x_2, x_3 \dots$ соответственно в течение времени свободного пробега $t_1, t_2, t_3 \dots$. Тогда все электроны, двигаясь в поле, за время $T = t_1 + t_2 + t_3 + \dots$ пройдут суммарное расстояние $X = x_1 + x_2 + x_3 + \dots$. Скорость дрейфа есть скорость направленного движения электронов, она равна:

$$v = \frac{X}{T}. \quad (1-12)$$

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

$$T = n_0 \langle t \rangle. \quad (1-13)$$

С другой стороны, есть определенная вероятность того, что среди электронов имеются такие, которые обладают одним и тем же временем свободного пробега t . Это электроны, испытавшие соударения в момент времени от t до $t + dt$. Их число согласно (1-10) составит $n dt/\tau$, и в суммарное время T они внесут время, равное $tn dt/\tau$. Тогда, интегрируя это выражение по всем временам свободного пробега, которые вследствие случайного характера столкновений могут принимать значения от 0 до ∞ , будем иметь:

$$T = \int_0^{\infty} \frac{tn dt}{\tau}. \quad (1-14)$$

Воспользовавшись выражениями (1-13), (1-14) и (1-11), получим:

$$\langle t \rangle = \frac{1}{n_0} \int_0^{\infty} \frac{tn dt}{\tau} = \frac{1}{n_0} \int_0^{\infty} \tau n_0 \frac{t}{\tau} e^{-\frac{t}{\tau}} \frac{dt}{\tau} = \tau. \quad (1-15)$$

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

Электрическое поле напряженности \mathcal{E} сообщит электрону с массой m ускорение, равное

$$a = \frac{e\mathcal{E}}{m}. \quad (1-16)$$

За время свободного пробега t электрон приобретает скорость

$$v_t = at = \frac{et}{m} \mathcal{E} \quad (1-17)$$

и пройдет путь

$$x = \frac{e\mathcal{E}}{2m} t^2. \quad (1-18)$$

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

$$\begin{aligned} X &= x_1 + x_2 + \dots = \int_0^{\infty} xn \frac{dt}{\tau} = \\ &= \frac{en_0\tau^2\mathcal{E}}{2m} \int_0^{\infty} \left(\frac{t}{\tau}\right)^2 e^{-t/\tau} \frac{dt}{\tau} = \frac{en_0\tau^3\mathcal{E}}{m}. \end{aligned} \quad (1-19)$$

Следовательно, *скорость дрейфа электронов* равна:

$$v = \frac{X}{T} = \frac{e\tau\mathcal{E}}{m}, \quad (1-20)$$

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

Величина, связывающая дрейфовую скорость носителей заряда с напряженностью электрического поля, называется подвижностью носителей заряда. Обозначим ее буквой μ , тогда

$$v = \mu\mathcal{E}, \quad (1-21)$$

откуда

$$\mu = \frac{v}{\mathcal{E}} = \frac{e\tau}{m}, \quad (1-22)$$

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

С учетом равенства (1-22) выражение (1-9) для плотности тока примет вид:

$$\mathbf{j} = -en\mathbf{v} = en\mu\mathcal{E}, \quad (1-23)$$

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

Удельная проводимость на основании закона Ома может быть выражена при помощи (1-23) как

$$\sigma = \frac{j}{\mathcal{E}} = en\mu. \quad (1-24)$$

С учетом соотношения (1-22) удельная проводимость будет равна:

$$\sigma = \frac{e^2n\tau}{m}. \quad (1-25)$$

Если выразить τ из (1-4), то соотношение (1-25) можно записать в виде

$$\sigma = \frac{e^2 n l}{m v_0} . \quad (1-26)$$

В действительности же при определении τ надо учитывать полную скорость электрона, которая определяется тепловой скоростью v_0 и скоростью дрейфа v , т. е.

$$\tau = \frac{l}{v_0 + v} . \quad (1-27)$$

Поскольку дрейфовая скорость электрона зависит от напряженности электрического поля, то его полная скорость ($v_0 + v$) и, следовательно, время свободного пробега τ являются функцией энергии частицы. Поэтому для нахождения дрейфовой скорости необходимо усреднять время свободного пробега с учетом функции распределения электронов по энергии. Это будет рассмотрено в гл. 6.

Глава вторая

ОСНОВЫ ЗОННОЙ ТЕОРИИ ПОЛУПРОВОДНИКОВ

2-1. УРАВНЕНИЕ ШРЕДИНГЕРА ДЛЯ КРИСТАЛЛА

Твердое тело, как известно, состоит из атомов, т. е. из ядер атомов и электронов. Ядра атомов образуют кристаллическую решетку, которая обладает свойством пространственной периодичности. При наложении внешнего электрического поля решетка практически не деформируется, хотя ядра атомов и заряжены. Это происходит потому, что силы, удерживающие ядра атомов в узлах кристаллической решетки, обычно значительно больше тех сил, которые создаются внешними электрическими полями. Те из электронов, которые не находятся близко к ядру атома и поэтому к нему не очень сильно притягиваются, могут передвигаться по твердому телу, создавая электрический ток. Однако при количественном описании этого явления возникают серьезные трудности. Они связаны с тем, что электроны являются заряженными частицами и при своем движении по твердому телу встречаются с другими электронами. Но так как между электронами действуют электрические силы отталкивания, то движение электрона оказывается зависимым от движения окружающих его электронов. Иными словами, в рассматриваемом случае надо решать не одноэлектронную, а многоэлектронную задачу. Поэтому для определения стационарных состояний и энергетического спектра совокупности

большого числа атомных ядер и электронов в кристалле пужно решить уравнение Шредингера:

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

где \hat{H} --- гамильтониан кристалла; Φ --- собственная волновая функция гамильтониана; \mathcal{E} --- энергия кристалла.

Значения волновой функции кристалла зависят от координат всех электронов \mathbf{r}_i и всех атомных ядер \mathbf{R}_α :

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

Оператор Гамильтона включает в себя следующие операторы:

1) оператор кинетической энергии электронов

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

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

2) оператор кинетической энергии ядер

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

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

3) потенциальную энергию попарного взаимодействия электронов

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

4) потенциальную энергию попарного взаимодействия ядер

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

5) потенциальную энергию взаимодействия электронов с ядрами

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

С учетом этих составляющих гамильтониана уравнение Шредингера примет вид:

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

Число независимых переменных в уравнении (2-2) определяется полным числом частиц в кристалле, которые в 1 см^3 вещества составляют величину порядка 10^{23} . Такая задача в настоящее время не может быть решена в общем виде. 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 этих частиц одного порядка. Так как масса ядра намного больше массы электрона, т. е. $M_\alpha \gg m_0$, то скорости движения электронов намного превосходят скорости ядер (приблизительно на два порядка). При каждом изменении положения атомных ядер практически мгновенно устанавливается пространственное распределение электронов, соответствующее новому положению ядер. Это позволяет в первом приближении рассматривать движение электронов в потенциальном поле фиксированных ядер. В этом случае волновая функция и энергия электронов будут некоторыми функциями, адиабатически меняющимися с изменением расположения ядер, координаты которых будут входить в эти функции как параметры. При изучении движения ядер, напротив, следует учитывать не мгновенное положение электронов, а поле, создаваемое их средним пространственным распределением. Такое приближенное рассмотрение называется адиабатическим, или приближением Борна—Оппенгеймера.

Воспользуемся наиболее грубым приближением, полагая, что ядра покоятся. Для этого случая координаты ядер $\mathbf{R}_1, \dots, \mathbf{R}_N$ уже не переменные, а параметры $\mathbf{R}_{10}, \dots, \mathbf{R}_{N0}$, представляющие собой координаты узлов кристаллической решетки. Для случая покоящихся ядер, когда $\mathbf{R}_\alpha = \mathbf{R}_{\alpha 0}$, кинетическая энергия ядер обращается в нуль, а энергия их взаимодействия V_0 принимает постоянное значение. Выбором начала отсчета энергии ее можно обратить в нуль. С учетом этого уравнение (2-2) упрощается и теперь описывает движение электронов в поле покоящихся ядер. Для этого случая энергия электронов E_e и их волновая функция Ψ_e уже параметрически зависят от расположения ядер $\mathbf{R}_{\alpha 0}$. В результате уравнение Шредингера для электронов, движущихся в поле покоящихся ядер, запишется в виде

$$\left\{ \sum_i \left(-\frac{\hbar^2}{2m_0} \Delta_i \right) + \frac{1}{2} \sum_i \sum_j \frac{e^2}{r_{ij}} + \right. \\ \left. + U(\mathbf{r}_1, \dots, \mathbf{r}_n, \mathbf{R}_{10}, \dots, \mathbf{R}_{N0}) \right\} \Psi_e = E_e \Psi_e. \quad (2-3)$$

Для дальнейшего сокращения числа переменных в уравнении (2-3) используется валентная аппроксимация. Принимается, что все электроны атома, кроме валентных, вместе с ядром образуют неподвижный атомный остаток (ион). При этом уравнение (2-3) записывается только для валентных электронов, движущихся в потенциальном поле фиксированных ионов.

Несмотря на проведенные упрощения, уравнение (2-3) не может быть решено в общем виде: по-прежнему мы имеем дело с задачей многих частиц, не поддающейся непосредственному решению.

2-3. ОДНОЭЛЕКТРОННОЕ ПРИБЛИЖЕНИЕ

Одним из наиболее распространенных методов решения многоэлектронной задачи для кристалла является метод Хартри—Фока, позволяющий многоэлектронную задачу свести к одноэлектронной. Его идея заключается в том, что энергия попарного взаимодействия электронов заменяется взаимодействием каждого электрона с усредненным полем всех остальных электронов. Пусть потенциальная энергия i -го электрона в этом поле будет Ω_i . Она зависит не только от движения всех остальных электронов, но зависит и от движения данного электрона, так как его движение оказывает влияние на движение остальных электронов. Поскольку поле определяет не только движение данного электрона, но и само зависит от его движения, то это поле получило название с а м о с о г л а с о в а н н о г о. Введение самосогласованного поля позволяет в уравнении (2-3) двойную сумму $\frac{1}{2} \sum_i \sum_{i'} \frac{e^2}{r_{ij}}$ заменить суммой $\sum_i \Omega_i(\mathbf{r}_i)$,

каждый член которой зависит от координат одного электрона, т. е.

$$\frac{1}{2} \sum_i \sum_{i'} \frac{e^2}{r_{ij}} = \sum_i \Omega_i(\mathbf{r}_i). \quad (2-4)$$

В этом случае $\Omega_i(\mathbf{r}_i)$ есть потенциальная энергия i -го электрона в поле всех остальных электронов.

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

$$U(\mathbf{r}_1, \mathbf{r}_2, \dots) = \sum_i U_i(\mathbf{r}_i), \quad (2-5)$$

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

С учетом выражений (2-4) и (2-5) уравнение (2-3) преобразуется к виду

$$\left\{ \sum_i \left[-\frac{\hbar^2}{2m_0} \Delta_i + \Omega_i(\mathbf{r}_i) + U_i(\mathbf{r}_i) \right] \right\} \Psi_e = E_e \Psi_e. \quad (2-6)$$

Это уравнение можно записать как

$$\hat{H}\Psi_e = (\sum_i \hat{H}_i) \Psi_e = E_e \Psi_e, \quad (2-7)$$

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

$$\hat{H}_i = -\frac{\hbar^2}{2m_0} \Delta_i + \Omega_i(\mathbf{r}_i) + U_i(\mathbf{r}_i). \quad (2-8)$$

Из соотношения (2-7) следует, что гамильтониан кристалла представляет собой сумму гамильтонианов (2-8), каждый из которых зависит от координат одной частицы. Для этого случая волновая функция системы частиц может быть представлена как произведение волновых функций, описывающих состояние отдельных частиц системы:

$$\Psi_e(\mathbf{r}_1, \mathbf{r}_2, \dots) = \Psi_1(\mathbf{r}_1) \Psi_2(\mathbf{r}_2) \dots = \Pi \Psi_i(\mathbf{r}_i), \quad (2-9)$$

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

$$E_e = E_1 + E_2 + \dots = \sum_i E_i. \quad (2-10)$$

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

Обозначим через $\Pi \Psi_i(i)$ произведение всех волновых функций, кроме Ψ_j . Тогда

$$[\Pi \Psi_{i(1)}] \Psi_1 = [\Pi \Psi_{i(2)}] \Psi_2 = \dots = \Pi \Psi_i. \quad (2-11)$$

Учитывая, что оператор \hat{H}_i действует только на волновую функцию Ψ_i , на основании уравнений (2-7), (2-9) и (2-11) получаем:

$$[\Pi \Psi_{i(1)}] \hat{H}_1 \Psi_1 + [\Pi \Psi_{i(2)}] \hat{H}_2 \Psi_2 + \dots = E_e \Pi \Psi_i. \quad (2-12)$$

Разделив обе части уравнения (2-12) на $\Pi \Psi_i$ и учитывая соотношение (2-10), будем иметь:

$$\frac{1}{\Psi_1} \hat{H}_1 \Psi_1 + \frac{1}{\Psi_2} \hat{H}_2 \Psi_2 + \dots = E_1 + E_2 + \dots \quad (2-13)$$

Каждый член в левой части уравнения (2-13) зависит от координат лишь одного электрона, поэтому это уравнение эквивалентно системе следующих одноэлектронных уравнений:

$$\left. \begin{aligned} \hat{H}_1 \Psi_1 &= E_1 \Psi_1; \\ \hat{H}_2 \Psi_2 &= E_2 \Psi_2; \\ \vdots & \\ \vdots & \end{aligned} \right\} \quad (2-14)$$

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

$$\hat{H}\Psi = E\Psi, \quad (2-15)$$

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

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

$$V(\mathbf{r}) = U(\mathbf{r}) + \Omega(\mathbf{r}), \quad (2-16)$$

то уравнение Шредингера для электрона кристалла запишется в виде

$$\left[-\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}) \right] \Psi(\mathbf{r}) = E\Psi(\mathbf{r}). \quad (2-17)$$

Существуют различные способы рационального выбора вида функции $\Psi(\mathbf{r})$ 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 (2-16) и (2-17) есть полный потенциал кристалла, обладающий трехмерной периодичностью решетки. Качественный вид $V(\mathbf{r})$ для одномерной решетки показан на рис. 2-1, где пунктирной кривой изображен потенциал изолированного атома $U_a(\mathbf{r})$.

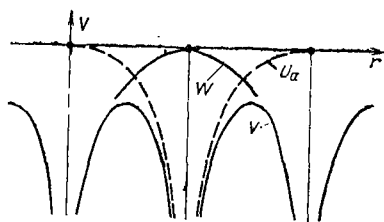


Рис. 2-1. Схематическое изображение периодического потенциала одномерной решетки.

Сплошная кривая — потенциал решетки $V(\mathbf{r})$; пунктирная — атомный потенциал $U_a(\mathbf{r})$; штрихпунктирная — возмущающий потенциал $W(\mathbf{r})$.

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

$$\mathbf{a}_n = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \quad (2-18)$$

где \mathbf{a}_1 , \mathbf{a}_2 , \mathbf{a}_3 — периоды идентичности решетки по трем произвольным направлениям, а n_1 , n_2 , n_3 — произвольные целые числа, кристалл совмещается сам с собой. Из этого следует, что точка с радиус-вектором \mathbf{r} и точка с радиус-вектором $\mathbf{r} + \mathbf{a}_n$ физически эквивалентны, поэтому

$$V(\mathbf{r} + \mathbf{a}_n) = V(\mathbf{r}). \quad (2-19)$$

Соотношение (2-19) выражает условие периодичности потенциального поля кристалла. В периодическом поле кристалла, если

волновая функция электрона $\Psi(\mathbf{r})$ невырождена, то она отличается от волновой функции $\Psi(\mathbf{r} + \mathbf{a}_n)$ только постоянным множителем, т. е.

$$\Psi(\mathbf{r} + \mathbf{a}_n) = C\Psi(\mathbf{r}). \quad (2-20)$$

Но из условия нормировки волновой функции следует, что

$$|C|^2 = 1, \quad (2-21)$$

поэтому можно положить, что

$$C = e^{i\mathbf{k}\mathbf{a}_n}, \quad (2-22)$$

так как

$$|C|^2 = |e^{i\mathbf{k}\mathbf{a}_n}|^2 = |\cos \mathbf{k}\mathbf{a}_n + i \sin \mathbf{k}\mathbf{a}_n|^2 = \cos^2 \mathbf{k}\mathbf{a}_n + \sin^2 \mathbf{k}\mathbf{a}_n = 1;$$

здесь \mathbf{k} — постоянный вектор, характеризующий квантовое состояние электрона в кристалле, называется волновым вектором. Поскольку произведение $(\mathbf{k}\mathbf{a}_n)$ должно быть безразмерным, то волновой вектор имеет размерность обратной длины (см^{-1}) и численно равен $k = 2\pi/\lambda$.

Из уравнений (2-20) и (2-22) следует, что

$$\Psi(\mathbf{r} + \mathbf{a}_n) = e^{i\mathbf{k}\mathbf{a}_n}\Psi(\mathbf{r}), \quad (2-23)$$

откуда

$$\Psi(\mathbf{r}) = e^{-i\mathbf{k}\mathbf{a}_n}\Psi(\mathbf{r} + \mathbf{a}_n) = e^{i\mathbf{k}\mathbf{r}}U_{\mathbf{k}}(\mathbf{r}), \quad (2-24)$$

где

$$U_{\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{k}(\mathbf{r} + \mathbf{a}_n)}\Psi(\mathbf{r} + \mathbf{a}_n). \quad (2-25)$$

Функция $U_{\mathbf{k}}(\mathbf{r})$ обладает трехмерной периодичностью кристаллической решетки, так как согласно (2-23) и (2-25)

$$\begin{aligned} U_{\mathbf{k}}(\mathbf{r} + \mathbf{a}_l) &= e^{-i\mathbf{k}(\mathbf{r} + \mathbf{a}_l + \mathbf{a}_n)}\Psi(\mathbf{r} + \mathbf{a}_l + \mathbf{a}_n) = \\ &= e^{-i\mathbf{k}(\mathbf{r} + \mathbf{a}_n + \mathbf{a}_l)}e^{i\mathbf{k}\mathbf{a}_l}\Psi(\mathbf{r} + \mathbf{a}_n) = e^{-i\mathbf{k}(\mathbf{r} + \mathbf{a}_n)}\Psi(\mathbf{r} + \mathbf{a}_n) = U_{\mathbf{k}}(\mathbf{r}). \end{aligned} \quad (2-26)$$

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

$$\Psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}}U_{\mathbf{k}}(\mathbf{r}), \quad (2-27)$$

где $e^{i\mathbf{k}\mathbf{r}}$ представляет собой плоскую волну, бегущую в направлении вектора \mathbf{k} , а $U_{\mathbf{k}}(\mathbf{r})$ — некая функция координат, зависящая от волнового вектора \mathbf{k} и имеющая периодичность решетки. Выражение (2-27) для $\Psi_{\mathbf{k}}(\mathbf{r})$ носит название волны (или функции) Блоха.

Если функцию Блоха (2-27) подставить в уравнение (2-15), то будем иметь:

$$\hat{H}\Psi_{\mathbf{k}}(\mathbf{r}) = E\Psi_{\mathbf{k}}(\mathbf{r}). \quad (2-28)$$

Из равенства (2-28) следует, что энергия электрона в кристалле должна зависеть от волнового вектора \mathbf{k} , т. е. $E = E(\mathbf{k})$.

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

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

2-4. ПРИБЛИЖЕНИЕ СИЛЬНО СВЯЗАННЫХ ЭЛЕКТРОНОВ

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

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

Рассмотрим теперь, что происходит с энергетическими уровнями при взаимодействии большого числа атомов, образующих кристалл. Уровни энергии внутренних электронов, расположенных ближе к ядру, при этом почти не изменяются. Об этом можно судить по рентгеновским характеристическим спектрам, вид которых почти не зависит от соединения или агрегатного состояния вещества. Однако оптический спектр, обусловленный переходом самых внешних эквивалентных электронов, резко меняется.

Если считать, что кинетическая энергия электронов значительно больше пространственных изменений его потенциальной энергии, то периодический потенциал $V(\mathbf{r})$ можно рассматривать как малое

возмущение свободного движения электронов. Такой подход, получивший название приближения почти свободных электронов, дает более или менее удовлетворительные результаты при решении некоторых задач для металлов.

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

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

$$\Psi_g = \Psi_g(\mathbf{r} - \mathbf{R}_g). \quad (2-29)$$

В приближении сильно связанных электронов считается, что состояние электрона в атоме кристалла мало меняется по сравнению с его состоянием в изолированном атоме. Поэтому волновую функцию электрона в кристалле можно представить в виде линейной комбинации атомных волновых функций (2-29):

$$\Psi(\mathbf{r}) = \sum_g a_g \Psi_g. \quad (2-30)$$

Чтобы выражение (2-30) для $\Psi(\mathbf{r})$ имело вид блоховской функции (2-27), необходимо положить:

$$a_g = e^{i\mathbf{kR}_g}. \quad (2-31)$$

Тогда

$$\Psi(\mathbf{r}) = \sum_g e^{i\mathbf{kR}_g} \Psi_g(\mathbf{r} - \mathbf{R}_g). \quad (2-32)$$

Выбор волновой функции электрона в кристалле в виде (2-32) вполне законен. Поскольку атомная волновая функция экспоненциально убывает, то вблизи g -го узла решетки

$$\Psi(\mathbf{r}) \approx \text{const } \Psi_g(\mathbf{r} - \mathbf{R}_g), \quad (2-33)$$

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

Подставим выражение (2-30) в уравнение Шредингера (2-15) для электрона в кристалле. В результате получим:

$$\sum_g a_g (\hat{H} \Psi_g - E \Psi_g) = 0. \quad (2-34)$$

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

$$\hat{H} = -\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}). \quad (2-35)$$

Для нахождения собственных значений энергии E электрона в кристалле при решении уравнения (2-34) будем считать, что энергетический спектр изолированного атома известен. Обозначим гамильтониан изолированного атома через

$$\hat{H}_a = -\frac{\hbar^2}{2m_0} \Delta + U_a(\mathbf{r}) \quad (2-36)$$

и запишем уравнение Шредингера для изолированного атома:

$$\hat{H}_a \Psi_a = E_a \Psi_a; \quad (2-37)$$

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

Вычитая выражение (2-36) из (2-35), будем иметь:

$$\hat{H} = \hat{H}_a + W(\mathbf{r}). \quad (2-38)$$

Здесь через

$$W(\mathbf{r}) = V(\mathbf{r}) - U_a(\mathbf{r}) \quad (2-39)$$

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

Уравнение Шредингера (2-34) с учетом соотношения (2-38) примет вид:

$$\sum_g a_g \{ \hat{H}_g \Psi_g + W(\mathbf{r}) \Psi_g - E \Psi_g \} = 0. \quad (2-40)$$

Если учесть равенство (2-37), то это выражение преобразуется следующим образом:

$$\sum_g a_g \{ W(\mathbf{r}) \Psi_g + (E_a - E) \Psi_g \} = 0. \quad (2-41)$$

Умножим уравнение (2-41) на функцию $\Psi_{g'}^*$, комплексно-сопряженную с Ψ_g , и проинтегрируем по всему объему кристалла τ :

$$\sum_g a_g \left\{ \int_{\tau} \Psi_g W(\mathbf{r}) \Psi_{g'}^* d\tau + (E_a - E) \int_{\tau} \Psi_g \Psi_{g'}^* d\tau \right\} = 0. \quad (2-42)$$

Введем обозначения:

$$\int_{\tau} \Psi_g W(\mathbf{r}) \Psi_{g'}^* d\tau = A(\mathbf{R}_g - \mathbf{R}_{g'}). \quad (2-43)$$

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

$$\int \Psi_{\mathbf{g}} \Psi_{\mathbf{g}'}^* d\tau = S(\mathbf{R}_{\mathbf{g}} - \mathbf{R}_{\mathbf{g}'}). \quad (2-44)$$

Этот интеграл зависит только от перекрытия атомных волновых функций.

С учетом соотношений (2-31), (2-43) и (2-44) уравнение (2-42) можно записать в виде

$$\sum_{\mathbf{g}} A(\mathbf{R}_{\mathbf{g}} - \mathbf{R}_{\mathbf{g}'}') e^{i\mathbf{k}\mathbf{R}_{\mathbf{g}}} + (E_a - E) \sum_{\mathbf{g}} S(\mathbf{R}_{\mathbf{g}} - \mathbf{R}_{\mathbf{g}'}') e^{i\mathbf{k}\mathbf{R}_{\mathbf{g}}} = 0. \quad (2-45)$$

Умножим полученное выражение на $e^{-i\mathbf{k}\mathbf{R}_{\mathbf{g}'}}$. Это возможно, так как суммирование в (2-45) производится только по \mathbf{g} , а \mathbf{g}' — неизменно. Тогда

$$\begin{aligned} & \sum_{\mathbf{g}} A(\mathbf{R}_{\mathbf{g}} - \mathbf{R}_{\mathbf{g}'}) e^{i\mathbf{k}(\mathbf{R}_{\mathbf{g}} - \mathbf{R}_{\mathbf{g}'})} + \\ & + (E_a - E) \sum_{\mathbf{g}} S(\mathbf{R}_{\mathbf{g}} - \mathbf{R}_{\mathbf{g}'}) e^{i\mathbf{k}(\mathbf{R}_{\mathbf{g}} - \mathbf{R}_{\mathbf{g}'})} = 0. \end{aligned} \quad (2-46)$$

Упростим выражение (2-46), введя вектор \mathbf{q} , соединяющий атомы \mathbf{g} и \mathbf{g}' :

$$\mathbf{q} = \mathbf{R}_{\mathbf{g}} - \mathbf{R}_{\mathbf{g}'}, \quad (2-47)$$

при этом получим:

$$\sum_{\mathbf{q}} A(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}} + (E_a - E) \sum_{\mathbf{q}} S(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}} = 0. \quad (2-48)$$

Отсюда можно найти собственные значения энергии электронов в кристалле:

$$E = E_a + \frac{\sum_{\mathbf{q}} A(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}}}{\sum_{\mathbf{q}} S(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}}}. \quad (2-49)$$

Из равенства (2-49) следует, что энергия электрона в кристалле складывается из его энергии на соответствующем уровне в изолированном атоме и добавочного члена, представляющего собой периодическую функцию волнового вектора \mathbf{k} . Теперь вместо одного атомного уровня для электронов в кристалле имеем энергетическую зону, границы которой определяются минимальным и максимальным значениями второго члена в соотношении (2-49).

Рассмотрим более подробно добавочный член в выражении (2-49). Для упрощения задачи будем считать, что атомные волновые функции очень быстро спадают и можно пренебречь их перекрытием даже для соседних атомов. Такое допущение является

обоснованным в приближении сильно связанных электронов. При этом все члены в выражении обращаются в нуль, кроме члена, соответствующего $q = 0$, когда он равен единице из условия нормировки атомных волновых функций, т. е.

$$S(\mathbf{q}) = \begin{cases} 1 & \text{при } \mathbf{q} = 0; \\ 0 & \text{при } \mathbf{q} \neq 0. \end{cases} \quad (2-50)$$

Отсюда следует, что

$$\sum_{\mathbf{g}} S(\mathbf{q}) e^{i\mathbf{k}\mathbf{q}} = 1. \quad (2-51)$$

Первый член суммы числителя дроби (2-49) для $\mathbf{q} = 0$

$$A(0) = \int_{\tau} \Psi_0(\mathbf{r}) W(\mathbf{r}) \Psi_0^*(\mathbf{r}) d\tau = \int_{\tau} |\Psi_0(\mathbf{r})|^2 W(\mathbf{r}) d\tau = C \quad (2-52)$$

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

Хотя мы пренебрегли перекрытием волновых функций соседних атомов, но малое значение волновой функции атома вблизи нулевого узла частично компенсируется большой разностью $W(\mathbf{r}) = V(\mathbf{r}) - U_a(\mathbf{r})$ в этой области (см. рис. 2-1). Поэтому в сумме, стоящей в числителе дроби (2-49), следует учитывать значения обменных интегралов для узлов с $\mathbf{q} \neq 0$, расположенных по соседству с нулевым узлом g_0 (рис. 2-2).

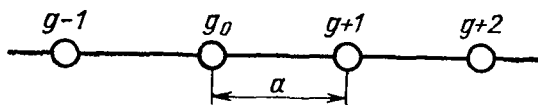


Рис. 2-2. Схематическое изображение одномерной решетки.

Если волновая функция соответствует s -состоянию, то эти обменные интегралы будут одинаковы для всех атомов. Пусть величина этого обменного интеграла равна A . Тогда с учетом соотношений (2-51) и (2-52) выражение (2-49) для энергии электрона в периодическом поле кристалла примет вид:

$$E = E_a + C + A \sum_{\mathbf{g}} e^{i\mathbf{k}\mathbf{q}}. \quad (2-53)$$

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

$$\begin{aligned} \sum_{\mathbf{g}} e^{i\mathbf{k}\mathbf{q}} &= e^{ik_x a} + e^{-ik_x a} + e^{ik_y a} + e^{-ik_y a} + e^{ik_z a} + e^{-ik_z a} = \\ &= 2(\cos k_x a + \cos k_y a + \cos k_z a), \end{aligned} \quad (2-54)$$

так как $e^{ika} = \cos ka + i \sin ka$, а $e^{-ika} = \cos ka - i \sin ka$. Поэтому выражение для энергии электрона в периодическом поле простой кубической решетки с учетом соотношения (2-54) примет вид:

$$E = E_a + C + 2A (\cos k_x a + \cos k_y a + \cos k_z a). \quad (2-55)$$

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 уровня зависит от знака C .

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

3. Экстремальные значения выражения (2-55), 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:

$$E_{\text{макс}} - E_{\text{мин}} = 12A, \quad (2-56)$$

т. е. зависит от величины обменного интеграла.

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

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

6. С ростом энергии ширина энергетических зон увеличивается, а ширина запрещенных зон уменьшается (рис. 2-3).

7. Уровень E_a в изолированном атоме может быть вырожденным. В кристаллической решетке вырождение может быть частично или полностью снято. При этом атомный уровень расщепляется на несколько зон, число которых соответствует степени вырождения. Например, для p -состояния фактор вырождения $g = 3$, так как $g = 2l + 1$, где l — азимутальное квантовое число, которое для p -состояния равно 1. Следовательно, из атомного p -состояния в кристалле возможно образование трех зон.

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

$$E(\mathbf{k}) = E(-\mathbf{k}).$$

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

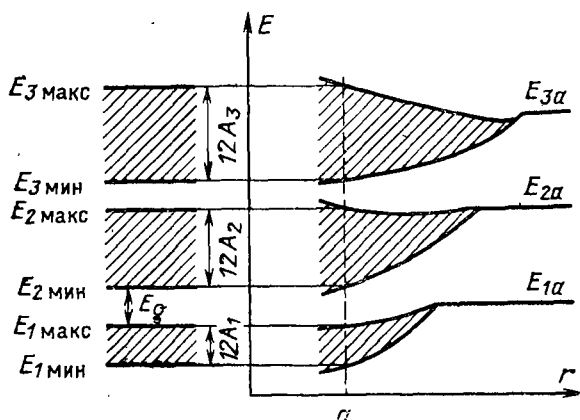


Рис. 2-3. Образование зон энергии из энергетических уровней при сближении атомов.

a — постоянная решетки кристалла.

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

2-5. ЧИСЛО СОСТОЯНИЙ ЭЛЕКТРОНОВ В ЭНЕРГЕТИЧЕСКОЙ ЗОНЕ,

Для подсчета числа состояний в зоне нужно при решении уравнения Шредингера для электрона учесть граничные условия на краях кристалла. Любые условия на границах кристалла не отразятся существенным образом на состояниях электронов в его объеме, так как число узлов в любом реальном кристалле очень 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

$$L_x = aN_x; \quad L_y = aN_y; \quad L_z = aN_z, \quad (2-57)$$

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

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

$$\Psi(x + L_x, y + L_y, z + L_z) = \Psi(x, y, z). \quad (2-58)$$

Уравнение (2-58) по существу выражает граничные условия цикличности Борна—Кармана.

Учитывая вид волновой функции для кристаллов (2-27), получаем:

$$\begin{aligned} & \Psi(x + L_x, y + L_y, z + L_z) = \\ & = U_{\mathbf{k}}(\mathbf{r}) e^{ik_x L_x} e^{ik_y L_y} e^{ik_z L_z} = \Psi(x, y, z). \end{aligned} \quad (2-59)$$

Для выполнения условия (2-58) необходимо в выражении (2-59) принять:

$$e^{ik_x L_x} = e^{ik_y L_y} = e^{ik_z L_z} = 1. \quad (2-60)$$

Это равенство выполнимо, если показатель экспоненты есть целое число, умноженное на $2\pi i$, т. е.

$$k_x L_x = 2\pi n_1; \quad k_y L_y = 2\pi n_2; \quad k_z L_z = 2\pi n_3, \quad (2-61)$$

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

$$k_x = \frac{2\pi}{L_x} n_1; \quad k_y = \frac{2\pi}{L_y} n_2; \quad k_z = \frac{2\pi}{L_z} n_3, \quad (2-62)$$

где

$$\begin{aligned} n_1 &= 0, \pm 1, \pm 2 \dots; & n_2 &= 0, \pm 1, \pm 2 \dots; \\ n_3 &= 0, \pm 1, \pm 2 \dots \end{aligned}$$

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

С учетом значений волнового вектора \mathbf{k} (2-62) можно записать волновую функцию (2-32) для одномерной задачи в виде

$$\Psi_{\mathbf{k}} = \sum_g e^{ik_x ag} \Psi_g = \sum_g e^{i \frac{2\pi n_1 g}{N_x}} \Psi_g, \quad (2-63)$$

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

Из выражения (2-63) видно, что волновая функция для $n_1 = \pm N_x$ будет Ψ_g и совпадает с функцией для $n_1 = 0$, а волновая функция для $n_1 = \pm(N_x + 1)$ совпадает с функцией для $n_1 = \pm 1$ и т. д. Это означает, что компонента k_x имеет N_x значений, соответствующих различным n_1 . При этом n_1 может принимать только значения $0, 1, 2, \dots, (N_x - 1)$, так как $E(\mathbf{k}) = E(-\mathbf{k})$, т. е. n_1 изменяется только в пределах

$$0 \leq n_1 < N_x \text{ или } -\frac{N_x}{2} \leq n_1 < \frac{N_x}{2}. \quad (2-64)$$

Из соотношений (2-64), (2-62) и (2-57) получаем, что компоненты вектора \mathbf{k} находятся в следующих интервалах значений:

$$-\frac{\pi}{a} \leq k_x < \frac{\pi}{a}; \quad -\frac{\pi}{a} \leq k_y < \frac{\pi}{a}; \quad -\frac{\pi}{a} \leq k_z < \frac{\pi}{a}, \quad (2-65)$$

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

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

Согласно квантовой механике состояние электрона в атоме характеризуется главным квантовым числом n , азимутальным квантовым числом l , магнитным квантовым числом m и s_z — проекцией спина на ось z . Состояние электрона в кристалле согласно принципу Паули также должно описываться четырьмя квантовыми числами. Как следует из соотношения (2-62), тремя квантовыми числами являются проекции волнового вектора k_x, k_y, k_z , а четвертым квантовым числом должно быть s_z . Проекция s_z может принимать только два значения: $+1/2$ и $-1/2$. Это означает, что в состоянии (k_x, k_y, k_z) может быть не более двух электронов. Но набор (k_x, k_y, k_z) согласно уравнению (2-55) определяет величину энергии $E(\mathbf{k})$ для данной зоны. Следовательно, на каждом энергетическом уровне зоны, который определяется волновым вектором \mathbf{k} , в соответствии с принципом Паули может находиться не более двух электронов с противоположно направленными спинами. Таким образом, в простой энергетической зоне, возникшей из невырожденного атомного уровня, имеется $2N$ квантовых состояний, соответственно N энергетических уровней, и в зоне может быть не более $2N$ электронов. Если зона g — кратко вырождена, то в ней может быть $2gN$ электронов. Из этого следует, что число квантовых состояний в зоне равно общему числу мест на уровнях изолированных атомов, из которых образовалась эта зона, т. е. имеет место сохранение числа состояний при образовании кристалла из атомов.

Оценим среднее расстояние между соседними уровнями энергии в разрешенной зоне. Примем значение параметра решетки $a =$

$= 4\text{Å} = 4 \cdot 10^{-8}$ см, что дает объем элементарной ячейки $a^3 = 64 \cdot 10^{-24}$ см³. Число элементарных ячеек в кристалле единичного объема $V = 1$ см³, равное числу состояний в зоне, составит:

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

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

2-6. КВАЗИИМПУЛЬС

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

$$E = \frac{P^2}{2m_0}. \quad (2-66)$$

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

$$\lambda = \frac{h}{P} = \frac{h}{m_0 v}, \quad (2-67)$$

где h — постоянная Планка.

Так как волновое число k — число волн, укладываемых на длине 2π см, равно:

$$k = \frac{2\pi}{\lambda}, \quad (2-68)$$

то импульс свободного электрона

$$\mathbf{P} = \hbar k, \quad (2-69)$$

а его энергия

$$E = \frac{\hbar^2 k^2}{2m_0}, \quad (2-70)$$

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

Для электрона, движущегося в периодическом поле кристалла, можно ввести величину $\mathbf{p} = \hbar \mathbf{k}$, называемую *к в а з и и м п у л ь с о м*. В соответствии с дискретным спектром k квазиимпульс \mathbf{p} также квантован. Согласно неравенствам (2-65) в кубической решетке квазиимпульс должен изменяться в пределах

$$-\frac{\pi\hbar}{a} \leq p_i < \frac{\pi\hbar}{a}; \quad i = x, y, z. \quad (2-71)$$

Если в формулу (2-55) вместо k_i подставить значение p_i , найденное из соотношения (2-71), то для энергии электрона в кубическом кристалле получим:

$$E = E_a + C + 2A \left(\cos \frac{a}{\hbar} p_x + \cos \frac{a}{\hbar} p_y + \cos \frac{a}{\hbar} p_z \right). \quad (2-72)$$

Из соотношения (2-72) следует, что энергия электрона в кристалле — четная функция квазиимпульса, т. е.

$$E(\mathbf{p}) = E(-\mathbf{p}). \quad (2-73)$$

2.7. ЗОНЫ БРИЛЛЮЭНА

Как было показано в § 2-5, в энергетической зоне кристалла имеется N энергетических состояний, которым соответствуют значения компонент волнового вектора

$$k_i = \frac{2\pi}{L_i} n_j \quad (2-74)$$

и компонент квазиимпульса

$$p_i = \frac{2\pi\hbar}{L_i} n_j, \quad (2-75)$$

где $i = x, y, z$, а $j = 1, 2, 3$.

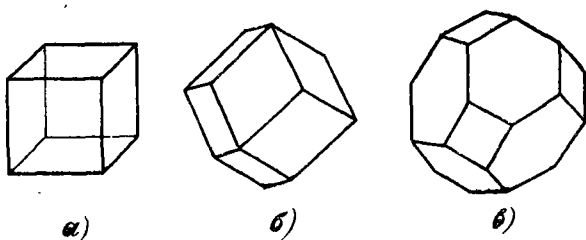


Рис. 2-4. Первая зона Бриллюэна для кристалла с простой кубической решеткой (а), с кубической объемно-центрированной (б) и кубической гранецентрированной решеткой (в).

Для кристалла с простой кубической решеткой согласно соотношениям (2-65) и (2-71) достаточно рассматривать изменение компонент k_i и p_i в пределах

$$\left. \begin{aligned} -\frac{\pi}{a} &\leq k_i < \frac{\pi}{a}; \\ -\frac{\pi\hbar}{a} &\leq p_i < \frac{\pi\hbar}{a}. \end{aligned} \right\} \quad (2-76)$$

Этим значениям квазимпульса в системе координат (p_x, p_y, p_z) будет соответствовать некоторая область, построенная вокруг начала координат и содержащая все возможные различные состояния. Эта область называется первой, или основной, зоной Бриллюэна.

Для кристалла с простой кубической решеткой первая зона Бриллюэна представляет собой куб (рис. 2-4, а) объемом

$$V_{\text{Б}}^{(p)} = \left(\frac{2\pi\hbar}{a}\right)^3 = \frac{\hbar^3}{a^3}. \quad (2-77)$$

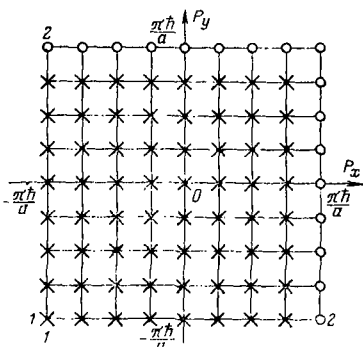


Рис. 2-5. Сечение первой зоны Бриллюэна для кристалла с простой кубической решеткой плоскостью $p_z = 0$.

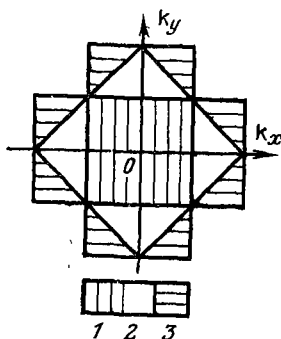


Рис. 2-6. Первые три зоны Бриллюэна для плоской квадратной решетки.

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

$$V_{\text{Б}}^{(k)} = \left(\frac{2\pi}{a}\right)^3 = \frac{8\pi^3}{a^3}. \quad (2-78)$$

На рис. 2-5 схематично показано сечение первой зоны Бриллюэна для простой кубической решетки плоскостью $p_z = 0$. Крестиками отмечены точки, соответствующие различным состояниям электронов. Центр зоны Бриллюэна (начало координат) изображает состояние $n_1 = n_2 = n_3 = 0$, т. е. $p = 0$. Расстояние между ближайшими точками в зоне Бриллюэна согласно соотношению (2-75) равно $2\pi\hbar/L = \hbar/L$. При этом, как следует из равенства (2-63), точки, отмеченные на рисунке светлыми кружками, изображают состояния, которые физически эквивалентны соответствующим состояниям на противоположных гранях зоны Бриллюэна. Например, точки 1 и 2 изображают эквивалентные состояния, имеющие одну и ту же волновую функцию, и электроны, находящиеся в этих двух состояниях, будут иметь одинаковую энергию.

Первую зону Бриллюэна можно разбить на элементарные кубические ячейки объемом

$$\Delta V_{\text{Б}}^{(p)} = \left(\frac{2\pi\hbar}{L}\right)^3 = \frac{\hbar^3}{a^3 N} = \frac{\hbar^3}{V} \quad (2-79)$$

и

$$\Delta V_{\text{Б}}^{(k)} = \left(\frac{2\pi}{L}\right)^3 = \frac{8\pi^3}{a^3 N} = \frac{8\pi^3}{V}, \quad (2-79a)$$

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

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

Вторая и последующие зоны Бриллюэна, соответствующие второй и последующим энергетическим зонам, имеют более сложную конфигурацию, но их объем остается постоянным. Они также содержат N элементарных ячеек, каждой из которых можно сопоставить ячейку в первой зоне, изображающую эквивалентное состояние. В качестве примера на рис. 2-6 изображены в k -пространстве три первые бриллюэновские зоны для плоской квадратной решетки. Первая зона Бриллюэна для кристалла с простой кубической решеткой изображена на рис. 2-4, а, с кубической объемно-центрированной решеткой — на рис. 2-4, б и с кубической грацецентрированной решеткой на рис. 2-4, в.

2.8. ВОЗМОЖНОЕ ЗАПОЛНЕНИЕ ЭЛЕКТРОННЫХ СОСТОЯНИЙ ВАЛЕНТНОЙ ЗОНЫ

Рассмотрим несколько примеров возможного заполнения электронами квантовых состояний валентной зоны.

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

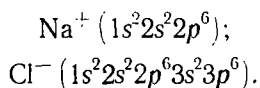
Типичными представителями веществ, обладающих хорошей проводимостью, являются одновалентные щелочные металлы Li, Na, K, Rb, Cs и благородные — Cu, Ag, Au. У щелочных элементов валентный электрон находится в ns -состоянии, а остальные электроны занимают состояния, которые они полностью заполняют. В качестве примера рассмотрим натрий, у которого 11 электронов распределены следующим образом: $\text{Na}^{(11)} (1s^2 2s^2 2p^6 3s)$. При образовании кристалла уровни атомов натрия расщепляются в зоны.

Зоны, образованные из уровней $1s$, $2s$ и $2p$, полностью заполнены, так как в их $2N$ -, $2N$ - и $6N$ -состояниях находятся $2N$, $2N$ и $6N$ электронов. В зоне валентных электронов, образованной из уровня $3s$, в $2N$ -состояниях находятся N электронов, которые заполняют энергетическую зону только наполовину.

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

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

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



У иона хлора в $3p$ -состоянии имеется 6 электронов, что соответствует наиболее устойчивому энергетическому состоянию, т. е. все электронные состояния заняты и на каждом уровне находится по 2 электрона с противоположно направленными спинами. Из этого состояния в кристалле образуется зона, в которой все $6N$ -состояний заняты электронами. Это будет валентная зона. У иона натрия в $3s$ -состоянии не имеется электрона, а поэтому из этого уровня в кристалле образуется зона, в которой все $2N$ электронных 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 такова (см. рис. 1-4 и 1-5), что в твердом состоя-

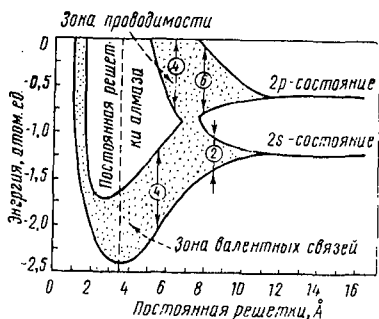


Рис. 2-7. Образование энергетических зон из атомных уровней в алмазе.

нии у них в образовании ковалентной связи принимают участие четыре электрона каждого атома. При этом, как следует из рис. 2-7, зоны, образованные из ns - и np -состояний, перекрываются, образуя общую зону с числом состояний $8N$. С уменьшением межатомного расстояния эта зона затем расщепляется на две зоны с $4N$ квантовыми состояниями в каждой. Нижняя зона содержит $4N$ заполненных электронами состояний — это валентная зона, а у верхней зоны $4N$ электронных состояний свободны — это зона проводимости.

2-9. ЗАВИСИМОСТЬ ЭНЕРГИИ ЭЛЕКТРОНА ОТ ВОЛНОВОГО ВЕКТОРА У ДНА И ПОТОЛКА ЭНЕРГЕТИЧЕСКОЙ ЗОНЫ

Энергия электрона в простой кубической решетке, которая выражается уравнением (2-55), определяется величиной и знаком обменного интеграла. В свою очередь знак обменного интеграла A ,

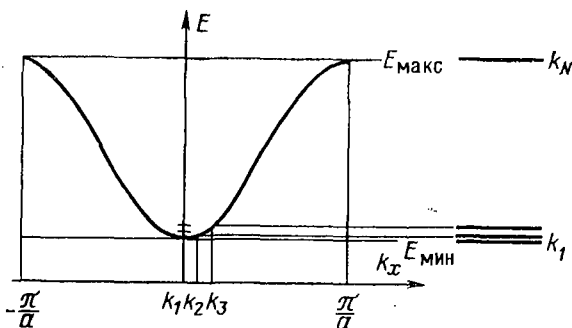


Рис. 2-8. Зависимость $E(k_x)$ для зоны, образованной из s -состояния.

как следует из равенства (2-43), зависит от степени перекрытия атомных волновых функций соответствующего состояния и знака энергии возмущения $W(\mathbf{r})$. Энергия возмущения $W(\mathbf{r})$ согласно рис. 2-1 всегда является величиной отрицательной. Поэтому для s -состояния, имеющего сферическую электронную оболочку, $A < 0$. Если учесть знак обменного интеграла, то энергия электрона в зоне, образованной из s -состояния, будет равна:

$$E = E' - 2|A|(\cos k_x a + \cos k_y a + \cos k_z a), \quad (2-80)$$

где $E' = E_a + C$.

Эта зависимость для значений волнового вектора k_x , меняющегося в пределах первой зоны Бриллюэна, изображена на рис. 2-8. Кривая $E(k)$ для s -зоны обращена вогнутоостью вверх и для компонент волновых векторов k_y и k_z .

На рис. 2-8 точка $k_1 = 0$ определяет состояние с минимальной величиной энергии. Разрешенные уровни энергии соответствуют значениям $k = \frac{2\pi}{L_x} n$. Точка $k_N = \frac{2\pi}{L_x} (N-1)$ отвечает состоянию с максимальным значением энергии.

Для p -состояния, атомные функции которого имеют вид:

$$\frac{x}{r} f(\mathbf{r}), \quad (2-81)$$

перекрывающиеся части функций Ψ_a для атомов с различными координатами, как видно из рис. 2-9, имеют в основном противоположные знаки.

Рис. 2-9. Атомные функции p -состояния, описываемые уравнением (2-81).

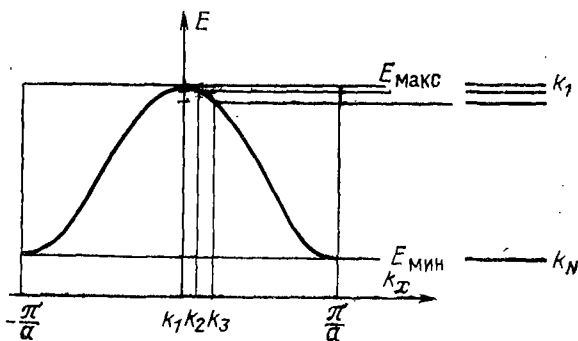
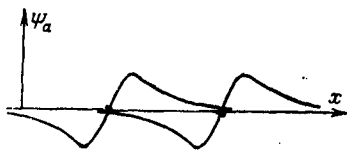


Рис. 2-10. Зависимость $E(k_x)$ для зоны, образованной из p -состояния.

Как видно из рис. 2-9, произведение перекрывающихся частей функций одинакового знака мало, поэтому можно считать, что $A > 0$; тогда вместо (2-80) получаем:

$$E = E'' + 2|A|(\cos k_x a + \cos k_y a + \cos k_z a). \quad (2-82)$$

Зависимость энергии от k_x для зоны, образованной из p -состояния, представлена на рис. 2-10. Здесь при $k_1 = 0$ имеется максимум энергии, а k_N соответствует минимуму энергии.

Для зоны, образованной из d -состояния, зависимость $E(k)$ аналогична зависимости для зоны из s -состояния и график $E(k)$ имеет вид, подобный изображенному на рис. 2-8.

Рассмотрим более детально выражение (2-55) для энергии электронов, движущихся в потенциальном поле простой кубической решетки. Если выбрать начало отсчета энергии так, чтобы $E_a + C = 0$, то энергия электронов в зоне, образованной из s -состояния, запишется в виде

$$E = -2A (\cos k_x a + \cos k_y a + \cos k_z a), \quad (2-83)$$

где $A > 0$.

Обозначим:

$$E_{\text{мин}} = -6A = E_d; \quad (2-84)$$

$$E_{\text{макс}} = 6A = E_b. \quad (2-85)$$

Проанализируем равенство (2-83) в окрестности точки $\mathbf{k} = 0$, т. е. вблизи центра первой зоны Бриллюэна. Разложим функцию $E(\mathbf{k})$ в ряд и ограничимся первыми членами, приняв во внимание, что $k_i a \ll 1$:

$$\begin{aligned} E &= -2A \left\{ \left[1 - \frac{(k_x a)^2}{2} \right] + \left[1 - \frac{(k_y a)^2}{2} \right] + \left[1 - \frac{(k_z a)^2}{2} \right] \right\} = \\ &= -6A + Aa^2 (k_x^2 + k_y^2 + k_z^2) = -6A + Aa^2 k^2. \end{aligned} \quad (2-86)$$

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

По аналогии с выражением (2-70) для энергии свободного электрона запишем E для электрона в кристалле, используя равенства (2-84) и (2-86), в виде

$$E = E_d + \frac{\hbar^2}{2m_n^*(d)} k^2, \quad (2-87)$$

где

$$m_n^*(d) = \frac{\hbar^2}{2Aa^2} \quad (2-88)$$

есть эффективная масса электрона у дна энергетической зоны.

Рассмотрим теперь формулу (2-83) для состояний, близких к верхнему краю зоны. Введем вектор \mathbf{k}' с компонентами $k'_i = \frac{\pi}{a} - k_i$. При этом у потолка зоны $k'_i a \ll 1$ и разложение (2-83) в ряд с учетом соотношения (2-85) примет вид:

$$\begin{aligned} E &= -2A [\cos(\pi - k'_x a) + \cos(\pi - k'_y a) + \cos(\pi - k'_z a)] = \\ &= 2A \left\{ \left[1 - \frac{(k'_x a)^2}{2} \right] + \left[1 - \frac{(k'_y a)^2}{2} \right] + \left[1 - \frac{(k'_z a)^2}{2} \right] \right\} = \\ &= 6A - Aa^2 (k')^2 = E_b - \frac{\hbar^2}{2m_n^*(b)} (k')^2. \end{aligned} \quad (2-89)$$

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

$$m_n(b) = -\frac{\hbar^2}{2Aa^2}. \quad (2-90)$$

Поскольку масса по самой своей сути не может быть отрицательна, 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:

$$m_p(b) = -m_n(b) = \frac{\hbar^2}{2Aa^2}. \quad (2-91)$$

Следовательно, эффективная масса дырки у верхнего края зоны по величине совпадает с эффективной массой электрона у дна этой зоны и обратно пропорциональна величине обменного интеграла. Но обменный интеграл согласно соотношению (2-56) определяет ширину энергетической зоны. Отсюда следует, что эффективная масса обратно пропорциональна ширине энергетической зоны. Это означает, что чем выше зона энергии, тем она шире и тем меньше в ней эффективная масса носителей заряда.

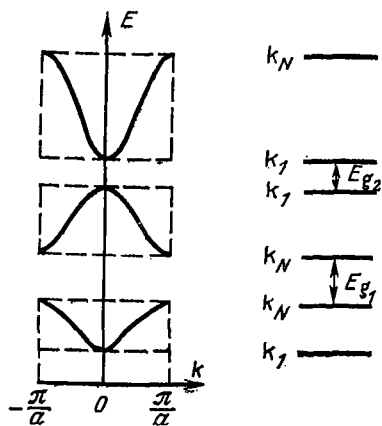


Рис. 2-11. Зависимость $E(k)$ для трех соседних зон, образованных из s -, p - и d -состояний атомов одномерной решетки.

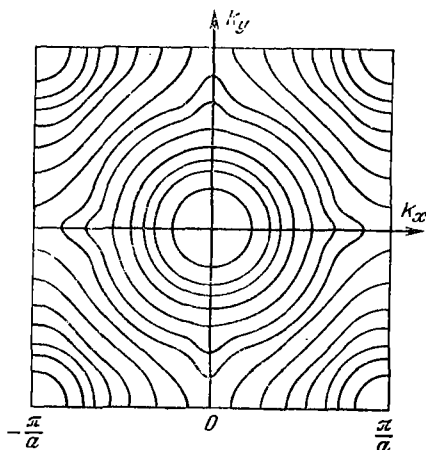


Рис. 2-12. Изоэнергетические линии для квадратной плоской решетки.

На рис. 2-11 показана зависимость $E(k)$ для трех соседних зон одномерной решетки, образованных из s -, p - и d -состояний атомов для области изменения k в пределах $\pm \pi/a$. Каждая разрешенная зона содержит N энергетических уровней, среднее расстояние между которыми по порядку величины равно 10^{-22} эВ. В точ-

как $k = 0$ и $k = \pm \pi/a$ энергия электрона принимает попеременно минимальное и максимальное значения.

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

$$E(\mathbf{k}) = \text{const} \text{ или } E(\mathbf{p}) = \text{const} \quad (2-92)$$

определяет поверхность постоянной энергии или изоэнергетическую поверхность. Согласно выражениям (2-86) и (2-89) энергия электрона у дна и потолка зоны не зависит от направления вектора \mathbf{k} , и изоэнергетические поверхности вблизи центра зоны Бриллюэна при $k = 0$ и около точек $k = \pm \pi/a$ являются сферами. По мере удаления от этих точек сферы начинают деформироваться.

Вид изоэнергетических поверхностей (кривых) для квадратной плоской решетки приведен на рис. 2-12. В центре и вблизи углов квадрата изоэнергетические кривые представляют собой окружности, которые искажаются по мере удаления от этих точек.

2-10. ДВИЖЕНИЕ ЭЛЕКТРОНОВ В КРИСТАЛЛЕ ПОД ДЕЙСТВИЕМ ВНЕШНЕГО ЭЛЕКТРИЧЕСКОГО ПОЛЯ

Рассмотрев задачу о движении электрона в периодическом поле кристалла, мы нашли, что энергетический спектр электронов имеет вид квазинепрерывных зон, разделенных промежутками запрещенных значений энергии (см. рис. 2-3).

Поместим кристалл во внешнее постоянное электрическое поле, направленное вдоль оси x , $\mathcal{E} = \mathcal{E}_x = \text{const}$. В этом случае на электрон, находящийся в периодическом поле кристалла, будет действовать внешняя сила F . Будем считать, что сила F достаточно мала и не может вызвать переход электрона между энергетическими зонами, т. е.

$$Fa \ll E_g, \quad (2-93)$$

где a — постоянная решетки, а E_g — ширина соответствующей запрещенной зоны.

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

$$U(\mathbf{r}) = -e\mathcal{E}x. \quad (2-94)$$

Уравнение Шредингера для этого случая имеет вид:

$$\left\{ -\frac{\hbar^2}{2m_0} \Delta + [V(\mathbf{r}) + U(\mathbf{r})] \right\} \Psi = E' \Psi. \quad (2-95)$$

Для упрощения задачи определения собственных значений энергии E' из уравнения (2-95) примем, что потенциал внешнего поля U мало изменяется в интервале изменения r , равном постоянной решетки, так что в области каждого атома можно считать $U = \text{const}$.

При этом мы получаем уравнение Шредингера, подобное уравнению для стационарного случая:

$$\left[-\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}) \right] \Psi = (E' - U) \Psi = E \Psi, \quad (2-96)$$

где E — энергия электрона в отсутствие внешнего поля.

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

$$E' = E + U(\mathbf{r}). \quad (2-97)$$

Тогда вместо (2-55) с учетом (2-94) можно написать:

$$E' = E_a + C + 2A (\cos k_x a + \cos k_y a + \cos k_z a) - e \mathcal{E} x. \quad (2-98)$$

Из выражения (2-98) следует, что если задать все возможные значения энергии E' , то получим график зависимости положения энергетических зон от координаты (рис. 2-13, б). Этот график можно интерпретировать как наклон зон под действием внешнего поля.

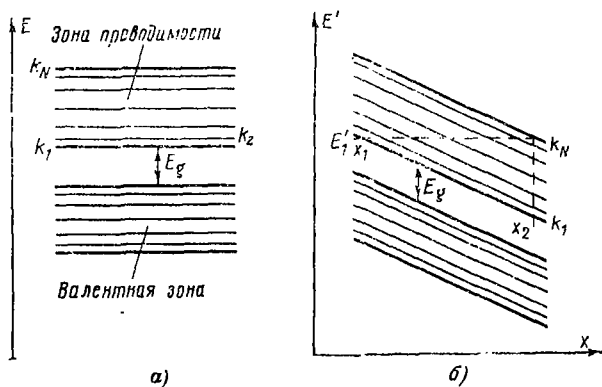


Рис. 2-13. Наклон энергетических зон во внешнем электрическом поле.

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

$$\left. \begin{aligned} E_{\text{макс}} &= E_a + C + 6A - e \mathcal{E} x; \\ E_{\text{мин}} &= E_a + C - 6A - e \mathcal{E} x. \end{aligned} \right\} \quad (2-99)$$

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

В отсутствие внешнего электрического поля электрон находится на одном из уровней энергетической зоны и обладает определенным значением волнового вектора k , например k_2 на рис. 2-13, а. Это означает, что в стационарном состоянии электрон перемещается вдоль кристалла на любые расстояния, имея постоянные значения квазиимпульса и волнового вектора.

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

$$E' = E + U = \text{const}, \quad (2-100)$$

но изменяется соотношение между кинетической и потенциальной энергией.

На рис. 2-13, б значение полной энергии показано пунктирной прямой, параллельной оси абсцисс. Но в этом случае электрон при движении по кристаллу переходит с одного уровня на другие уровни зоны, в результате чего имеет место изменение его волнового вектора от k_1 до k_N , а следовательно, и квазиимпульса от p_1 до p_N . При этом энергия $E(k)$ электрона меняется от $E_{\text{мин}}$ до $E_{\text{макс}}$. Теперь электрон, как следует из рис. 2-13, б, перемещается не по всему кристаллу, а только в пределах от x_1 до x_2 , т. е. в области $\Delta x = x_2 - x_1$, равной:

$$\Delta x = \frac{E_{\text{макс}} - E_{\text{мин}}}{e\mathcal{E}} = \frac{12A}{e\mathcal{E}}. \quad (2-101)$$

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

$$\mathbf{k} = \mathbf{k}(t); \quad \mathbf{p} = \mathbf{p}(t). \quad (2-102)$$

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

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

Волновая функция волнового пакета для одномерного случая имеет вид:

$$\Psi(x, t) = \int_{k-\Delta k}^{k+\Delta k} a(k) e^{i(kx - \omega t)} dk, \quad (2-103)$$

где ω — циклическая частота.

Движение волнового пакета характеризуется групповой скоростью $v_{\text{гп}} = \frac{d\omega}{dk}$, которая равна скорости перемещения какой-

либо точки пакета, например его максимума. Координату этого максимума можно найти из условия $\bar{x} - \frac{d\omega}{dk} t = 0$. Отсюда следует, что

$$v = \frac{\bar{x}}{t} = \frac{d\omega}{dk} = v_{\text{гр}}, \quad (2-104)$$

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

$$v = v_{\text{гр}} = \frac{d\omega}{dk}. \quad (2-105)$$

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

$$v = \frac{d\omega}{dk} = \frac{1}{\hbar} \frac{dE}{dk} = \frac{dE}{dP}, \quad (2-106)$$

где $\mathbf{P} = \hbar\mathbf{k}$ — импульс.

Движение электрона в кристалле описывается волновой функцией (2-27), которая определяется набором атомных волновых функций с разным значением k . Поскольку $k = \frac{2\pi}{aN} n$, где $n = 0, 1, \dots, (N-1)$, а $a \approx 10^{-8}$ см, $N \approx 10^{22}$ см $^{-3}$ и $aN \approx 10^{14}$, то волновую функцию Ψ можно рассматривать как совокупность плоских волн, для которых k меняется почти непрерывно. В силу этого движение электрона в кристалле можно охарактеризовать волновым пакетом, составленным из блоховских функций. Поэтому выражение (2-106) будет справедливо и для средней скорости движения электрона в кристалле

$$\langle v \rangle = v = \frac{1}{\hbar} \frac{dE}{dk} = \frac{dE}{dp} \quad (2-107)$$

или для трехмерного случая

$$\mathbf{v} = \frac{1}{\hbar} \text{grad}_{\mathbf{k}} E(\mathbf{k}) = \frac{1}{\hbar} \frac{dE}{d\mathbf{k}} = \frac{dE}{d\mathbf{p}}, \quad (2-108)$$

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

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

Вычислим среднюю квантовомеханическую скорость электрона в простой кубической решетке в приближении сильной связи. Из уравнений (2-83) и (2-108) получаем:

$$\begin{aligned} \mathbf{v} &= \frac{2aA}{\hbar} (\mathbf{i} \sin k_x a + \mathbf{j} \sin k_y a + \mathbf{k} \sin k_z a) = \\ &= \frac{2aA}{\hbar} \left(\mathbf{i} \sin \frac{a}{\hbar} p_x + \mathbf{j} \sin \frac{a}{\hbar} p_y + \mathbf{k} \sin \frac{a}{\hbar} p_z \right), \end{aligned} \quad (2-109)$$

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

Из (2-109) следует, что *средняя скорость электронов в кристалле зависит как от абсолютной величины, так и от направления волнового вектора \mathbf{k} (или квазиимпульса \mathbf{p})*.

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

$$E = -2A \cos ka; \quad (2-110)$$

$$v = \frac{2aA}{\hbar} \sin ka. \quad (2-111)$$

Эти зависимости представлены на рис. 2-14. Из рисунка видно, что на краях зоны ($k = \pm \pi/a$) $E(k)$ экстремальна, т. е. $\text{grad}_k E(k) = 0$. Следовательно, в точках экстремума энергии у нижнего и верхнего края зоны скорость электрона равна нулю. В точках же внутри зоны Бриллюэна электроны имеют наибольшие значения скорости. Средняя квантовомеханическая скорость электрона по всей зоне

$$\langle v \rangle = \int_{-\pi/a}^{\pi/a} v(k) dk = \frac{2aA}{\hbar} \int_{-\pi/a}^{\pi/a} \sin ka dk = 0. \quad (2-112)$$

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

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

$$\frac{dE}{dt} = Fv, \quad (2-113)$$

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

Так как

$$\frac{dE}{dt} = \frac{dE}{dp} \frac{dp}{dt}, \quad (2-114)$$

то из сравнения равенств (2-113) и (2-114) с учетом (2-107) следует, что

$$\frac{dp}{dt} = F \quad \text{или} \quad \frac{d\mathbf{p}}{dt} = \mathbf{F}. \quad (2-115)$$

Рис. 2-14. Зависимость энергии и скорости движения электрона от волнового вектора для одномерной решетки.

Это уравнение по форме совпадает с уравнением движения Ньютона: скорость изменения квазиимпульса равна приложенной силе. Следовательно, в периодическом поле кристалла $V(\mathbf{r})$ квазиимпульс электрона $\mathbf{p} = \hbar\mathbf{k}$ в уравнении движения (2-115) выполняет роль импульса свободного электрона.

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

$$\mathbf{p}(t) = \mathbf{p}_0 + \mathbf{F}t \quad (2-116)$$

и траекторией движения будет прямая линия, направление которой задается направлением внешней силы \mathbf{F} . Но электрон, двигаясь в \mathbf{p} -пространстве (или \mathbf{k} -пространстве), переходит от одной изоэнергетической поверхности к другой (см. рис. 2-12). Следовательно, *внешняя сила \mathbf{F} меняет не только квазиимпульс, но также и энергию электрона.*

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

$$\frac{d\mathbf{P}}{dt} = \mathbf{F}_{\text{кр}}. \quad (2-117)$$

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

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

$$\frac{d\mathbf{P}}{dt} = \mathbf{F}_{\text{кр}} + \mathbf{F}. \quad (2-118)$$

Итак, *если структура кристаллической решетки идеальна, то ее потенциальное поле $V(\mathbf{r})$ строго периодическое. В периодическом поле решетки электрон движется вдоль всего кристалла, имея постоянный квазиимпульс, а значит имея постоянную энергию. В этом случае электрон, перемещаясь по кристаллу, остается на одном и том же уровне зоны. Но поскольку $\mathbf{p} = \text{const}$, а значит $\mathbf{k} = \text{const}$, то согласно соотношению (2-111) скорость движения электрона постоянна: $v = \text{const}$, следовательно, его ускорение $a = \frac{dv}{dt} = 0$. Это значит, что в периодическом поле решетки электрон движется без ускорения. Другими словами, в строго периодическом поле решетки электрон движется как свободная частица, без сопротивления, не рассеиваясь.*

Если кристалл с идеальной структурой поместить во внешнее электрическое поле, то, как следует из равенства (2-115), движение

электрона будет подобно движению свободной частицы под действием внешней силы F . Только теперь электрон, перемещаясь в ограниченной области кристалла Δx (2-101), в энергетической зоне переходит с уровня на уровень и его энергия изменяется.

Допустим, что приобретенную во внешнем поле энергию электрон передает решетке (или дефекту). В результате этого энергия электрона будет оставаться постоянной, а это значит, что он будет обладать постоянным во времени квазиимпульсом. Из этого следует, что электрон будет оставаться на одном и том же уровне зоны и будет перемещаться вдоль кристалла на любые расстояния.

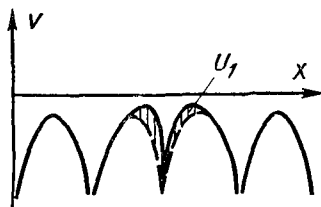


Рис. 2-15. Схематическое изображение искаженного потенциала одномерной решетки.

Рассмотрим теперь случай, когда поле кристаллической решетки не является строго периодическим. Пусть поле решетки подобно изображенному на рис. 2-15. Для участка кристалла, где нарушена идеальность структуры, поле решетки можно представить как поле периодическое $V(\mathbf{r})$, на которое наложено некоторое дополнительное силовое поле U_1 , не обладающее периодичностью решетки. Но согласно (2-115) под действием непериодической части потенциального поля U_1 квазиимпульс электрона

должен измениться. Из этого следует, что при любых нарушениях идеальной периодичности потенциального поля решетки происходит изменение квазиимпульса. Это значит, что на любых нарушениях идеальности структуры решетки происходит рассеяние электронных волн.

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

2-11. ЭФФЕКТИВНАЯ МАССА НОСИТЕЛЕЙ ЗАРЯДА

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

$$a = \frac{1}{m_0} F = -\frac{1}{m_0} e\mathcal{E}, \quad (2-119)$$

направленное так же, как и внешняя сила, т. е. против поля \mathcal{E} .

Для электрона в кристалле, находящегося во внешнем электрическом поле, учитывая равенства (2-107) и (2-115), можно записать:

$$a = \frac{dv}{dt} = \frac{\partial v}{\partial p} \frac{\partial p}{\partial t} = \frac{\partial^2 E}{\partial p^2} F. \quad (2-120)$$

Обобщая равенство (2-120) для трехмерного случая, получаем:

$$\left. \begin{aligned} a_x &= \frac{\partial^2 E}{\partial p_x^2} F_x + \frac{\partial^2 E}{\partial p_x \partial p_y} F_y + \frac{\partial^2 E}{\partial p_x \partial p_z} F_z; \\ a_y &= \frac{\partial^2 E}{\partial p_y \partial p_x} F_x + \frac{\partial^2 E}{\partial p_y^2} F_y + \frac{\partial^2 E}{\partial p_y \partial p_z} F_z; \\ a_z &= \frac{\partial^2 E}{\partial p_z \partial p_x} F_x + \frac{\partial^2 E}{\partial p_z \partial p_y} F_y + \frac{\partial^2 E}{\partial p_z^2} F_z. \end{aligned} \right\} \quad (2-121)$$

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

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

$$m_{ij}^{-1} = \begin{vmatrix} \frac{\partial^2 E}{\partial p_x^2} & \frac{\partial^2 E}{\partial p_x \partial p_y} & \frac{\partial^2 E}{\partial p_x \partial p_z} \\ \frac{\partial^2 E}{\partial p_y \partial p_x} & \frac{\partial^2 E}{\partial p_y^2} & \frac{\partial^2 E}{\partial p_y \partial p_z} \\ \frac{\partial^2 E}{\partial p_z \partial p_x} & \frac{\partial^2 E}{\partial p_z \partial p_y} & \frac{\partial^2 E}{\partial p_z^2} \end{vmatrix} = \begin{vmatrix} m_{xx}^{-1} & m_{xy}^{-1} & m_{xz}^{-1} \\ m_{yx}^{-1} & m_{yy}^{-1} & m_{yz}^{-1} \\ m_{zx}^{-1} & m_{zy}^{-1} & m_{zz}^{-1} \end{vmatrix}. \quad (2-122)$$

Величины $\frac{\partial^2 E}{\partial p_i \partial p_j} = m_{ij}^{-1}$ называются компонентами или элементами тензора. Поскольку размерность квазиимпульса совпадает с размерностью импульса, то размерность компонент тензора m_{ij}^{-1} есть размерность обратной массы, а размерность $m_{ij} = \left(\frac{\partial^2 E}{\partial p_i \partial p_j}\right)^{-1}$ есть размерность массы. Поэтому по аналогии с выражением (2-119) для свободного электрона тензор (2-122) называется тензором обратной эффективной массы.

Запишем выражение для энергии электрона в кристалле в малой окрестности экстремальной точки \mathbf{k}_0 . Для этого разложим $E(\mathbf{k})$ в окрестности точки \mathbf{k}_0 в ряд по степеням компонент разности $\mathbf{k} - \mathbf{k}_0$ и ограничимся членами второго порядка малости по $\mathbf{k} - \mathbf{k}_0$.

Поскольку точка \mathbf{k}_0 соответствует экстремуму энергии $E(\mathbf{k})$, то, учитывая, что для экстремума $\frac{\partial E}{\partial \mathbf{k}} = 0$, будем иметь:

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{1}{2} \sum_{ij} \left[\left(\frac{\partial^2 E}{\partial k_i \partial k_j} \right)_{\mathbf{k}_0} (k_i - k_{0i})(k_j - k_{0j}) \right]. \quad (2-123)$$

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

Учитывая, что

$$\frac{1}{\hbar^2} \left(\frac{\partial^2 E}{\partial k_i \partial k_j} \right)_{\mathbf{k}_0} = \left(\frac{\partial^2 E}{\partial p_i \partial p_j} \right)_{\mathbf{p}_0} = \frac{1}{m_{ij}}, \quad (2-124)$$

получаем:

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{\hbar^2}{2} \sum_{ij} \frac{1}{m_{ij}} (k_i - k_{0i})(k_j - k_{0j}) \quad (2-125)$$

или

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{1}{2} \sum_{ij} \frac{1}{m_{ij}} (p_i - p_{0i})(p_j - p_{0j}). \quad (2-126)$$

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

$$v_i = \frac{\partial E}{\partial p_i} = \sum_j \frac{p_j - p_{0j}}{m_{ij}}, \quad (2-127)$$

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

$$\mathbf{v} = \frac{1}{\mathbf{m}} (\mathbf{p} - \mathbf{p}_0). \quad (2-128)$$

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

Тензор (2-122) симметричен относительно главной диагонали, т. е. $m_{ij}^{-1} = m_{ji}^{-1}$, так как $\frac{\partial^2 E}{\partial p_i \partial p_j} = \frac{\partial^2 E}{\partial p_j \partial p_i}$. Выбрав соответствующую систему координат, можно свести симметричный тензор к диагональному виду, когда отличны от нуля только члены, расположенные на главной диагонали:

$$\mathbf{m}^{-1} = \begin{vmatrix} m_{xx}^{-1} & 0 & 0 \\ 0 & m_{yy}^{-1} & 0 \\ 0 & 0 & m_{zz}^{-1} \end{vmatrix} = \begin{vmatrix} m_1^{-1} & 0 & 0 \\ 0 & m_2^{-1} & 0 \\ 0 & 0 & m_3^{-1} \end{vmatrix}, \quad (2-129)$$

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

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

$$\{\mathbf{m}^{-1}\}^{-1} = \mathbf{m} = \begin{vmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{vmatrix}. \quad (2-130)$$

Величины m_i ($i = 1, 2, 3$) называются компонентами тензора эффективной массы.

Если учесть компоненты тензора эффективной массы (2-130), то энергия электрона вблизи экстремальной точки \mathbf{k}_0 запишется в виде

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{\hbar^2 (k_x - k_{0x})^2}{2m_1} + \frac{\hbar^2 (k_y - k_{0y})^2}{2m_2} + \frac{\hbar^2 (k_z - k_{0z})^2}{2m_3} \quad (2-131)$$

или

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{(p_x - p_{0x})^2}{2m_1} + \frac{(p_y - p_{0y})^2}{2m_2} + \frac{(p_z - p_{0z})^2}{2m_3}, \quad (2-132)$$

а скорость

$$v_i = \frac{p_i - p_{0i}}{m_i} = \frac{\partial E}{\partial p_i}. \quad (2-133)$$

Из равенства (2-132) следует, что поверхности равной энергии в \mathbf{p} -пространстве $E(\mathbf{p}) - E(\mathbf{p}_0) = \text{const}$ вблизи экстремума энергии представляют собой эллипсоиды, а m_1, m_2, m_3 — эффективные массы вдоль основных осей эллипсоидальных поверхностей энергии.

Запишем уравнение эллипсоида в канонической форме:

$$\frac{(p_x - p_{0x})^2}{a^2} + \frac{(p_y - p_{0y})^2}{b^2} + \frac{(p_z - p_{0z})^2}{c^2} = 1, \quad (2-134)$$

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

$$\begin{aligned} a^2 &= 2m_1 [E(\mathbf{p}) - E(\mathbf{p}_0)]; \\ b^2 &= 2m_2 [E(\mathbf{p}) - E(\mathbf{p}_0)]; \\ c^2 &= 2m_3 [E(\mathbf{p}) - E(\mathbf{p}_0)]. \end{aligned} \quad (2-135)$$

Как известно, градиент направлен по нормали к поверхности уровня. Поскольку скорость движения электрона есть градиент энергии в пространстве квазиимпульса (2-133), следовательно, она направлена по нормали к изоэнергетической поверхности. Поэтому для эллипсоидальных поверхностей энергии (2-134) в общем случае направления векторов скорости \mathbf{v} и квазиимпульса \mathbf{p} не совпадают (рис. 2-16 и 2-17, а). Они будут совпадать по направлению только вдоль осей эллипсоидов. При этом, если принять $E(\mathbf{p}_0) = E_0$, то

$$p_i - p_{0i} = \sqrt{2m_i (E - E_0)}; \quad (2-136)$$

$$v = \frac{p_i - p_{0i}}{m_i} = \sqrt{\frac{2(E - E_0)}{m_i}}. \quad (2-137)$$

Из соотношения (2-137) следует, что при одной и той же энергии скорость вдоль осей эллипсоида обратно пропорциональна корню из соответствующей компоненты эффективной массы. Но компонента эффективной массы согласно соотношению (2-135) определяет величину оси эллипсоида. Поэтому, чем сильнее вытянут эллипсоид, т. е. чем больше m_i , тем меньше скорость движения в этом направлении.

Если симметрия кристалла такова, что две главные оси тензора физически эквивалентны (например, x и y), соответствующие ком-

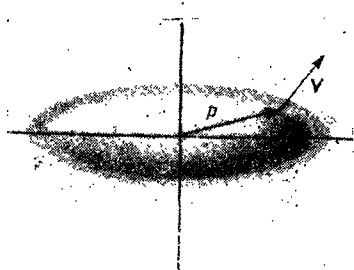


Рис. 2-16. Скорость и квазиимпульс электрона на эллипсоидальной поверхности. Вектор \mathbf{v} нормален к поверхности; векторы \mathbf{v} и \mathbf{p} в общем случае неколлинеарны.

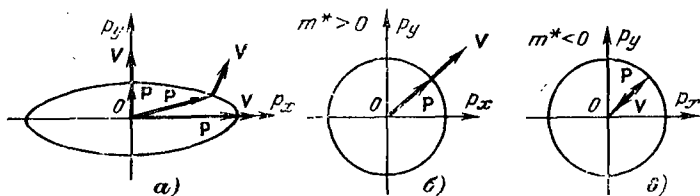


Рис. 2-17. Направления радиус-вектора и нормали к поверхности энергии. a — эллипсоидальные поверхности энергии; $б$ — сферические поверхности энергии, $m^* > 0$; $в$ — сферические поверхности энергии, $m^* < 0$.

поненты эффективной массы равны, т. е. $m_1 = m_2$, а следовательно, $a = b$. Уравнение изоэнергетической поверхности в этом случае упрощается:

$$E(\mathbf{p}) = E(\mathbf{p}_0) + \frac{(p_x - p_{0x})^2 + (p_y - p_{0y})^2}{2m_1} + \frac{(p_z - p_{0z})^2}{2m_3}. \quad (2-138)$$

Изоэнергетическая поверхность, описываемая уравнением (2-138), является эллипсоидом вращения, у которого ось вращения совпадает с осью p_z . Так как m_3 — эффективная масса вдоль оси симметрии, т. е. вдоль оси вращения эллипсоида, то она получила название продольной эффективной массы и обозначается m_l . Величина $m_1 = m_2$ — эффективная масса поперек осей симметрии, а поэтому она называется поперечной эффективной массой и обозначается m_t . Если $m_1 < m_3$, то эллипсоид вращения вытянут вдоль оси вращения, при этом он вытянут тем

больше, чем больше отношение эффективных масс m_3/m_1 . Если $m_1 \gg m_3$, то эллипсоид сжат вдоль оси вращения.

Для кристаллов, обладающих кубической симметрией, все главные оси эквивалентны. При этом $m_1 = m_2 = m_3 = m^*$ и тензор эффективной массы вырождается в скаляр (тензор нулевого ранга). В этом случае изоэнергетические поверхности описываются уравнением

$$E(\mathbf{p}) - E(\mathbf{p}_0) + \frac{p^2}{2m^*} = \text{const} \quad (2-139)$$

и представляют собой сферы (рис. 2-17, б, в). Выражение для эффективной массы имеет вид:

$$\frac{1}{m^*} = \frac{\partial^2 E}{\partial p^2} = \frac{1}{\hbar^2} \frac{\partial^2 E}{\partial k^2}. \quad (2-140)$$

Следовательно, для кристалла кубической симметрии эффективная масса носителей заряда обратно пропорциональна кривизне изоэнергетических поверхностей в k -пространстве (или p -пространстве). В том случае, когда электрон находится в окрестности минимума энергии, т. е. в окрестности дна зоны проводимости,

$$\frac{\partial^2 E}{\partial k^2} > 0 \quad \text{и} \quad m^* = \text{const} > 0, \quad (2-141)$$

т. е. электроны ведут себя как отрицательно заряженные частицы с положительной эффективной массой. При этом согласно выражениям (2-120) и (2-140) получаем:

$$\mathbf{F} = m^* \mathbf{a}, \quad (2-142)$$

т. е. ускорение направлено по направлению внешней силы, а значит против внешнего поля \mathcal{E} .

Так как ускорение $\mathbf{a} = \frac{d\mathbf{v}}{dt}$, а $\mathbf{F} = \frac{d\mathbf{p}}{dt}$, то квазимпульс электрона в кубическом кристалле равен:

$$\mathbf{p} = m^* \mathbf{v}, \quad (2-143)$$

т. е. скорость совпадает по направлению с квазимпульсом (рис. 2-17, б).

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

Из равенства (2-142) следует, что ускорение электрона в кристалле сообщает только внешняя сила. Действие поля решетки проявляется в том, что при наличии внешней силы движение электрона определяется не его обычной массой m_0 , а эффективной массой m^* .

В окрестности максимума энергии, т. е. в окрестности вершины валентной зоны,

$$\frac{\partial^2 E}{\partial k^2} < 0 \quad \text{и} \quad m^* = \text{const} < 0, \quad (2-144)$$

а следовательно, скорость направлена против квазиимпульса (рис. 2-17, в). Из равенства (2-144) следует, что направление ускорения электрона, находящегося у верхнего края зоны, противоположно направлению действующей на него внешней силы:

$$\mathbf{a} = - \frac{\mathbf{F}}{m^*}. \quad (2-145)$$

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

$$\mathbf{F} = -e\mathcal{E}, \quad (2-146)$$

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

$$\mathbf{a} = \frac{-|e|\mathcal{E}}{-|m^*|} = \frac{e\mathcal{E}}{m^*}, \quad (2-147)$$

т. е. ускорение направлено по полю \mathcal{E} . Такой носитель заряда в окрестности вершины валентной зоны ведет себя как частица с положительным зарядом и положительной эффективной массой и носит название дырки.

По мере удаления от краев зоны зависимость $E(\mathbf{k})$ отклоняется от квадратичного закона. В точках перегиба $k = \pm \pi/2a$ кривой $E(\mathbf{k})$ (рис. 2-14)

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

Из этого следует, что применение соотношения (2-115) для описания движения носителей заряда в кристалле под действием внешнего поля возможно только для носителей заряда, находящихся вблизи дна или потолка энергетической зоны.

На практике наиболее важными оказываются два случая заполнения энергетической зоны.

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

2. Зона почти заполнена электронами. В такой зоне имеется некоторое число свободных состояний. Состояния, свободные от электронов, находятся в самой верхней части зоны, т. е. дырки располагаются у потолка зоны, где их энергия минимальна. Следовательно, энергия дырок отсчитывается в направле-

нии, противоположном отсчету энергии для электрона. В этом случае $m_p^* = \text{const} > 0$.

При температуре абсолютного нуля валентная зона полупроводника заполнена электронами, а зона проводимости — свободная. При повышении температуры небольшая часть электронов из валентной зоны в результате теплового возбуждения переходит в зону проводимости. При этом реализуется рассмотренный выше случай — наличие небольшого количества электронов у дна зоны проводимости и дырок у верхнего края валентной зоны. Так как ширина зоны проводимости больше ширины валентной зоны, а 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, но они имеют и разные величины эффективных масс.

2-12. ЦИКЛОТРОННЫЙ РЕЗОНАНС

Существует несколько экспериментальных методов определения эффективной массы носителей заряда. Одним из них является метод циклотронного резонанса. Существо метода состоит в следующем. Пусть полупроводник находится в постоянном магнитном поле с индукцией $\mathbf{B} = \text{const}$. На электрон, движущийся со скоростью \mathbf{v} , будет действовать сила Лоренца

$$\mathbf{F} = -e [\mathbf{vB}]. \quad (2-148)$$

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

$$\frac{m^* v^2}{r} = evB, \quad (2-149)$$

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

$$\omega_c = \frac{v}{r} = \frac{eB}{m^*}. \quad (2-150)$$

Если одновременно полупроводник поместить в слабое высокочастотное электромагнитное поле, колеблющееся в плоскости, перпендикулярной к \mathbf{B} , то при приближении частоты поля ω к ω_c электрон будет сильно поглощать энергию электромагнитного поля. Это явление резонансного поглощения излучения называется ц и к л о т р о н н ы м р е з о н а н с о м. Из измерений резонансной частоты поглощения радиоизлучения определяется ω_c , а затем

по соотношению (2-150) 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 равной энергии.

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

$$\tau \gg \frac{1}{\omega_c}, \quad (2-151)$$

т. е. за время τ электрон должен успеть совершить по меньшей мере один оборот. Чем больше электрон делает оборотов, тем резче,

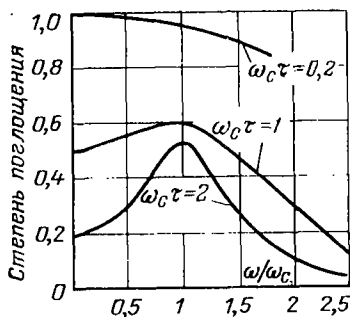


Рис. 2-18. Степень поглощения при циклотронном резонансе в зависимости от напряженности постоянного магнитного поля (в единицах ω/ω_c).

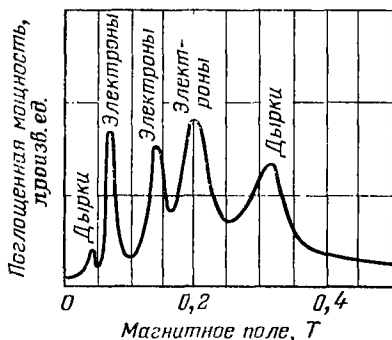


Рис. 2-19. Поглощение при циклотронном резонансе в германии.

острее проявляется резонанс. На рис. 2-18 изображены кривые поглощения электромагнитного поля в зависимости от ω/ω_c для трех значений $\omega_c \tau$ — 0,2; 1 и 2.

Время свободного пробега τ носителя заряда тем больше, чем меньше в кристалле структурных несовершенств. Поэтому эксперименты по циклотронному резонансу необходимо проводить при очень низких температурах (обычно при 4,2 К — температуре жидкого гелия) и на особо чистых кристаллах. Обычно в опытах задается частота радионизлучения, а резонанс достигается изменением индукции магнитного поля. На рис. 2-19 приведены пики поглощения при циклотронном резонансе в германии.

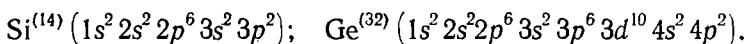
2-13. ЗОННАЯ СТРУКТУРА НЕКОТОРЫХ ПОЛУПРОВОДНИКОВ

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

Кремний и германий имеют кристаллическую структуру типа алмаза, представляющую собой две 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. 2-20, где значения составляющих волнового вектора (k_x, k_y, k_z) даны в единицах $2\pi/a$ (a — ребро куба решетки).

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

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



У них последняя оболочка не заполнена, в ней в p -состоянии имеется два электрона с параллельными спинами. Поскольку зона проводимости и валентная зона кремния и германия включают p -состояние (см. рис. 2-7), для которого в кристалле вырождение снимается, то каждая из них представляет собой наложение трех различных зон. На рис. 2-21 они представлены тремя ветвями $E(\mathbf{k})$. Эта зависимость неодинакова для разных кристаллографических направлений.

Одна из ветвей $E(\mathbf{k})$ зоны проводимости как у кремния, так и у германия (рис. 2-21) лежит значительно ниже других. Положение абсолютного минимума энергии определяет дно зоны проводимости. Минимумы энергии называют также д о л и н а м и.

Абсолютный минимум зоны проводимости у германия расположен в направлении осей $[111]$ (рис. 2-21), поэтому имеется восемь эквивалентных минимумов энергии, координаты которых в единицах $2\pi/a$ имеют вид $(1/2, 1/2, 1/2)$ и лежат на границе зоны Бриллюэна (на рис. 2-20 они показаны жирной точкой). *Изоэнергетические поверхности около абсолютных минимумов энергии (или долин) представляют собой эллипсоиды вращения относительно боль-*

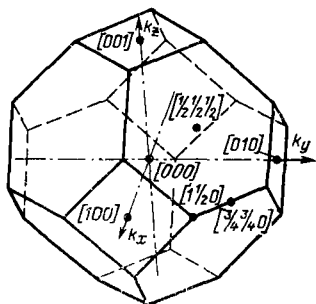


Рис. 2-20. Первая зона Бриллюэна полупроводника типа алмаза.

шой полуоси. У германия она совпадает с направлением осей [111] и на первую зону Бриллюэна приходится половина каждого эллипсоида энергии, а следовательно, в германии имеется не восемь, а только четыре полных эллипсоида энергии (рис. 2-22).

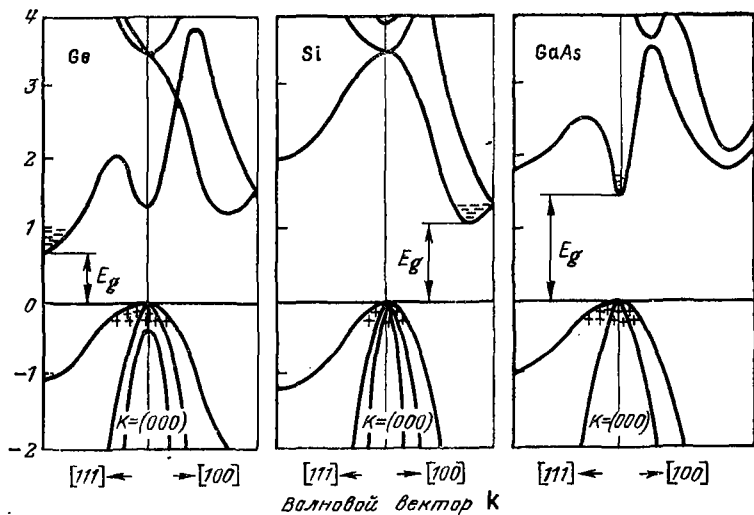


Рис. 2-21. Энергетическая зонная структура германия, кремния и арсенида галлия.

E_g — ширина запрещенной зоны; знаком плюс обозначены дырки в валентных зонах; знаком минус — электроны в зонах проводимости.

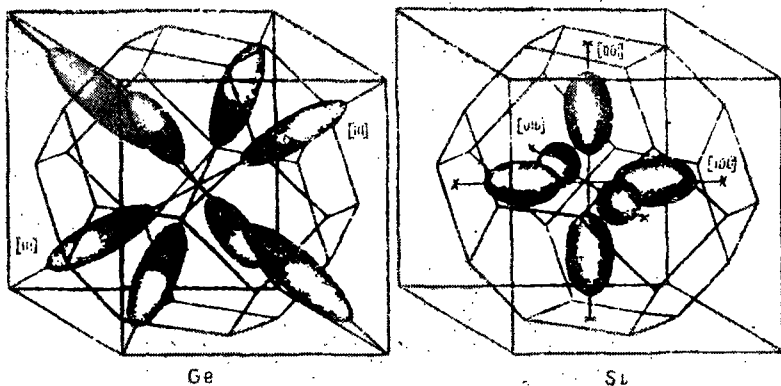


Рис. 2-22. Формы поверхностей постоянной энергии в зонах проводимости германия и кремния.

Абсолютный минимум зоны проводимости у кремния лежит в направлении осей [100] недалеко от границы зоны Бриллюэна (рис. 2-21). Поэтому у кремния имеется шесть эквивалентных минимумов энергии, а следовательно, у кремния на первую зону Бриллюэна приходится шесть эллипсоидальных поверхностей постоянной энергии, вытянутых вдоль осей [100] (рис. 2-22). Центры эллипсоидов расположены на расстоянии трех четвертей от центра зоны Бриллюэна.

У арсенида галлия (рис. 2-21) и антимонида индия (рис. 2-25) абсолютный минимум зоны проводимости лежит при $k = 0$, поэтому у них поверхностями постоянной энергии являются сферы с центром в центре зоны Бриллюэна и эффективные массы электронов соответственно равны $0,07 m_0$ и $0,013 m_0$.

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

$$E(k) = E(k_0) + \frac{\hbar^2 [(k_x - k_{0x})^2 + (k_y - k_{0y})^2]}{2m_1} + \frac{\hbar^2 (k_z - k_{0z})^2}{2m_3}. \quad (2-152)$$

Значения компонент тензора эффективной массы электрона, определенные по циклотронному резонансу, составляют у германия $m_1 = m_2 = m_t = 0,082 m_0$; $m_3 = m_l = 1,60 m_0$, а у кремния $m_1 = m_2 = m_t = 0,19 m_0$; $m_3 = m_l = 0,97 m_0$. Отношение m_3/m_1 характеризует анизотропию свойств изоэнергетических поверхностей. Для германия это отношение равно 19,3, а у кремния — 5,16.

Минимальное расстояние между дном зоны проводимости и потолком валентной зоны называется шириной запрещенной зоны. У кремния и германия экстремумы энергии электронов и дырок лежат в различных точках зоны Бриллюэна (рис. 2-20). При комнатной температуре и нормальном атмосферном давлении ширина запрещенной зоны у германия достаточной степени чистоты равна 0,67 эВ, у кремния она составляет 1,11 эВ, арсенида галлия 1,43 эВ и у антимонида индия 0,17 эВ.

В приближении сильной связи валентная зона германия и кремния образуется из атомных p -орбиталей, а поэтому валентная зона их включает три подзоны, которые условно можно обозначить как $P_{3/2}^{(3/2)}$ -, $P_{3/2}^{(1/2)}$ - и $P_{1/2}^{(1/2)}$ -зоны. Для всех трех зон максимум энергии находится в центре зоны Бриллюэна $k = 0$ (рис. 2-21). При этом обе $P_{3/2}$ -зоны вырождены при $k = 0$. Для $P_{1/2}$ -зоны вырождение частично снижается за счет спин-орбитального взаимодействия, в результате этого она расположена ниже на величину E_{s0} .

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

$$E_{1,2}(k) = E(0) - \frac{\hbar^2}{2m_0} \left[Ak^2 \pm \sqrt{B^2 k^4 + C^2 (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2)} \right], \quad (2-153)$$

где A, B, C — безразмерные константы, равные соответственно

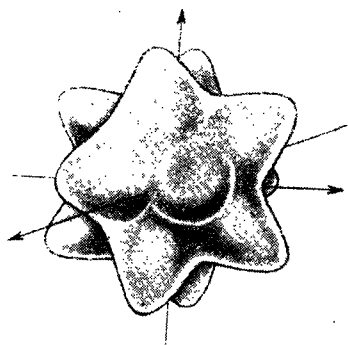


Рис. 2-23. Поверхность постоянной энергии тяжелых дырок в кремнии.

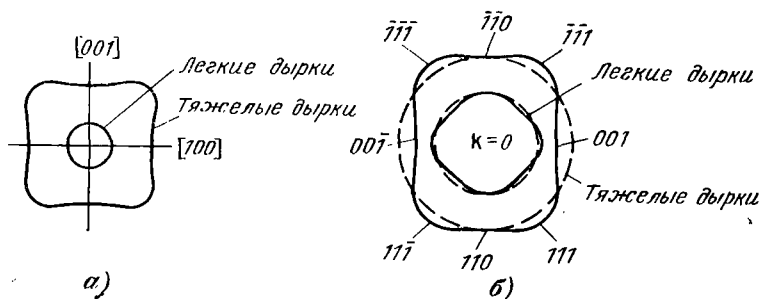


Рис. 2-24. Изоэнергетические поверхности в валентных зонах германия (а) и кремния (б).

для германия 13; 8,9 и 10,3 и для кремния 4,1; 1,6 и 3,3. Поверхности равной энергии (2-153) представляют собой гофрированные поверхности (рис. 2-23 и 2-24).

Усреднение зависимости (2-153) по различным направлениям в k -пространстве позволяет заменить гофрированную поверхность сферической (рис. 2-24), для которой $E_{1,2}$ может быть представлено выражением

$$E_{1,2} = E(0) - \frac{\hbar^2 k^2}{2m_0} \left[A \pm \sqrt{B^2 + \frac{1}{5} C^2} \right]. \quad (2-154)$$

В этом случае эффективная масса является скалярной величиной

$$m_p^* = \frac{m_0}{A \pm \sqrt{B^2 + \frac{1}{5} C^2}} \quad (2-155)$$

и должны существовать два вида дырок:
тяжелые

$$m_{pt}^* = \frac{m_0}{A - \sqrt{B^2 + \frac{1}{5} C^2}} \quad (2-156)$$

и легкие

$$m_{рл}^* = \frac{m_0}{A + \sqrt{B^2 + \frac{1}{5} C^2}}. \quad (2-157)$$

Согласно рис. 2-21 зоной тяжелых дырок является широкая зона (с малым значением $\frac{\partial^2 E}{\partial k^2}$), а зона легких дырок — более узкая зона (с большим значением $\frac{\partial^2 E}{\partial k^2}$).

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

для германия

$$m_{рг}^* = 0,34m_0; \quad m_{рл}^* = 0,04m_0,$$

для кремния

$$m_{рг}^* = 0,50m_0; \quad m_{рл}^* = 0,16m_0,$$

для арсенида галлия

$$m_{рг} = 0,50m_0; \quad m_{рл} = 0,12m_0.$$

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

$$E_3(k) = E(0) - E_{s0} - \frac{\hbar^2 k^2}{2m_0} A, \quad (2-158)$$

где E_{s0} — энергия спин-орбитального взаимодействия, для германия равная 0,28 эВ, для кремния 0,035 эВ. Поверхности равной энергии третьей валентной зоны согласно (2-158) представляют собой сферы. Поэтому эффективная масса третьего типа дырок является скалярной величиной и равна $m_p^* = m_0/A$.

Как показывают эксперименты, зона проводимости антимонида индия обладает сферической симметрией и минимум ее находится в центре зоны Бриллюэна (рис. 2-25). Вблизи минимума энергия подчиняется параболическому закону, а кривизна зоны очень велика. В силу этого величина эффективной массы электронов, находящихся на дне этой зоны, очень мала и составляет 0,013 m_0 по данным измерений из циклотронного резонанса. Для больших энергий зона проводимости имеет непараболическую форму, и кривизна ее уменьшается. Это подтверждают данные рис. 2-26, из которого видно, что энергия непропорциональна k^2 . В этом случае энергия вблизи минимума так называемой нестандартной зоны имеет вид:

$$E = E_c + \frac{\hbar^2 k^2}{2m_n^*} (1 - \alpha k^2), \quad (2-159)$$

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

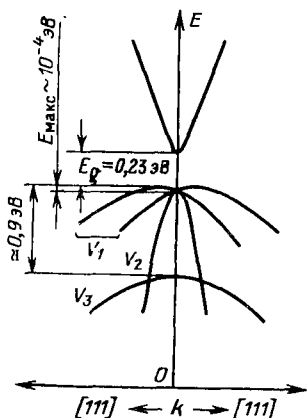


Рис. 2-25. Энергетическая зонная структура антимионда индия.

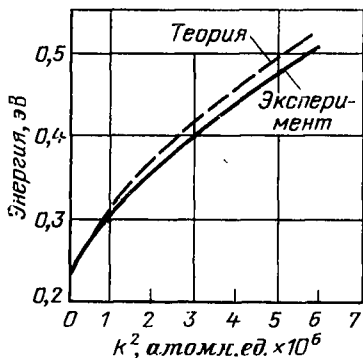


Рис. 2-26. Отклонение формы зоны проводимости в антимионде индия от параболической.

Валентная зона антимионда индия, зонная структура которого для направления $[111]$ изображена на рис. 2-25, включает три ветви $E(k)$. Зона тяжелых дырок V_1 ($m_{pг}^* = 0,6m_0$) и зона легких дырок V_2 ($m_{pл}^* = 0,012m_0$) вырождены при $k = 0$. Предполагается, что максимумы валентной зоны находятся не точно в центре зоны Бриллюэна, а смещены от него на малое расстояние. Третья зона V_3 отщеплена вследствие спин-орбитального взаимодействия.

2-14. МЕТОД ЭФФЕКТИВНОЙ МАССЫ

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

$$\left[-\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}) + U(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2-160)$$

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

Из предыдущего рассмотрения следует, что уравнение движения электрона в периодическом поле решетки при наличии внеш-

него поля аналогично классическому уравнению движения свободного электрона, если его массу m_0 заменить в общем случае тензором эффективной массы m^* .

Уравнение Шредингера для свободного электрона имеет вид:

$$-\frac{\hbar^2}{2m_0} \Delta \Psi_0(\mathbf{r}) = E \Psi_0(\mathbf{r}). \quad (2-161)$$

Тогда для электрона в кристалле уравнение Шредингера можно записать аналогично (2-161):

$$-\frac{\hbar^2}{2m^*} \Delta \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2-162)$$

Необходимо выяснить, при каких условиях уравнение (2-162) дает решения, подобные решениям уравнения Шредингера с периодическим потенциалом $V(\mathbf{r})$:

$$\left[-\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}). \quad (2-163)$$

Известно, что уравнению (2-163) удовлетворяют блоховские функции $\Psi(\mathbf{r}) = U_{\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\mathbf{r}}$, а энергетический спектр электрона вблизи экстремальной точки ($\mathbf{k} = \mathbf{k}_0$) описывается выражением:

$$E = E(\mathbf{k}_0) + \frac{\hbar^2}{2} \left[\frac{(k_x - k_{0x})^2}{m_1} + \frac{(k_y - k_{0y})^2}{m_2} + \frac{(k_z - k_{0z})^2}{m_3} \right]. \quad (2-164)$$

В уравнении (2-164) предполагается, что тензор обратной эффективной массы приведен к диагональному виду, и поэтому соотношение (2-162) может быть записано в виде

$$\sum_{i=1}^3 \frac{1}{m_i} \frac{\partial^2 \Psi}{\partial x_i^2} + \frac{2E}{\hbar^2} \Psi = 0. \quad (2-165)$$

Нетрудно убедиться, что решением этого уравнения будет:

$$\Psi = A e^{i(\alpha x + \beta y + \gamma z)}, \quad (2-166)$$

при этом

$$E = \frac{\hbar^2}{2} \left[\frac{\alpha^2}{m_1} + \frac{\beta^2}{m_2} + \frac{\gamma^2}{m_3} \right]. \quad (2-167)$$

Сравнивая выражения (2-164) и (2-167), можно сделать вывод, что если положить $\alpha = k_x - k_{0x}$; $\beta = k_y - k_{0y}$; $\gamma = k_z - k_{0z}$, а начало отсчета энергии вести от экстремальной точки, т. е. $E(\mathbf{k}_0) = 0$, то решение уравнения (2-162) дает тот же спектр энергии электрона в кристалле вблизи экстремальной точки $\mathbf{k} = \mathbf{k}_0$, что и уравнение (2-163).

Таким образом, для решения задачи определения спектра энергии электрона кристалла при наличии внешних полей можно вместо уравнения (2-160) решать уравнение

$$\left[-\frac{\hbar^2}{2m^*} \Delta + U(\mathbf{r}) \right] \Psi(\mathbf{r}) = E \Psi(\mathbf{r}), \quad (2-168)$$

в которое уже не входит неизвестный потенциал поля решетки $V(\mathbf{r})$. Подобный метод решения получил название метода эффективной массы. Большое достоинство этого метода состоит в том, что значения компонентов тензора эффективной массы вблизи экстремальной точки можно определить из эксперимента (например, из опытов по циклотронному резонансу). То, что метод эффективной массы справедлив лишь для электронов вблизи экстремума энергии, не накладывает серьезных ограничений при решении большинства задач теории полупроводников, поскольку концентрация электронов (или дырок) в невырожденных полупроводниках много меньше числа состояний в зоне и, следовательно, электроны (или дырки) располагаются в разрешенных зонах в состояниях с минимальной энергией.

2-15. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ПРИМЕСНЫХ СОСТОЯНИЙ

До сих пор 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 окажется сравнительно слабо связанным с ионом мышьяка силами кулоновского взаимодействия. Его поведение подобно состоянию 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 при описании его движения необходимо пользоваться не действительной массой, а эффективной массой 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 $U(\mathbf{r})$ электрона примеси следует ввести величину ϵ , тогда энергия будет равна:

$$U(\mathbf{r}) = -\frac{Ze^2}{4\pi\epsilon\epsilon_0 r}, \quad (2-169)$$

где ϵ_0 — диэлектрическая проницаемость вакуума.

Соотношение (2-169) 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 пятого электрона атома мышьяка, не участвующего в образовании валентных связей, согласно выражению (2-161) примет вид:

$$\left(-\frac{\hbar^2}{2m^*} \Delta - \frac{Ze^2}{4\pi\epsilon_0 r} \right) \Psi_a = E_n \Psi_a. \quad (2-170)$$

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

$$E_n = E_c - \frac{m^* Z^2 e^4}{8\hbar^2 \epsilon^2 \epsilon_0^2} \frac{1}{n^2}, \quad (2-171)$$

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

Эту формулу можно записать следующим образом:

$$E_n = E_c - \frac{m_0 Z^2 e^4}{8\hbar^2 \epsilon_0^2} \left(\frac{m^*}{m_0} \right) \frac{1}{\epsilon^2 n^2}. \quad (2-172)$$

Если подставить численные значения e , m_0 , \hbar , ϵ_0 и выразить энергию в электронвольтах, то получим:

$$E_n = E_c - \frac{13,52 Z^2}{\epsilon^2} \left(\frac{m^*}{m_0} \right) \frac{1}{n^2} = E_c - \frac{E_d}{n^2}, \quad (2-173)$$

где 13,52 соответствует значению энергии ионизации атома водорода в электронвольтах, а

$$E_d = \frac{13,52 Z^2}{\epsilon^2} \left(\frac{m^*}{m_0} \right) \quad (2-174)$$

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

Из (2-173) следует, что *уровень энергии донорной примеси, соответствующий основному состоянию ($n = 1$), лежит в запрещенной зоне полупроводника ниже дна зоны проводимости E_c на величину E_d .*

Энергия ионизации донорной примеси, равная E_d , согласно равенству (2-174) в ϵ^2 раз меньше энергии ионизации атома водорода. Кроме того, она зависит от Z^2 . Это означает, что уровень двукратно заряженного иона донорной примеси должен быть расположен в запрещенной зоне ниже уровня однократно заряженного иона этой примеси.

Если положить $m^* = 0,25 m_0$, то в случае германия, у которого $\epsilon = 15,8$, энергия ионизации пятивалентных атомов примеси должна быть примерно 0,01 эВ. Значения энергии ионизации донорных примесей, полученные экспериментально, достаточно хорошо согласуются с рассчитанным значением.

Расчет радиуса орбиты электрона примеси по водородоподобной модели дает формулу

$$a = \frac{\epsilon_0 h^2}{\pi e^2 m^*} \epsilon n^2 = a_0 \left(\frac{m_0}{m^*} \right) \epsilon n^2, \quad (2-175)$$

где $a_0 = \frac{\epsilon_0 h^2}{\pi e^2 m_0} = 0,053$ нм (0,53 Å) — радиус первой боровской орбиты атома водорода.

Радиус первой орбиты электрона донорной примеси в германии составляет: $a_1 = 64$, $a_0 = 3,4$ нм (34 Å). Если учесть, что постоянная решетки германия равна 0,562 нм (5,62 Å), то эта орбита охватывает приблизительно 200 узлов решетки. Это подтверждает возможность рассмотрения движения электрона примеси в среде с постоянной диэлектрической проницаемостью.

Рассмотрим теперь 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 атома кремния, расположенного рядом с акцепторной примесью. 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

$$E_p = E_v + \frac{m^* e^4 Z^2}{8 h^2 \epsilon^2 \epsilon_0^2 n^2} = E_v + \frac{E_a}{n^2}, \quad (2-176)$$

где

$$E_a = \frac{13,52 Z^2}{\epsilon^2} \left(\frac{m^*}{m_0} \right), \quad (2-177)$$

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, такие же значения энергии ионизации акцепторов получены экспериментально.

Таким образом, «мелкие» уровни донорной и акцепторной примесей, которые образуют атомы элементов V и III групп в ковалентных полупроводниках типа германия, достаточно хорошо описываются водородоподобной моделью. 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. Эти элементы могут входить в решетку в виде атома (донора D или акцептора A), однократно заряженного иона ($D^{(1)}$ или $A^{(1)}$) и двукратно заряженного иона ($D^{(2)}$ или $A^{(2)}$). Кроме того, одна и та же примесь может образовывать как донорные, так и акцепторные уровни.

На рис. 2-27 показаны измеренные значения энергии ионизации для различных примесей в германии и кремнии. Примесные уровни, лежащие выше середины запрещенной зоны, являются донорными. Донорные уровни, лежащие ниже середины запрещенной зоны, отмечены значком *D*. Энергия ионизации примесей, уровни кото-

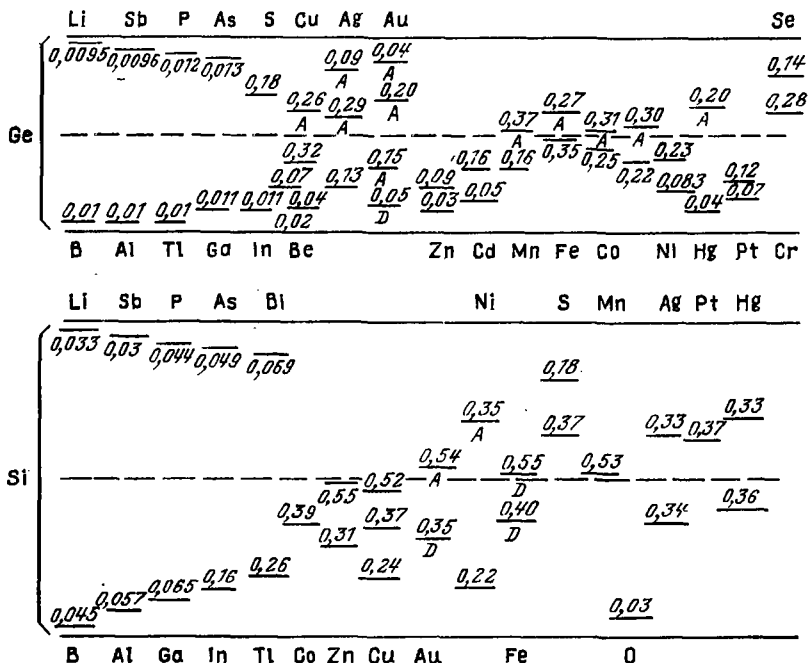


Рис. 2-27. Измеренные энергии ионизации для различных примесей в германии и кремнии.

рых лежат ниже середины запрещенной зоны, отсчитывается от потолка валентной зоны. Такие уровни являются акцепторными. Акцепторные уровни, лежащие выше середины запрещенной зоны, отмечены значком *A*.

Глава третья

КОЛЕБАНИЯ АТОМОВ КРИСТАЛЛИЧЕСКОЙ РЕШЕТКИ

3-1. ОДНОМЕРНЫЕ КОЛЕБАНИЯ ОДНОРОДНОЙ СТРУНЫ

Как известно, атомы в кристалле не неподвижны, они совершают хаотические тепловые колебания относительно своих положений равновесия. Эти колебания нарушают строгую периодичность поля решетки. Рассеяние носителей заряда на этих нерегу-

лярностях решетки оказывает существенное влияние на многие физические явления в полупроводниках (см. гл. 5 и 6).

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

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

$$e(x) = \frac{\partial u}{\partial x}, \quad (3-1)$$

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

$$e(x + \Delta x) = e(x) + \frac{\partial^2 u}{\partial x^2} \Delta x. \quad (3-2)$$

Если сила, вызывающая деформацию, равна F , то упругая постоянная (модуль Юнга) T_{II} определяется отношением $T_{II} = F/e$. Поэтому результирующая сила $F_{\Delta x}$, действующая на элемент струны Δx , равна:

$$F_{\Delta x} = T_{II} [e(x + \Delta x) - e(x)] = T_{II} \frac{\partial^2 u}{\partial x^2} \Delta x. \quad (3-3)$$

С другой стороны, эту силу можно определить как произведение массы элемента струны $\rho \Delta x$ на ускорение $\frac{\partial^2 u}{\partial t^2}$, т. е.

$$F_{\Delta x} = \rho \frac{\partial^2 u}{\partial t^2} \Delta x. \quad (3-4)$$

В результате получим уравнение свободных колебаний струны:

$$\frac{\partial^2 u}{\partial t^2} = \frac{T_{II}}{\rho} \frac{\partial^2 u}{\partial x^2}. \quad (3-5)$$

Здесь

$$\sqrt{\frac{T_{II}}{\rho}} = v_{зв} \quad (3-6)$$

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

Решение уравнения колебания струны имеет вид бегущей монокроматической волны:

$$u = A \sin 2\pi \left(\frac{x}{\lambda} - vt \right) = A \sin(qx - \omega t) = A e^{i(qx - \omega t)}, \quad (3-7)$$

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

Для упругой волны, распространяющейся в непрерывной струне, частота колебаний

$$\omega = 2\pi\nu = 2\pi \frac{v_{зв}}{\lambda} = v_{зв}q \quad (3-8)$$

пропорциональна волновому числу q . При этом абсолютная величина волнового числа q может меняться от 0 до ∞ , а следовательно, частота колебаний меняется непрерывно от нуля до бесконечного значения (см. рис. 3-2). Каждому значению q соответствует волна, характеризующаяся определенной частотой колебаний.

3-2. КОЛЕБАНИЯ ОДНОАТОМНОЙ ЛИНЕЙНОЙ ЦЕПочки

Рассмотрим теперь линейную цепочку одинаковых атомов массы m , расположенных на расстоянии a друг от друга, как показано на рис. 3-1. Пусть n -й атом смещен из положения равновесия на расстояние u_n . Если отклонение атомов от положения равновесия мало, то силы взаимодействия можно рассматривать как квазиупругие, т. е. пропорциональные отклонению (закон Гука).

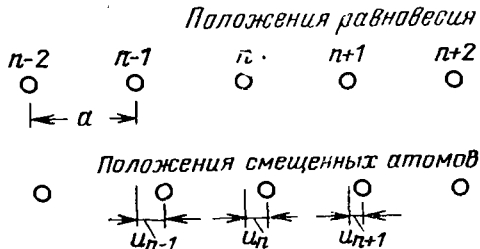


Рис. 3-1. Схематическое изображение линейной решетки из одинаковых атомов.

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

$$F_n = -\beta(u_n - u_{n+1}) - \beta(u_n - u_{n-1}) = \beta(u_{n+1} + u_{n-1} - 2u_n), \quad (3-9)$$

где β — коэффициент квазиупругой силы. Уравнение движения для этого атома имеет вид:

$$m \frac{d^2 u_n}{dt^2} = \beta(u_{n+1} + u_{n-1} - 2u_n). \quad (3-10)$$

Решение уравнения будем искать в виде

$$u_n = A e^{i(qan - \omega t)}. \quad (3-11)$$

Подставляя (3-11) в (3-10), получаем:

$$-m\omega^2 = \beta (e^{iqa} + e^{-iqa} - 2). \quad (3-12)$$

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

$$\omega = \pm 2 (\beta/m)^{1/2} \sin \frac{aq}{2} = \pm \omega_m \sin \frac{aq}{2}, \quad (3-13)$$

где

$$\omega_m = 2 (\beta/m)^{1/2}. \quad (3-14)$$

Зависимость частоты ω от волнового числа q носит название закона дисперсии для колебаний. Таким образом, решения (3-11) типа упругой волны удовлетворяют уравнениям движения (3-10) для любого n , если частота колебаний волны связана с волновым числом q дисперсионным соотношением вида (3-13).

Скорость распространения упругой волны вдоль линейной цепочки атомов

$$v = \frac{\omega \lambda}{2\pi} = \frac{\lambda}{\pi} (\beta/m)^{1/2} \sin \frac{\pi a}{\lambda} \quad (3-15)$$

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

Установим пределы изменения волнового числа q . Для этого сравним уравнения колебаний при волновых числах q и $q' = q + \frac{2\pi}{a}g$, где g — целое положительное или отрицательное число. В этом случае

$$(u_n)_g = A e^{i(qan - \omega t)} e^{i2\pi ng} = A e^{i(qan - \omega t)} = u_n, \quad (3-16)$$

так как n — порядковый номер атома, то ng — целое число и $e^{i2\pi ng} = 1$. Из соотношения (3-16) следует, что волна $(u_n)_g$ во всех точках и во все моменты времени совпадает с волной u_n . Это значит, что волновые числа q и q' физически неразличимы и, следовательно, характеризуют одно и то же колебание. Поэтому достаточно рассматривать дисперсионное соотношение при изменении волнового числа в пределах

$$-\frac{\pi}{a} \leq q < \frac{\pi}{a}. \quad (3-17)$$

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

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 атомы находились в равновесии на расстоянии постоянной решетки a . Так как число атомов N велико, то радиус окружности будет большим. Поскольку в цепочке взаимодействуют лишь атомы, находящиеся на расстоянии нескольких постоянных решетки, то условия, в которых будут находиться граничные атомы, не влияют на движение атомов внутри цепочки. Для этого случая граничные условия можно заменить условиями цикличности Борна—Кармана, согласно которым

$$u_{n+N} = u_n. \quad (3-18)$$

Подставив (3-18) в выражение (3-11), получим:

$$u_{n \pm N} = A e^{i[qa(n \pm N) - \omega t]} = A e^{(gan - \omega t)} e^{\pm iqaN} = u_n. \quad (3-19)$$

Из равенства (3-19) следует, что $e^{\pm iqaN} = 1$. Это значит, что

$$qaN = 2\pi g, \quad (3-20)$$

откуда

$$q = \frac{2\pi}{aN} g. \quad (3-21)$$

С учетом (3-17) найдем, что g должно меняться в пределах

$$-\frac{N}{2} \leq g < \frac{N}{2}. \quad (3-22)$$

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

Для линейной решетки, состоящей из одинаковых атомов, закон дисперсии представлен на рис. 3-2. На этом рисунке дана зависимость $\omega(q)$ для упругой однородной струны (континуум).

Для длинных волн, или малых значений q , можно положить $\sin \frac{aq}{2} \approx \frac{aq}{2}$. Тогда из (3-13) будем иметь:

$$\omega = 2(\beta/m)^{1/2} \frac{aq}{2} = \sqrt{\beta/m} aq. \quad (3-23)$$

Для линейной атомной цепочки линейная плотность $\rho = m/a$, а модуль Юнга

$$T_{II} = \frac{F_n}{\frac{|u_n - u_{n+1}|}{a}} = \beta a, \quad (3-24)$$

поэтому

$$v_{об} = \sqrt{T_{II}/\rho} = \sqrt{\beta/m} a. \quad (3-25)$$

С учетом (3-25) соотношение (3-23) принимает вид:

$$\omega = v_{зв} q, \quad (3-26)$$

а максимальная частота ω_m равна:

$$\omega_m = 2 (\beta/m)^{1/2} = \frac{2v_{зв}}{a}. \quad (3-27)$$

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

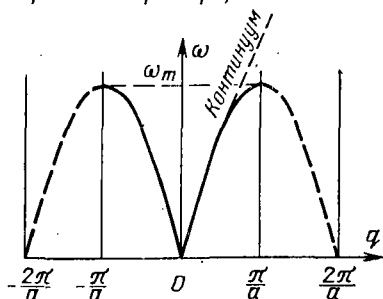


Рис. 3-2. Зависимость частоты колебания от волнового числа для линейной решетки, состоящей из одинаковых атомов.

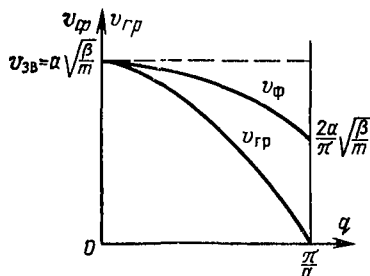


Рис. 3-3. Зависимость фазовой и групповой скорости от волнового числа.

В случае, если имеет место дисперсия волны, надо различать фазовую скорость, с которой распространяется фаза монохроматической волны, которая равна $v_\phi = \omega/q$, и групповую скорость, с которой распространяется волновой пакет, а следовательно, энергия волны, $v_{гр} = d\omega/dq$ (рис. 3-3). Для линейной цепочки одинаковых атомов согласно (3-13) и (3-27)

$$v_\phi = \omega_m \left| \frac{\sin \frac{aq}{2}}{q} \right| = v_{зв} \left| \frac{\sin \frac{aq}{2}}{\frac{1}{2} aq} \right|; \quad (3-28)$$

$$v_{гр} = \frac{a\omega_m}{2} \left| \cos \frac{aq}{2} \right| = v_{зв} \left| \cos \frac{aq}{2} \right|. \quad (3-29)$$

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

$$v_\phi = v_{гр} = v_{зв}, \quad (3-30)$$

скорость волны и групповая скорость равны и представляют собой скорость звука. Это означает, что в пределах длинных волн одноатомная цепочка ведет себя как непрерывный объект.

Как видно из рис. 3-2, с увеличением значений q , т. е. с уменьшением длины волны, частота колебаний одноатомной цепочки отклоняется от линейной зависимости. Минимальная длина волны

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

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

3-3. КИНЕТИЧЕСКАЯ И ПОТЕНЦИАЛЬНАЯ ЭНЕРГИЯ ОДНОМЕРНОЙ РЕШЕТКИ. НОРМАЛЬНЫЕ КООРДИНАТЫ

В атомной цепочке выражение (3-11) удовлетворяет уравнению движения (3-10), если выполнено соотношение дисперсии (3-13). В случае малых колебаний движение атомов решетки может быть представлено посредством линейной суперпозиции (суммы) бегущих волн смещения типа (3-11). Каждая из этих волн будет характеризоваться волновым числом q , соответствующей ему частотой ω_q и амплитудой A_q . В этом случае смещение n -го атома цепочки запишется в виде

$$u_n = \sum_q A_q e^{i(qan - \omega_q t)}. \quad (3-31)$$

Введем новые координаты, положив

$$Q_q = C_q e^{-i\omega_q t}, \quad (3-32)$$

где $C_q = \sqrt{N} A_q$. Независимые переменные Q_q , изменяющиеся во времени по гармоническому закону, называются нормальными, или главными, координатами решетки кристалла.

Подставим значения Q_q и A_q в (3-31), в результате будем иметь:

$$u_n = \frac{1}{\sqrt{N}} \sum_q Q_q e^{iqan}. \quad (3-33)$$

Определим теперь внутреннюю (полную) энергию колебаний линейной решетки. Она складывается из кинетической T и потен-

циальной U энергии всех атомов. Кинетическая энергия согласно (3-33) равна:

$$T = \sum_{n=1}^N \frac{m}{2} \dot{u}_n^2 = \frac{m}{2N} \sum_q \dot{Q}_q^2 \sum_{n=1}^N e^{i2qan}. \quad (3-34)$$

Подсчитаем сомножитель $\sum_{n=1}^N e^{i2qan}$, учитывая выражение (3-21). Он вычисляется как сумма геометрической прогрессии

$$\begin{aligned} \sum_{n=1}^N e^{i2qan} &= \sum_{n=1}^N e^{i \frac{4\pi}{N} ng} = e^{i \frac{4\pi}{N} g \cdot 1} + e^{i \frac{4\pi}{N} g \cdot 2} + \dots + e^{i \frac{4\pi}{N} g N} = \\ &= \frac{e^{i \frac{4\pi}{N} g} - e^{i \frac{4\pi}{N} g N}}{1 - e^{i \frac{4\pi}{N} g}} = \frac{e^{i \frac{4\pi}{N} g} (1 - e^{i 4\pi g})}{1 - e^{i \frac{4\pi}{N} g}} = \\ &= \begin{cases} 0, & \text{когда } q \neq 0 \text{ (или } g \neq 0); \\ N, & \text{когда } q = 0 \text{ (или } g = 0), \end{cases} \end{aligned} \quad (3-35)$$

так как $e^{i 4\pi g} = 1$, если $g \neq 0$; а сумма ряда равна N , если $g = 0$. С учетом (3-34) кинетическая энергия атомов одномерной решетки примет вид:

$$T = \sum_q \frac{m}{2} \dot{Q}_q^2. \quad (3-36)$$

Так как сила $F_n = -\frac{\partial U}{\partial u_n}$, то для потенциальной энергии решетки U с учетом (3-9) и (3-11) получим:

$$U = U_0 + \frac{\beta}{2} \sum_{n=1}^N (2u_n - u_{n+1} - u_{n-1}). \quad (3-37)$$

Для определения β используем уравнение движения (3-10), в которое подставим величины смещения (3-31):

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

Отсюда найдем:

$$\beta = \sum_q \frac{m\omega_q^2}{2 - e^{iqa} - e^{-iqa}}. \quad (3-38)$$

Теперь в уравнение (3-37) подставим значение β (3-38) и величины смещения (3-33), выраженные через нормальные координаты. В результате, используя (3-34), будем иметь:

$$U = U_0 + \sum_q \frac{m}{2} \omega_q^2 Q_q^2. \quad (3-39)$$

Полная энергия колебаний решетки

$$E = T + U = U_0 + \sum_q \left(\frac{m}{2} \dot{Q}_q^2 + \frac{m}{2} \omega_q^2 Q_q^2 \right). \quad (3-40)$$

При $U_0 = 0$

$$E = \sum_q \left(\frac{m}{2} \dot{Q}_q^2 + \frac{m}{2} \omega_q^2 Q_q^2 \right) = \sum_q E_q. \quad (3-41)$$

Выражение

$$E_q = \frac{m}{2} \dot{Q}_q^2 + \frac{m}{2} \omega_q^2 Q_q^2 \quad (3-42)$$

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

$$E_q = \hbar \omega_q \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (3-43)$$

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

$$E = \sum_q E_q = \sum_q \hbar \omega_q \left(n + \frac{1}{2} \right). \quad (3-44)$$

Из (3-41) следует, что полная энергия колебаний решетки, выраженной через нормальные координаты, имеет вид энергии совокупности гармонических осцилляторов. Это означает, что одномерный кристалл, содержащий N атомов, можно рассматривать как динамическую систему с N степенями свободы, и ее движение можно описать с помощью независимых друг от друга N нормальных координат. Каждая нормальная координата описывает определенную конфигурацию всех атомов кристалла, совершающую колебания по простому гармоническому закону (3-32). Такое колебание носит название нормального колебания. Координата Q удовлетворяет уравнению

$$\ddot{Q}_q + \omega_q^2 Q_q = 0 \quad (q = 1, 2, \dots, N), \quad (3-45)$$

которое является уравнением движения гармонического осциллятора.

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

3-4. КОЛЕБАНИЯ ДВУХАТОМНОЙ ЛИНЕЙНОЙ ЦЕПОЧКИ

Рассмотрим теперь одномерный кристалл, состоящий из атомов двух сортов с массами m_1 и m_2 (рис. 3-4). В качестве примеров веществ, у которых в элементарной ячейке имеется по два атома, можно привести ионные щелочно-галогенидные соединения, ковалент-

ные 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 .

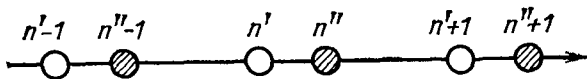


Рис. 3-4. Одномерная модель решетки, имеющей два атома в элементарной ячейке.

Уравнения движения атомов обоих сортов имеют вид:

$$\left. \begin{aligned} m_1 \ddot{u}'_n &= \beta (u''_n + u''_{n-1} - 2u'_n); \\ m_2 \ddot{u}''_n &= \beta (u'_n + u'_{n+1} - 2u''_n). \end{aligned} \right\} \quad (3-46)$$

Решение этих уравнений будем искать в виде

$$u_n^\alpha = A_\alpha e^{i(qan - \omega t)}. \quad (3-47)$$

Подставляя выражение (3-47) в (3-46), получаем:

$$\left. \begin{aligned} (2\beta - m_1 \omega^2) A_1 &= \beta (1 + e^{-iqa}) A_2; \\ (2\beta - m_2 \omega^2) A_2 &= \beta (1 + e^{iqa}) A_1. \end{aligned} \right\} \quad (3-48)$$

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

$$\left\| \begin{array}{cc} 2\beta - m_1 \omega^2 & \beta (1 + e^{-iqa}) \\ \beta (1 + e^{iqa}) & 2\beta - m_2 \omega^2 \end{array} \right\| = 0. \quad (3-49)$$

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

$$\omega^4 - 2\beta \frac{m_1 + m_2}{m_1 m_2} \omega^2 + \frac{4\beta^2}{m_1 m_2} \sin^2 \frac{aq}{2} = 0. \quad (3-50)$$

Из этого уравнения найдем:

$$\omega_1^2 = \frac{\omega_0^2}{2} \left[1 + \sqrt{1 - \gamma^2 \sin^2 \frac{aq}{2}} \right]; \quad (3-51)$$

$$\omega_2^2 = \frac{\omega_0^2}{2} \left[1 - \sqrt{1 - \gamma^2 \sin^2 \frac{aq}{2}} \right], \quad (3-52)$$

где

$$\omega_0^2 = 2\beta \frac{m_1 + m_2}{m_1 m_2}; \quad \gamma^2 = 4 \frac{m_1 m_2}{(m_1 + m_2)^2}. \quad (3-53)$$

Из формул (3-51) и (3-52) следует, что каждому значению q (или λ) соответствуют два значения ω (так как отрицательные значения ω не имеют смысла). Следовательно, *у одномерного кристалла, в элементарной ячейке которого находятся два атома, имеются две ветви колебаний; одна из них, характеризующаяся ω_1 , называется оптической, другая с ω_2 — акустической ветвью колебаний.*

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

$$\left. \begin{aligned} \omega_{\text{оп}}(0) &= \omega_0; & \omega_{\text{оп}}\left(\frac{\pi}{a}\right) &= \frac{\omega_0}{\sqrt{2}} \sqrt{1 + \sqrt{1 - \gamma^2}}; \\ \omega_{\text{ак}}(0) &= 0; & \omega_{\text{ак}}\left(\frac{\pi}{a}\right) &= \frac{\omega_0}{\sqrt{2}} \sqrt{1 - \sqrt{1 - \gamma^2}}. \end{aligned} \right\} \quad (3-54)$$

Отсюда видно, что

$$\omega_{\text{оп}}(0) = \omega_0 > \omega_{\text{оп}}\left(\frac{\pi}{a}\right) > \omega_{\text{ак}}\left(\frac{\pi}{a}\right) > \omega_{\text{ак}}(0) = 0. \quad (3-55)$$

Для бесконечно длинной волны акустической ветви колебаний, когда $\omega_{\text{ак}}(0) = 0$, согласно (3-47) и (3-48)

$$\left(\frac{u'_n}{u''_n}\right)_{\text{ак}} = \left(\frac{A_1}{A_2}\right)_{\text{ак}} = 1, \quad (3-56)$$

т. е. *смещение атомов одинаково. Это означает, что в бесконечно длинной волне акустической ветви колебаний атомы движутся синхронно и в любой момент времени их отклонения от положения равновесия одинаковы, т. е. ячейки смещаются как одно целое* (рис. 3-5).

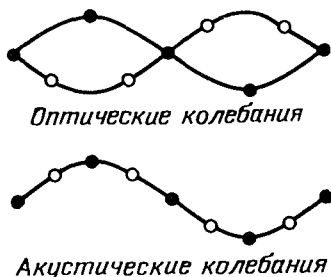


Рис. 3-5. Схематическое изображение колебаний частиц в двухатомной линейной решетке, иллюстрирующее природу различия между оптическими и акустическими колебаниями. Предполагается, что колебания поперечные и что длина волны обоих типов колебаний одинакова.

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

Бесконечно длинной волне оптической ветви колебаний соответствует максимальная частота, равная ω_0 . Для нее согласно (3-47)

и (3-48) справедливо отношение

$$\left(\frac{u'_n}{u''_n} \right)_{\text{оп}} = \left(\frac{A_1}{A_2} \right)_{\text{оп}} = - \frac{m_2}{m_1}. \quad (3-57)$$

Следовательно, в бесконечно длинной волне оптической ветви колебаний атомы ячейки смещаются в противоположных направлениях, они колеблются в противофазе, так что центр тяжести каждой пары атомов остается неподвижным, т. е. $m_1 u'_n + m_2 u''_n = 0$ (рис. 3-5). Если бы элементарную ячейку составляли ионы разного знака, то их колебания в противофазе вызвали бы смещение ионов и привели бы к изменению электрического дипольного момента ячейки. Это проявилось бы в поглощении и испускании инфракрасного излучения. Поэтому колебания этого типа и были названы оптическими.

Для длинных волн, а следовательно, для малых значений q , полагая $\sin \frac{aq}{2} \approx \frac{aq}{2}$, корни в (3-51) и (3-52) можно разложить в ряд. В результате для длинных волн оптической ветви колебаний

$$\omega_{\text{оп}} = \omega_0 \left(1 - \frac{\gamma^2 a^2}{32} q^2 \right) \quad (3-58)$$

частота колебаний с ростом q уменьшается по квадратичному закону. При $q \rightarrow 0$ фазовая скорость оптических колебаний $v_{\text{ф}} = \frac{\omega_{\text{оп}}}{q}$ стремится к бесконечности, а групповая скорость $v_{\text{гр}} = \frac{d\omega_{\text{оп}}}{dq}$ равна нулю.

Для длинных волн акустической ветви колебаний

$$\omega_{\text{ак}} \approx \frac{1}{4} \omega_0 \gamma a q = \sqrt{\frac{2\beta}{m_1 + m_2}} q = v_{\text{зв}} q \quad (3-59)$$

частота колебаний пропорциональна скорости звука, а фазовая и групповая скорость

$$v_{\text{ф}} = v_{\text{гр}} = v_{\text{зв}} \quad (3-60)$$

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

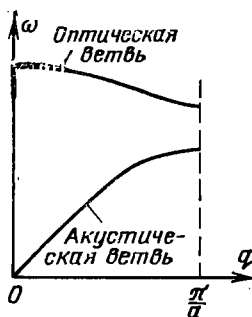
Для наименьшей длины волны $\lambda = 2a$, когда $q = \pi/a$, частота оптических колебаний будет минимальной:

$$\omega_{\text{оп}} \left(\frac{\pi}{a} \right) = \left(\frac{2\beta}{m_2} \right)^{1/2} = \omega_{0 \text{ мин}}, \quad (3-61)$$

а частота акустических колебаний будет максимальной:

$$\omega_{\text{ак}} \left(\frac{\pi}{a} \right) = \left(\frac{2\beta}{m_1} \right)^{1/2} = \omega_{\text{ак. макс.}} \quad (3-62)$$

Рис. 3-6. Оптические и акустические ветви колебаний одномерной цепочки, в элементарной ячейке которой находятся два атома (положительные ветви кривых).



Зависимость ω от q для оптической и акустической ветвей колебаний одномерной решетки, имеющей в элементарной ячейке два атома, изображена на рис. 3-6. Во всем диапазоне изменений q $\omega_{\text{оп}} > \omega_{\text{ак}}$. Частота оптических колебаний незначительно уменьшается при увеличении q . Наибольшее отличие величины $\omega_{\text{оп}}$ от $\omega_{\text{ак}}$ имеет место при $q = 0$.

3-5. КОЛЕБАНИЯ АТОМОВ ТРЕХМЕРНОЙ РЕШЕТКИ

Установленная закономерность колебаний одномерной цепочки атомов позволяет высказать общие соображения о тепловом движении атомов трехмерного кристалла. Смещение атомов трехмерной решетки также можно описать посредством суперпозиции нормальных колебаний решетки. Каждое нормальное колебание представляет собой волну, направление распространения которой определяется волновым вектором \mathbf{q} . Волна характеризуется длиной волны λ и частотой ω_q . Волновой вектор \mathbf{q} имеет столько разрешенных значений, сколько в кристалле имеется элементарных ячеек. При этом значения q меняются в пределах

$$-\pi \leq \mathbf{q}_i \mathbf{a}_i < \pi, \quad (3-63)$$

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

Объемный кристалл содержит $N = N_x N_y N_z$ элементарных ячеек, в каждой из которых может быть s атомов. Поскольку каждый атом обладает тремя степенями свободы, то кристалл характеризуется $3sN$ степенями свободы и в таком кристалле 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 $2s$ типа 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 при $q \rightarrow 0$ $v_{\phi} \rightarrow \infty$, а $v_{гр} \rightarrow 0$.

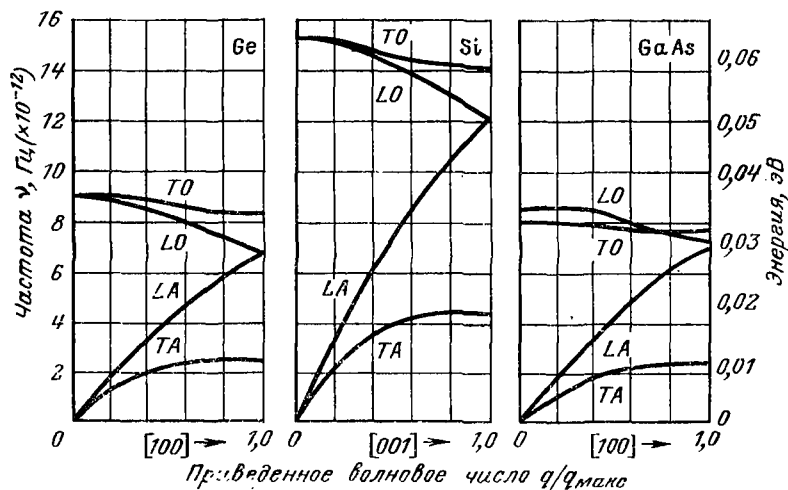


Рис. 3-7. Колебательные спектры решетки германия, кремния и арсенида галлия.

ТА и LA — поперечные и продольные акустические колебания; TO и LO — поперечные и продольные оптические колебания.

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

На рис. 3-7 приведены измеренные колебательные спектры решеток германия, кремния и арсенида галлия. У них в элементарной ячейке находятся два атома, поэтому эти полупроводники имеют оптические и акустические ветви колебаний, каждая из которых содержит одну продольную ветвь и две поперечные ветви колебаний, являющиеся вырожденными.

Как следует из рис. 3-7, 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 (вырожденные).

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

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

Таким образом, полная энергия колебаний кристалла

$$E = \sum_{qj} E_{qj} = \sum_{qj} \hbar\omega_j(\mathbf{q}) \left[n_{qj} + \frac{1}{2} \right] \quad (n_{qj} = 0, 1, 2, \dots). \quad (3-64)$$

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

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

$$\Delta n_{qj} = \pm 1. \quad (3-65)$$

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

По аналогии с фотоном — квантом электромагнитного поля — фонон можно рассматривать как носитель кванта энергии колебаний решетки, т. е. рассматривать как свободную квазичастицу.

Таким образом, колебания решетки в энергетическом отношении эквивалентны фоновому газу. В этом случае $\Delta n_{qj} = +1$ будет означать рождение фонона, а $\Delta n_{qj} = -1$ — уничтожение фонона.

3-6. СТАТИСТИКА ФОНОНОВ

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

Следовательно,

$$\langle n_{qj} \rangle = \frac{\langle E_{qj} \rangle}{\hbar\omega_j(\mathbf{q})}. \quad (3-66)$$

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

$$W_s = C e^{-E_s/kT}. \quad (3-67)$$

Коэффициент C определяется из условия нормировки вероятности

$$\sum_s W_s = 1, \quad (3-68)$$

из которого следует, что

$$C = \left(\sum_s e^{-E_s/kT} \right)^{-1}. \quad (3-69)$$

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

$$\langle E \rangle = \sum_{s=0}^{\infty} E_s W_s. \quad (3-70)$$

Учитывая равенства (3-68) и (3-69), выражение для равновесного значения энергии гармонического осциллятора запишется в виде

$$\langle E \rangle = \frac{\sum_{n=0}^{\infty} E_n e^{-E_n/kT}}{\sum_{n=0}^{\infty} e^{-E_n/kT}} = - \frac{d}{d \left(\frac{1}{kT} \right)} \ln \sum_{n=0}^{\infty} e^{-E_n/kT}. \quad (3-71)$$

Найдем значение суммы в (3-71), учитывая, что энергия гармонического осциллятора $E_s = \hbar\omega \left(n + \frac{1}{2} \right)$. Определить ее можно как сумму геометрической прогрессии:

$$\sum_{n=0}^{\infty} e^{-E_s/kT} = \sum_{n=0}^{\infty} e^{-[\hbar\omega (n + 1/2)]/kT} = e^{-\hbar\omega/2kT} (1 + e^{-\hbar\omega/kT} + e^{-2\hbar\omega/kT} + \dots) = \frac{e^{-\hbar\omega/2kT}}{1 - e^{-\hbar\omega/kT}}. \quad (3-72)$$

Если в (3-71) подставим выражение (3-72) и проведем дифференцирование по аргументу $1/kT$, то получим:

$$\begin{aligned} \langle E \rangle &= - \frac{d}{d\left(\frac{1}{kT}\right)} \ln \frac{e^{-\hbar\omega/2kT}}{1 - e^{-\hbar\omega/kT}} = \\ &= \frac{\hbar\omega}{2} + \frac{d}{d\left(\frac{1}{kT}\right)} \ln (1 - e^{-\hbar\omega/kT}) = \frac{\hbar\omega}{2} + \frac{\hbar\omega e^{-\hbar\omega/kT}}{1 - e^{-\hbar\omega/kT}} = \\ &= \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}. \end{aligned} \quad (3-73)$$

Член $\frac{1}{2} \hbar\omega$, не зависящий от температуры, называется нулевой энергией осциллятора. Поскольку во взаимодействии с электронами принимают участие только те нормальные колебания, которые возбуждены выше нулевого энергетического уровня, то нулевую энергию можно не учитывать. Поэтому будем считать, что средняя энергия гармонического осциллятора равна:

$$\langle E \rangle = \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}. \quad (3-74)$$

Тогда

$$\langle n \rangle = \frac{\langle E \rangle}{\hbar\omega} = \frac{1}{e^{\hbar\omega/kT} - 1}. \quad (3-75)$$

Из этого выражения следует, что фононы подчиняются статистике Бозе—Эйнштейна. $\langle n \rangle$ есть «среднее» значение колебательного квантового числа осциллятора. Другими словами, $\langle n \rangle$ есть среднее число фононов в данном нормальном колебании, т. е. среднее число фононов (их концентрация) с энергией $\hbar\omega$ при тепловом равновесии в одном квантовом состоянии — в ячейке фазового пространства объемом h^3 с энергией $\hbar\omega$.

При высоких температурах, когда $kT \gg \hbar\omega$, осциллятор находится в высоком энергетическом состоянии и число фононов в одной ячейке фазового пространства составляет:

$$\langle n \rangle \approx kT/\hbar\omega, \quad (3-76)$$

а средняя энергия осциллятора

$$\langle E \rangle = \langle n \rangle \hbar \omega = kT, \quad (3-77)$$

т. е. соответствует классической статистике.

При низких температурах, когда $kT \ll \hbar \omega$,

$$\langle n \rangle = e^{-\hbar \omega / kT}, \quad (3-78)$$

а средняя энергия осциллятора

$$\langle E \rangle = \hbar \omega e^{-\hbar \omega / kT}. \quad (3-79)$$

Итак, число фононов, находящихся в одном квантовом состоянии, зависит от температуры. Кроме того, при взаимодействии с электронами или друг с другом фононы возникают и исчезают. Следовательно, фононы ведут себя не как газ обыкновенных частиц, именно поэтому они и называются квазичастицами.

В твердом теле возможны акустические и оптические колебания. Но каждое колебание решетки в энергетическом отношении эквивалентно фонону. Поэтому можно говорить об акустических и оптических фононах. Их среднее число, приходящееся на одно квантовое состояние, согласно (3-75) зависит от энергии фононов $\hbar \omega_j(\mathbf{q})$ и температуры T .

Во всем диапазоне изменения q от 0 до π/a $\omega_{j \text{ оп}} \gg \omega_{j \text{ ак}}$, поэтому энергия оптических фононов больше энергии акустических фононов (см. рис. 3-7). При малых значениях q энергия акустических фононов, равная

$$E_{\text{ак.ф}} = \hbar \omega_{j \text{ ак}} = \hbar v_{j \text{ зв}} q,$$

значительно меньше энергии оптических фононов, для которых ее значение максимально: $E_{\text{оп.ф}} = \hbar \omega_{0j}$.

При низких температурах, когда $kT \ll \hbar \omega_{0j}$, оптические фононы отсутствуют, возбуждаются только акустические фононы, а число их в одном квантовом состоянии определяется равенством (3-78). При высоких температурах, когда $kT \gg \hbar \omega_{0j}$, кроме акустических фононов возбуждается большое количество и оптических фононов.

3-7. ТЕПЛОЕМОСТЬ КРИСТАЛЛИЧЕСКОЙ РЕШЕТКИ

По определению теплоемкость твердого тела

$$C = \frac{\partial E}{\partial T}. \quad (3-80)$$

Поскольку гармонические осцилляторы не взаимодействуют, то полная энергия кристалла в термодинамическом равновесии E равна сумме средних значений энергии всех осцилляторов. При этом суммирование нужно провести с учетом волнового вектора \mathbf{q} ,

имеющего N различных значений, и типов колебаний, которых может быть $3s$. Кроме того, нужно учесть нулевую энергию $E_0 = \sum_{\mathbf{q}l} \frac{1}{2} \hbar \omega_l(\mathbf{q})$. Поэтому на основании (3-74) энергия колебаний решетки может быть записана в виде

$$E = E_0 + \sum_{j=1}^{3s} \sum_{\mathbf{q}} \frac{\hbar \omega_j(\mathbf{q})}{e^{\hbar \omega_j(\mathbf{q})/kT} - 1} = E_0 + \langle E_A \rangle + \langle E_0 \rangle. \quad (3-81)$$

Здесь равновесное значение энергии акустических колебаний решетки

$$\langle E_A \rangle = \sum_{j=1}^3 \sum_{\mathbf{q}} \frac{\hbar \omega_j(\mathbf{q})}{e^{\hbar \omega_j(\mathbf{q})/kT} - 1}, \quad (3-82)$$

а равновесное значение энергии оптических колебаний решетки

$$\langle E_0 \rangle = \sum_{j=4}^{3s} \sum_{\mathbf{q}} \frac{\hbar \omega_j(\mathbf{q})}{e^{\hbar \omega_j(\mathbf{q})/kT} - 1}. \quad (3-83)$$

Определим $\langle E_A \rangle$. Так как волновой вектор \mathbf{q} меняется квазинепрерывно от 0 до π/a , имея N значений, равных числу элементарных ячеек в кристалле, то и частота колебаний $\omega_j(\mathbf{q})$ также меняется квазинепрерывно. В акустической ветви колебаний она меняется от 0 до $\omega_{A \text{ макс}}$. Поскольку N велико, то в (3-82) сумму по $\omega_j(\mathbf{q})$, выражающую энергию акустических колебаний $\langle E_A \rangle$, можно заменить интегралом:

$$\langle E_A \rangle = \int_0^{\omega_{A \text{ макс}}} dZ \hbar \omega \langle n \rangle = \int_0^{\omega_{A \text{ макс}}} dZ \frac{\hbar \omega}{e^{\hbar \omega/kT} - 1}, \quad (3-84)$$

где dZ — число нормальных колебаний в интервале от q до $q + dq$, а $\omega_{A \text{ макс}}$ соответствует $q_{\text{макс}}$.

Для определения числа нормальных колебаний dZ во всех трех акустических ветвях поступим следующим образом. В \mathbf{q} -пространстве выделим слой, заключенный между сферическими поверхностями радиуса q и $q + dq$. Объем этого слоя (заштрихованная область на рис. 3-8) $dV_{\mathbf{q}} = 4\pi q^2 dq$. Объем элементарной ячейки зоны Бриллюэна кубического кристалла ΔV_B в \mathbf{q} -пространстве согласно (3-21) равен:

$$\Delta V_B = dq_x dq_y dq_z = \frac{2\pi}{aN_x} \frac{2\pi}{aN_y} \frac{2\pi}{aN_z} = \frac{8\pi^3}{V}, \quad (3-85)$$

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

Каждой элементарной ячейке зоны Бриллюэна соответствует одно возможное колебательное состояние кристалла (т. е. одно нор-

мальное колебание). Следовательно, в одной акустической ветви на интервал dq приходится нормальных колебаний в количестве

$$dZ = \frac{dV_q}{\Delta V_B} = V \frac{q^2 dq}{2\pi^2}. \quad (3-86)$$

В соответствии с теорией теплоемкости по Дебаю будем считать, что для трех акустических ветвей для всей области изменения q (от 0 до π/a) имеет место закон дисперсии

$$\omega_j(q) = v_{j\text{зв}} q \quad (j = 1, 2, 3). \quad (3-87)$$

Надо отметить, что в этом случае кристаллическое тело рассматривается как изотропный континуум, в котором отсутствует дисперсия. Теперь с учетом (3-87) число колебаний в одной акустической ветви в интервале dq , а следовательно, в интервале частот от ω до $\omega + d\omega$ равно:

$$dZ = V \frac{q^2 dq}{2\pi^2} = V \frac{\omega^2 d\omega}{2\pi^2 v_{j\text{зв}}^3}. \quad (3-88)$$

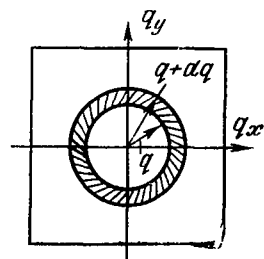


Рис. 3-8. Объем слоя, заключенного между двумя поверхностями.

Найдем теперь плотность состояний колебаний решетки $g(\omega)$, т. е. число нормальных колебаний, соответствующих единичному интервалу частоты кристалла единичного объема. Величина $g(\omega)$ также называется функцией распределения частот. Согласно (3-88) она имеет вид:

$$g(\omega) = \frac{dZ}{V d\omega} = \frac{\omega^2}{2\pi^2 v_{j\text{зв}}^3}. \quad (3-89)$$

Для трех возможных типов акустических колебаний — одного продольного со скоростью $v_{l\text{зв}}$ и двух поперечных со скоростью $v_{t\text{зв}}$ — плотность состояний колебаний в интервале $\omega, \omega + d\omega$ равна:

$$g(\omega) = \frac{\omega^2}{2\pi^2 v_{l\text{зв}}^3} + \frac{2\omega^2}{2\pi^2 v_{t\text{зв}}^3} = \frac{3\omega^2}{2\pi^2 v_0^3}, \quad (3-90)$$

где v_0 , определяемая условием

$$\frac{1}{v_0^3} = \frac{1}{3} \left[\frac{1}{v_{l\text{зв}}^3} + \frac{2}{v_{t\text{зв}}^3} \right], \quad (3-91)$$

есть скорость звука, усредненная по направлениям и типам колебаний. В этом случае для кристалла объемом V энергия акустиче-

ских колебаний (3-84) запишется в виде

$$\langle E_A \rangle = \int_0^{\omega_{A \text{ макс}}} Vg(\omega) \frac{\hbar \omega d\omega}{e^{\hbar \omega / kT} - 1} = \frac{3V\hbar}{2\pi^2 v_{0 \text{ зв}}^3} \int_0^{\omega_{A \text{ макс}}} \frac{\omega^3 d\omega}{e^{\hbar \omega / kT} - 1}. \quad (3-92)$$

Введем новые переменные

$$x = \frac{\hbar \omega}{kT} \quad \text{и} \quad \frac{\hbar \omega_{A \text{ макс}}}{k} = \theta,$$

тогда

$$x_{\text{макс}} = \frac{\hbar \omega_{A \text{ макс}}}{kT} = \frac{\theta}{T}. \quad (3-93)$$

В этом случае

$$\begin{aligned} \langle E_A \rangle &= \frac{3V\hbar}{2\pi^2 v_{0 \text{ зв}}^3} \left(\frac{kT}{\hbar} \right)^4 \int_0^{x_{\text{макс}}} \frac{x^3 dx}{e^x - 1} = \\ &= \frac{3V\hbar}{2\pi^2 v_{0 \text{ зв}}^3} \left(\frac{kT}{\hbar} \right)^4 \int_0^{\theta/T} \frac{x^3 dx}{e^x - 1}. \end{aligned} \quad (3-94)$$

Максимальная частота $\omega_{A \text{ макс}}$ может быть определена из условия, что полное число акустических колебаний равно $3N$. Тогда

$$\int_0^{\omega_{A \text{ макс}}} Vg(\omega) d\omega = \frac{3V}{2\pi^2 v_{0 \text{ зв}}^3} \int_0^{\omega_{A \text{ макс}}} \omega^2 d\omega = \frac{V\omega_{A \text{ макс}}^3}{2\pi^2 v_{0 \text{ зв}}^3} = 3N, \quad (3-95)$$

откуда

$$\omega_{A \text{ макс}} = v_{0 \text{ зв}} \left(\frac{6\pi^2 N}{V} \right)^{1/3}. \quad (3-96)$$

Параметр θ , равный:

$$\theta = \frac{\hbar \omega_{A \text{ макс}}}{k} = \frac{\hbar v_{0 \text{ зв}}}{k} \left(\frac{6\pi^2 N}{V} \right)^{1/3}, \quad (3-97)$$

называется *характеристической температурой твердого тела* или температурой Дебая. Ее физический смысл состоит в том, что $\hbar\theta$ — это энергия самых коротких акустических фононов в теле. Длина волны таких фононов $\lambda = 2a = 6 \cdot 10^{-8}$ см, где $a = 3 \cdot 10^{-8}$ см — постоянная решетки. Так как $\omega_{A \text{ макс}} \approx v_{0 \text{ зв}}/a \approx 10^{13}$ с⁻¹, $\hbar = 1,05 \cdot 10^{-34}$ Дж·с, $k = 1,38 \cdot 10^{-23}$ Дж/К, то $\theta \approx 100$ К. Для большинства твердых тел θ лежит в пределах 100—300 К (табл. 3-1).

Характеристическая температура Дебая θ некоторых материалов

Вещество	θ , К	Вещество	θ , К
Au	170	Ge	290
Cd	160	NaCl	280
Cu	315	KCl	230
Sn (серое)	260	KBr	177

Учитывая (3-97), коэффициент перед интегралом в выражении (3-94) можно записать в виде

$$\frac{3Vh}{2\pi^2v_0^3} \left(\frac{kT}{h}\right)^4 = 9kTN \left(\frac{T}{\theta}\right)^3, \quad (3-98)$$

поэтому

$$\langle E_\lambda \rangle = \frac{9NkT}{(\theta/T)^3} \int_0^{\theta/T} \frac{x^3 dx}{e^x - 1} = 3NkTD \left(\frac{\theta}{T}\right), \quad (3-99)$$

где функция Дебая

$$D\left(\frac{\theta}{T}\right) = \frac{3}{(\theta/T)^3} \int_0^{\theta/T} \frac{x^3 dx}{e^x - 1}. \quad (3-100)$$

Определим теперь $\langle E_0 \rangle$. Частоты оптических колебаний мало меняются во всем диапазоне изменений волнового вектора \mathbf{q} . Можно считать их равными предельным значениям $\omega_j(\mathbf{q}) = \omega_{0j}$. Введем температуру Дебая, соответствующую предельной частоте колебаний данной оптической ветви:

$$\theta_{0j} = \frac{\hbar \omega_{0j}}{k}. \quad (3-101)$$

Для большинства веществ θ_0 лежит в пределах 100—800 К. В табл. 3-2 приведены значения энергии продольных оптических фононов $\hbar\omega_0$ и соответствующие ей температуры Дебая.

Используя соотношение (3-101), энергию оптических колебаний $\langle E_0 \rangle$ можно представить в виде

$$\langle E_0 \rangle = \sum_{j=1}^{3s} \sum_{\mathbf{q}} \frac{\hbar \omega_j(\mathbf{q})}{e^{\hbar \omega_j(\mathbf{q})/kT} - 1} = NkT \sum_{j=1}^{3s} \frac{\theta_{0j}/T}{e^{\theta_{0j}/T} - 1}, \quad (3-102)$$

так как суммирование по \mathbf{q} для каждой оптической ветви колебаний эквивалентно умножению на N .

Ширина запрещенной зоны, максимальная частота продольных оптических фононов ω_0 , их энергия $\hbar\omega_0$ и температура Дебая θ_0 для некоторых полупроводников

Полупроводники	Ширина запрещенной зоны (300 К), эВ	$\omega_0 \cdot 10^{-13}$, Гц	$\hbar\omega_0$, эВ	θ_0 , К
Si	1,08	9,45	0,063	730
Ge	0,66	4,8	0,037	430
GaAs	1,35	5,35	0,0355	408
GaSb	0,67	4,4	0,029	336
InAs	0,35	4,4	0,029	334
InSb	0,17	3,5	0,0225	264

Следовательно, энергия колебаний решетки равна:

$$E = E_0 + \langle E_A \rangle + \langle E_O \rangle = E_0 + NkT \left\{ 3D\left(\frac{\theta}{T}\right) + \sum_{j=1}^{3s} \frac{\theta_{Oj}/T}{e^{\theta_{Oj}/T} - 1} \right\}. \quad (3-103)$$

В общем случае выражение для теплоемкости $C = \frac{\partial E}{\partial T}$ получается достаточно сложным. Рассмотрим область высоких и низких температур. Для области высоких температур, когда $T \gg \theta_{Oj}$ и тем более $T \gg \theta$ (так как $\theta_{Oj} > \theta$),

$$e^x \approx 1 + x \text{ и } 3D\left(\frac{\theta}{T}\right) = 9 \left(\frac{T}{\theta}\right)^3 \int_0^{\theta/T} x^2 dx = 3,$$

а сумма

$$\sum_{j=1}^{3s} \frac{\theta_{Oj}/T}{e^{\theta_{Oj}/T} - 1} = \sum_{j=1}^{3s} 1 = 3s - 3, \quad (3-104)$$

поэтому

$$E = E_0 + NkT [3 + (3s - 3)] = E_0 + 3sNkT = E_0 + 3RT, \quad (3-105)$$

где

$$R = skN = 1,98 \text{ кал/(град} \cdot \text{моль)}. \quad (3-106)$$

Теплоемкость решетки

$$C = \frac{\partial E}{\partial T} = 3skN = 3R. \quad (3-107)$$

Следовательно, при высоких температурах теплоемкость твердого тела не зависит от температуры (рис. 3-9). Для монокристаллического твердого тела $C = 5,96$ кал/(моль·град), что согласуется с законом Дюлонга и Пти.

При очень низких температурах, когда $T < \theta_{0j}$, оптические колебания не возбуждаются, а поэтому не дают вклада в энергию колебаний решетки. Для $T \ll \theta$ верхний предел интеграла в (3-100) можно положить равным ∞ . Так как

$$\int_0^{\infty} \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15}, \quad (3-108)$$

то

$$E = E_0 + \frac{3\pi^4 NkT^4}{5\theta^3}, \quad (3-109)$$

а теплоемкость решетки

$$C = \frac{12\pi^4 Nk}{5} \left(\frac{T}{\theta}\right)^3. \quad (3-110)$$

Таким образом, при низких температурах теплоемкость твердого тела пропорциональна T^3 . Закон $C \sim T^3$, полученный из теории Дебая, оправдывается только в области нескольких градусов вблизи абсолютного нуля (рис. 3-9). Отступление истинной теплоемкости от дебаевского значения (3-110) определяется тем, что

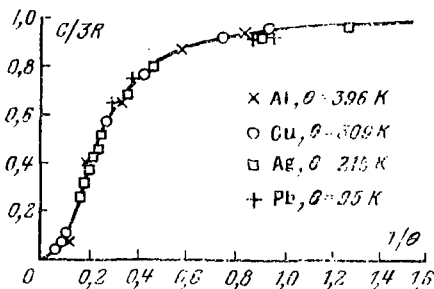


Рис. 3-9. Теплоемкость некоторых металлов.

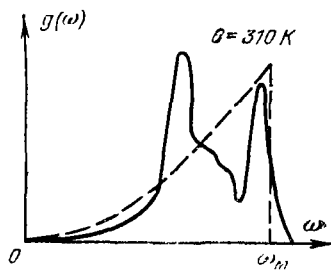


Рис. 3-10. Плотность колебаний $g(\omega)$ для решетки вольфрама (сплошная кривая) и дебаевская кривая (пунктирная кривая).

истинная плотность колебаний $g(\omega)$ для акустических ветвей в кристалле сильно отличается от зависимости (3-89). Это положение иллюстрирует рис. 3-10, где представлены рассчитанная функция $g(\omega)$ для решетки вольфрама (сплошная кривая) и дебаевская функция $g(\omega)$ (пунктирная кривая) для температуры Дебая вольфрама, равной 310 К.

3-8. ТЕПЛОЕ РАСШИРЕНИЕ И ТЕПЛОЕ СОПРОТИВЛЕНИЕ ТВЕРДОГО ТЕЛА

При определении потенциальной энергии в уравнении (3-37) был учтен только квадратичный член смещения, так называемый гармонический член. Пусть это смещение равно x , тогда потенциальная энергия $U = \beta x^2$.

Согласно Больцману вероятность отклонения атома от положения равновесия на величину x равна:

$$f(x) = Ae^{-U/kT}. \quad (3-111)$$

В этом случае *расстояние между атомами, совершающими гармоническое колебание, не изменяется, так как их среднее смещение, которое может быть определено по формуле*

$$\bar{x} = \frac{\int_{-\infty}^{\infty} xf(x) dx}{\int_{-\infty}^{\infty} f(x) dx} = \frac{\int_{-\infty}^{\infty} xe^{-\beta x^2/kT} dx}{\int_{-\infty}^{\infty} e^{-\beta x^2/kT} dx} = 0, \quad (3-112)$$

равно нулю (интеграл числителя дроби равен нулю ввиду нечетности подынтегральной функции).

Реальные тела испытывают тепловое расширение. Его можно объяснить, если учесть кубические члены разложения потенциальной энергии по смещениям атомов. Расчет ангармонических членов сил взаимодействия атомов твердого тела представляет очень сложную задачу. Однако рассмотрение простой модели двух взаимодействующих атомов позволяет понять причины теплового расширения. Кроме того, эта модель дает возможность определить правильный порядок величины температурного коэффициента линейного расширения.

Предположим, что два атома при смещении на расстояние x от положения их равновесия в решетке взаимодействуют друг с другом с силой

$$F = -\frac{dU}{dx} = -2\beta x + 3\gamma x^2 \quad (3-113)$$

и потенциальной энергией

$$U = \beta x^2 - \gamma x^3. \quad (3-114)$$

Если в уравнение (3-111) подставить значение U из (3-114), то в случае малых смещений экспоненту, соответствующую ангармоническому члену, можно разложить в ряд, т. е. $e^{\gamma x^3/kT} = 1 + \frac{\gamma x^3}{kT}$. В результате получим приближенно:

$$f(x) = Ae^{-\beta x^2/kT} \left(1 + \frac{\gamma x^3}{kT} \right). \quad (3-115)$$

Среднее отклонение атома от положения равновесия будет равно:

$$\bar{x} = \frac{\int_{-\infty}^{\infty} e^{-\beta x^2 kT} \left(x + \frac{\gamma x^4}{kT} \right) dx}{\int_{-\infty}^{\infty} e^{-\beta x^2 kT} \left(1 + \frac{\gamma x^3}{kT} \right) dx} = \frac{\frac{\gamma}{kT} \frac{3\sqrt{\pi}}{4} \left(\frac{kT}{\beta} \right)^{5.2}}{\left(\frac{\pi kT}{\beta} \right)^{1/2}} = \frac{3\gamma kT}{4\beta^2}, \quad (3-116)$$

где интегралы от первого слагаемого в числителе дроби, содержащего множитель x , и от второго слагаемого в знаменателе дроби, содержащего x^3 , равны нулю ввиду нечетности подынтегральной функции.

Как известно, линейное тепловое расширение определяется формулой

$$l_t = l_0 (1 + \alpha t), \quad (3-117)$$

поэтому в нашем случае температурный коэффициент линейного расширения α , определяемый как удлинение на единицу длины и на 1°C , равен:

$$\alpha = \frac{\bar{x}}{aT} = \frac{3\gamma}{4a\beta^2}, \quad (3-118)$$

где a — параметр решетки. Таким образом, температурный коэффициент линейного расширения твердого тела есть постоянная величина; он пропорционален коэффициенту ангармоничности и при $\gamma = 0$ равен нулю. Значения коэффициентов линейного расширения для некоторых веществ приведены в табл. 3-3.

Таблица 3-3

Температурные коэффициенты линейного расширения при комнатной температуре

Вещество	$\alpha \cdot 10^6, \text{K}^{-1}$	Вещество	$\alpha \cdot 10^6, \text{K}^{-1}$
Au	14	Pt	9
Li	56	CsCl	50
Ni	13	AgBr	400

Ангармонизм колебаний решетки позволяет также объяснить тепловое сопротивление твердого тела. При гармонических колебаниях решетки нормальные координаты являются независимыми, т. е. имеет место принцип линейных суперпозиций, согласно которому упругие волны распространяются в кристалле независимо, не рассеиваясь друг на друге. Поскольку каждой нормальной координате, представляющей собой гармонический осциллятор, со-

ответствует определенный набор фононов, то в гармоническом приближении фононы не взаимодействуют, т. е. не сталкиваются друг с другом. В этом случае тепловой поток — энергия, переносимая фононами, — распространяется со скоростью звука ($v_{гр} = v_{зв}$), поэтому тепловое сопротивление равно нулю. Учет ангармонических членов в потенциальной энергии делает нормальные координаты решетки, а следовательно, и осцилляторы ангармоническими и взаимно зависимыми. Ангармонизм осцилляторов и обуславливает рассеяние фононов на фононах. При этом в процессе столкновения участвуют три фонона, в результате чего либо два фонона превращаются в один, либо один фонон распадается на два. Рассеяние фононов на фононах, сопровождающееся рождением и исчезновением фононов, и приводит к появлению теплового сопротивления в твердом теле. С ростом температуры ангармонизм колебаний усиливается, так как возрастает амплитуда колебаний, поэтому увеличивается и тепловое сопротивление.

Глава четвертая

СТАТИСТИКА ЭЛЕКТРОНОВ И ДЫРОК В ПОЛУПРОВОДНИКАХ

4-1. ПЛОТНОСТЬ КВАНТОВЫХ СОСТОЯНИЙ

Важнейшая задача статистической физики состоит в определении числа частиц, энергия которых лежит в определенном интервале. Для ее решения необходимо знать число квантовых состояний и вероятность нахождения частиц в этих состояниях. Следовательно, для определения концентрации носителей заряда в полупроводнике необходимо знать фактическое число состояний, занятых электронами и дырками.

Пусть в кристалле единичного объема в интервале энергий от E до $E + dE$ имеется dZ квантовых состояний (с учетом спина). Обозначим через $N(E)$ плотность состояний, т. е. число состояний в единичном интервале энергии для единичного объема кристалла. Тогда согласно определению

$$N(E) = \frac{dZ}{dE}. \quad (4-1)$$

Если вероятность заполнения состояний с энергией E равна $f(E, T)$, то число электронов dn , находящихся в состояниях dZ , составит величину

$$dn = f(E, T) dZ = f(E, T) N(E) dE. \quad (4-2)$$

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

$$n = \int_{E_1}^{E_2} f(E, T) N(E) dE. \quad (4-3)$$

Найдем выражение для плотности квантовых состояний в случае, когда поверхности равной энергии зоны проводимости и валентной зоны являются сферами. Определим плотность состояний у нижнего края зоны проводимости. Согласно (2-139) энергия электронов у дна зоны может быть записана в виде

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

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

Выделим шаровой слой, заключенный между двумя изоэнергетическими поверхностями, соответствующими энергии $E(\mathbf{p}) = \text{const}$ и $E(\mathbf{p}) + dE = \text{const}$ (рис. 4-1). Объем этого слоя составляет величину

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

Рис. 4-1. Объем слоя в зоне Бриллюэна, заключенного между двумя изоэнергетическими поверхностями E и $E + dE$.

Объем элементарной ячейки зоны Бриллюэна кристалла единичного объема в \mathbf{p} -пространстве в соответствии с (2-79) равен h^3 . В каждой ячейке могут находиться два электрона с противоположно направленными спинами. С учетом этого число состояний в объеме dV_p равно:

$$dZ = 2 \frac{dV_p}{h^3} = \frac{8\pi}{h^3} p^2 dp. \quad (4-6)$$

Исходя из равенства (4-4) имеем:

$$p^2 = 2m_n^* (E - E_c), \quad (4-7)$$

откуда

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

и

$$dp = \frac{1}{2} (2m_n^*)^{1/2} (E - E_c)^{-1/2} dE. \quad (4-8)$$

Подставив равенства (4-6), (4-7) и (4-8) в соотношение (4-1), получим выражение для плотности квантовых состояний у дна зоны проводимости, обладающей сферической симметрией:

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

Определим плотность состояний вблизи верхнего края валентной зоны. С этой целью воспользуемся формулами (2-139) и (2-144) для энергии дырок:

$$E = E_v - \frac{p^2}{2m_p^*}. \quad (4-10)$$

Здесь $E_v = E(\mathbf{p}_0)$ — энергия дырки у потолка валентной зоны; m_p^* — эффективная масса дырки. Из расчетов, аналогичных проделанным выше, будем иметь:

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

Таким образом, если энергия носителей заряда является квадратичной функцией квазиимпульса, то плотность состояний $N(E)$ имеет зависимость от энергии вида $(E - E_c)^{1/2}$ или $(E_v - E)^{1/2}$. Следовательно, выражения (4-9) и (4-11) справедливы только для состояний вблизи экстремума энергии, т. е. у дна и потолка зоны.

Подсчитаем теперь плотность квантовых состояний для случая сложной зоны проводимости, когда зависимость $E(\mathbf{p})$ у дна зоны описывается соотношением вида (2-132). Поскольку для дна зоны $\rho_{0x} = \rho_{0y} = \rho_{0z} = 0$, а $E(\mathbf{p}_0) = E_c$, то

$$E = E_c + \frac{p_x^2}{2m_1} + \frac{p_y^2}{2m_2} + \frac{p_z^2}{2m_3}, \quad (4-12)$$

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

$$\frac{p_x^2}{a^2} + \frac{p_y^2}{b^2} + \frac{p_z^2}{c^2} = 1, \quad (4-13)$$

где полуоси эллипсоида

$$a_i = [2m_i(E - E_c)]^{1/2}. \quad (4-14)$$

Объем эллипсоида с полуосями a , b , c равен:

$$V_p = \frac{4\pi}{3} abc = \frac{8\pi}{3} (2m_1 m_2 m_3)^{1/2} (E - E_c)^{3/2}. \quad (4-15)$$

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

$$dV_p = 4\pi (2m_1 m_2 m_3)^{1/2} (E - E_c)^{1/2} dE. \quad (4-16)$$

В объеме dV_p с учетом спина заключено следующее количество квантовых состояний:

$$dZ = 2 \frac{dV_p}{h^3} = 4\pi \left(\frac{2}{h^2} \right)^{3/2} (m_1 m_2 m_3)^{1/2} (E - E_c)^{1/2} dE, \quad (4-17)$$

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

$$N(E) = 4\pi \left(\frac{2}{h^2} \right)^{3,2} (m_1 m_2 m_3)^{1,2} (E - E_c)^{1,2}. \quad (4-18)$$

Его можно свести к виду, аналогичному выражению (4-9), если положить:

$$m_1 m_2 m_3 = m_{dn}^*, \quad (4-19)$$

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

$$N(E) = 4\pi \left(\frac{2m_{dn}^*}{h^2} \right)^{3,2} (E - E_c)^{1,2}. \quad (4-20)$$

Таким образом, эллипсоидальную поверхность равной энергии зоны проводимости можно заменить сферой, если ввести эффективную массу плотности состояний для электронов, которая вычисляется по формуле (4-19).

Найдем плотность состояний для кремния и германия.

Кремний. Как следует из рис. 2-21, абсолютный минимум энергии в зоне проводимости кремния расположен в направлении [100] на расстоянии от точки $k = 0$, равном примерно 75% величины $2\pi/a$. Всего в первой зоне Бриллюэна располагается 6 минимумов, и для энергии, несколько превышающей величину E_c , имеется 6 эллипсоидов равной энергии, вытянутых вдоль главных кристаллографических осей (см. рис. 2-22). В соответствии с этим, плотность состояний, которая выражается для одного эллипсоида равенством (4-18), увеличится в 6 раз. И если учесть, что для кремния $m_1 = m_2$, то

$$N(E) = 6 \cdot 4\pi \left(\frac{2}{h^2} \right) (m_1^2 m_3)^{1,2} (E - E_c)^{1,2}, \quad (4-21)$$

а эффективная масса плотности состояний для электронов с учетом значений $m_1 = 0,19 m_0$ и $m_3 = 0,97 m_0$ будет:

$$m_{dn}^* = 6^{2,3} (m_1^2 m_3)^{1,3} = 1,08 m_0. \quad (4-22)$$

Следовательно, у кремния все 6 эллипсоидов изоэнергетической поверхности зоны проводимости можно заменить одной сферической поверхностью с эффективной массой плотности состояний для электронов, равной $1,08 m_0$.

В точке $k = 0$ валентная зона кремния дважды вырождена и поверхности равной энергии имеют довольно сложный вид, описываемый формулой (2-153). Однако их можно аппроксимировать двумя сферическими поверхностями вида (2-154), которым соответствуют эффективные массы тяжелых $m_{рт}^*$ и легких дырок $m_{рл}^*$.

В этом случае плотность состояний будет определяться суммой плотности состояний в зонах тяжелых и легких дырок:

$$N(E) = 4\pi \left(\frac{2}{h^2} \right)^{3/2} \left(m_{p\tau}^{*3/2} + m_{p\lambda}^{*3/2} \right) (E_v - E)^{1/2}. \quad (4-23)$$

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

$$N(E) = 4\pi \left(\frac{2m_{dp}^*}{h^2} \right)^{3/2} (E_v - E)^{1/2}, \quad (4-24)$$

для которой эффективная масса плотности состояний для дырок равна:

$$m_{dp}^* = \left(m_{p\tau}^{*3/2} + m_{p\lambda}^{*3/2} \right)^{2/3} = 0,56m_0, \quad (4-25)$$

так как $m_{p\tau}^* = 0,50 m_0$, а $m_{p\lambda} = 0,16 m_0$.

Германий. В первой зоне Бриллюэна германия имеется 8 минимумов энергии, расположенных в точках $[1/2, 1/2, 1/2]$. Соответствующие поверхности равной энергии изображены на рис. 2-22, где показаны восемь эллипсоидов энергии, расщепленные пополам граням зоны Бриллюэна. Две противоположные половины эллипсоидов рис. 2-22 могут рассматриваться как одна целая эллипсоидальная поверхность. Таким образом, в основной зоне Бриллюэна германия расположено четыре полных эллипсоида, два из которых $m_1 = m_2$. Следовательно, плотность состояний у потолка зоны проводимости германия будет определяться выражением (4-20) с эффективной массой плотности состояний для электронов, равной:

$$m_{dn}^* = (4^3 m_1^2 m_3)^{1/3} = 0,56m_0, \quad (4-26)$$

так как $m_1 = 0,082 m_0$, а $m_3 = 1,60 m_0$.

Структура валентной зоны германия подобна структуре зоны кремния (см. рис. 2-21). Эффективная масса плотности состояний для дырок, характеризующая плотность квантовых состояний у потолка валентной зоны германия, будет:

$$m_{dp}^* = \left(m_{p\tau}^{*3/2} + m_{p\lambda}^{*3/2} \right)^{2/3} = 0,35m_0, \quad (4-27)$$

так как $m_{p\tau}^* = 0,34m_0$; $m_{p\lambda}^* = 0,04m_0$.

4-2. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ ФЕРМИ—ДИРАКА

Для определения числа частиц, имеющих энергию в заданном интервале, помимо плотности квантовых состояний $N(E)$ необходимо знать вероятность того, что данное состояние с энергией E занято частицей, т. е. нужно знать функцию распределения $f(E)$. В условиях теплового равновесия для частиц с полуцелым спином,

подчиняющихся принципу Паули, справедливо распределение Ферми — Дирака

$$f_0(E) = \frac{1}{e^{(E-F)kT} + 1}, \quad (4-28)$$

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

Рассмотрим вид функции распределения Ферми—Дирака при различных температурах. Из соотношения (4-28) следует, что в случае $T = 0$ в интервале энергии $0 \leq E < F$ имеем $f_0 = 1$ и $f_0 = 0$

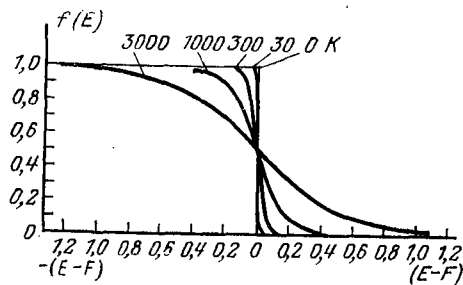


Рис. 4-2. Вид функции распределения Ферми—Дирака.

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 (4-28)

для значения энергии, равной значению энергии Ферми ($E = F$), имеем $f_0 = 1/2$. Таким образом, уровень Ферми есть энергетический уровень, вероятность заполнения которого при температуре, отличной от абсолютного нуля, равна 0,5. При $T > 0$ часть электронов в результате теплового движения перейдет в состояния с энергией, большей энергии Ферми ($E > F$), и соответственно часть состояний, находящихся ниже уровня Ферми, окажутся свободными. В этом случае число частиц, перешедших на более высокие энергетические уровни, будет равно количеству образовавшихся свободных состояний в области $E < F$.

Произведем оценку области изменения функции распределения $f_0(E)$ для случая $T > 0$. Для этого подсчитаем значение $f_0(E)$ для разных значений энергии. Для энергий, отличающихся от F на $\pm kT$, значение $f_0(E)$ составляет $(1 + e)^{-1} = 0,27$ и $(1 + 1/e)^{-1} = 0,73$. При изменении энергии на $\pm 2kT$ значение f_0 равно 0,118 и 0,882, а при изменении на $\pm 3kT$ составляет 0,047 и 0,953. Из этих данных следует, что вероятность заполнения состояний заметно отличается от единицы или нуля лишь в пределах $(2 \div 3) kT$ вблизи значения $E = F$ (рис. 4-2).

Функция распределения Ферми—Дирака характеризует вероятность заполнения данного квантового состояния электроном.

Вероятность того, что при тепловом равновесии в состоянии с энергией E электрон отсутствует, т. е. оно занято дыркой, будет равна:

$$f_{0p}(E) = 1 - f_0(E) = 1 - \frac{1}{e^{(E-F)/kT} + 1} = \frac{1}{e^{(F-E)/kT} + 1}. \quad (4-29)$$

Следовательно, функция распределения для дырок аналогична функции распределения для электронов, если отсчитывать энергию дырок от уровня Ферми в противоположную сторону по сравнению с направлением отсчета энергии для электронов.

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

$$f_0(E) = e^{(F-E)/kT} = Ce^{-E/kT}, \quad (4-30)$$

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

4-3. СТЕПЕНЬ ЗАПОЛНЕНИЯ ПРИМЕСНЫХ УРОВНЕЙ

Рассмотрим полупроводник, содержащий донорную примесь с концентрацией N_d . Донор, удерживающий электрон, электрически нейтрален. Это соответствует случаю, например, когда один из узлов кристаллической решетки кремния занят атомом мышьяка. При этом пятый валентный электрон атома донорной примеси не принимает участия в парноэлектронной связи и ему соответствует энергетический уровень, расположенный ниже дна зоны проводимости на величину E_d (рис. 4-3, а).

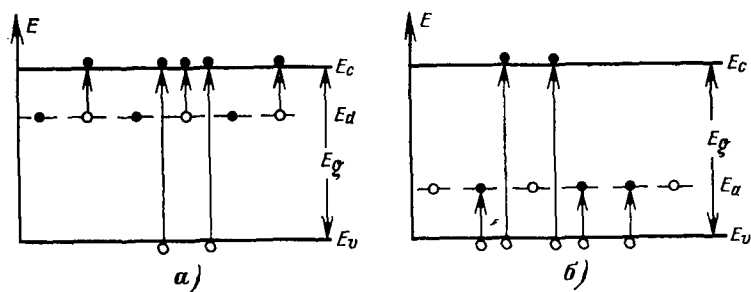


Рис. 4-3. Электронный (а) и акцепторный (б) полупроводники.

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

Пусть число электронов, находящихся на уровне донорной примеси, будет n_d . В этом случае концентрация ионизованных донорных атомов p_d , образовавшихся в результате тепловых переходов электронов с донорных уровней в зону проводимости и имеющих положительный заряд, равна:

$$p_d = N_d - n_d. \quad (4-31)$$

Если бы на примесном донорном уровне согласно принципу Паули могли расположиться два электрона с противоположными спинами, то вероятность его заполнения определялась бы функцией Ферми—Дирака (4-28), в которой вместо E следовало поставить E_d — энергию электрона на уровне примеси. Но на уровне E_d может быть только один электрон, который может быть захвачен двояким образом в зависимости от направления спина. Следовательно, *нейтральное состояние донорной примеси имеет вдвое больший статистический вес по сравнению с ионизованным состоянием*. Так как вероятность отсутствия электрона на уровне донорной примеси равна 1, то, исходя из принципа Больцмана, можно написать:

$$p_d : n_d = 1 : 2e^{-(E_d - F)kT}. \quad (4-32)$$

Используя (4-31), это равенство можно записать в виде

$$\frac{N_d - n_d}{n_d} = \left(2e^{-(E_d - F)kT} \right)^{-1}, \quad (4-33)$$

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

$$n_d = \frac{N_d}{\frac{1}{2} e^{(E_d - F)kT} + 1}, \quad (4-34)$$

а концентрация положительных ионов донорной примеси на основании равенств (4-32) и (4-34) будет выражаться соотношением вида

$$p_d = \frac{N_d}{2e^{(F - E_d)kT} + 1}. \quad (4-35)$$

Предэкспоненциальный множитель в выражении (4-34) в общем случае можно записать через g^{-1} . Тогда вероятность нахождения электрона на донорном уровне с энергией E_d будет определяться выражением

$$f_0(E_d) = \frac{1}{g^{-1} e^{(E_d - F)kT} + 1}, \quad (4-36)$$

а функция распределения для положительных ионов донорной примеси на основании равенств (4-29) и (4-35) будет:

$$f_{0p} = \frac{1}{ge^{(F - E_d)kT} + 1}. \quad (4-37)$$

Из равенства (4-36) следует, что для одновалентной донорной примеси, для которой примесный уровень двукратно вырожден, фактором (степенью) спинового вырождения является $g = 2$.

Рассмотрим теперь акцепторный полупроводник, например кремний, легированный бором. Пусть концентрация введенной примеси равна N_a . Энергетическая схема такого полупроводника представлена на рис. 4-3, б.

Нейтральный атом бора с соседними атомами кремния образует три парноэлектронные связи, четвертая связь одного из четырех соседних атомов кремния остается незавершенной и сна, располагаясь около атома бора, 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. Таким образом, акцепторный уровень может либо удерживать один электрон с произвольным направлением спина (нейтральное состояние акцепторной примеси), либо иметь два электрона со спаренными спинами в случае, когда атом акцепторной примеси для укомплектования парной связи захватывает электрон из валентной зоны (ионизованное состояние акцепторной примеси). Следовательно, степень вырождения акцепторного уровня $g = 2$. У германия и кремния примесные уровни дважды вырождены из-за двукратного вырождения валентной зоны при $k = 0$, вследствие чего для акцепторных уровней фактор вырождения $g = 4$. Поэтому концентрация электронов n_a на акцепторной примеси (или концентрация отрицательных ионов) при данной температуре будет определяться соотношением вида

$$n_a = \frac{N_a}{g e^{(E_a - F)kT} + 1}, \quad (4-38)$$

а концентрация дырок на акцепторной примеси p_a соответственно будет равна:

$$p_a = \frac{N_d}{g^{-1} e^{(F - E_a)kT} + 1}. \quad (4-39)$$

4-4. 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. 4-3, а.

При температуре абсолютного нуля электроны занимают состояния с минимальным значением энергии. Другими словами, при $T = 0$ К все состояния донорной примеси и в валентной зоне заняты электронами, поэтому в зоне проводимости нет электронов. При повышении температуры колебания атомов кристалла становятся более интенсивными. Это может привести к отрыву элект-

трона от донорной примеси и даже к разрыву валентных связей основного вещества, вследствие чего могут появиться 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 название 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. 4-3, 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 электронов из валентной зоны на уровень акцепторной примеси и в зону проводимости (рис. 4-3, б).

Определим при термодинамическом равновесии для примесного полупроводника значения концентрации электронов в зоне 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 можно найти, если в равенство (4-3) подставить плотность квантовых состояний для случая изоэнергетических поверхностей (4-9) и воспользоваться функцией распределения Ферми—Дирака (4-28). Интегрирование нужно провести от дна зоны проводимости E_c до ее потолка. Если учесть резкую зависимость функции $f_0(E)$ от энергии, то верхний предел интегрирования можно положить равным бесконечности. В результате будем иметь:

$$n_0 = \int_{E_c}^{\infty} f_0(E) N(E) dE = \int_{E_c}^{\infty} \frac{4\pi \left(\frac{2m_n^*}{h^2}\right)^{3/2} (E - E_c)^{1/2} dE}{e^{\frac{E-F}{kT}} + 1}. \quad (4-40)$$

Введем безразмерные величины

$$\frac{E - E_c}{kT} = \varepsilon; \quad \frac{F - E_c}{kT} = \eta. \quad (4-41)$$

Величину η называют приведенным уровнем Ферми. При этом равенство (4-40) преобразуется к виду

$$n_0 = 4\pi \left(\frac{2m_n^*}{h^2}\right)^{3/2} (kT)^{3/2} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1} = N_c F_{1/2}(\eta). \quad (4-42)$$

Здесь

$$N_c = 2 \left(\frac{2\pi m_n^* kT}{h^2} \right)^{3/2} \quad (4-43)$$

и называется эффективной плотностью состояний в зоне проводимости;

$$F_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon - \eta} + 1} \quad (4-44)$$

есть интеграл Ферми — Дирака порядка $1/2$. Он представлен на рис. 4-4.

Эффективная плотность состояний N_c зависит от температуры. Если в (4-43) подставить численные значения универсальных констант, то будем иметь:

$$\begin{aligned} N_c &= 2 \left(\frac{2\pi m_n^* kT}{h^2} \right)^{3/2} = \\ &= 4,82 \cdot 10^{15} \left(\frac{m_n^*}{m_0} \right)^{3/2} T^{3/2} = \\ &= 2,5 \cdot 10^{19} \left(\frac{m_n^*}{m_0} \right)^{3/2} \left(\frac{T}{300} \right)^{3/2}. \end{aligned} \quad (4-45)$$

Найдем количество дырок в валентной зоне. В соответствии с (4-11) и (4-29) получаем:

$$p_0 = \int_{E_{v \text{ мин}}}^{E_v} f_{0p}(E) N(E) dE = 4\pi \left(\frac{2m_p^*}{h^2} \right)^{3/2} \int_{-\infty}^{E_v} \frac{(E_v - E)^{1/2} dE}{e^{(F-E)/kT} + 1}. \quad (4-46)$$

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

Обозначим:

$$\frac{E_v - E}{kT} = \varepsilon_p; \quad \frac{E_c - E_v}{kT} = \frac{E_g}{kT} = \varepsilon_i; \quad \frac{F - E}{kT} = \eta + \varepsilon_i + \varepsilon_p. \quad (4-47)$$

Здесь $E_c - E_v = E_g$ — ширина запрещенной зоны. Теперь выражение для концентрации дырок примет вид:

$$p_0 = \frac{4}{\sqrt{\pi}} \left(\frac{2\pi m_p^* kT}{h^2} \right)^{3/2} \int_0^{\infty} \frac{\varepsilon_p^{1/2} d\varepsilon_p}{e^{\varepsilon_p + \eta + \varepsilon_i} + 1} = N_v F_{1/2}(-\eta - \varepsilon_i), \quad (4-48)$$

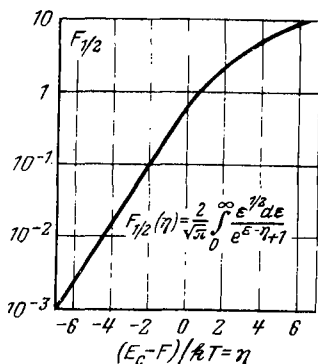


Рис. 4-4. Зависимость интеграла Ферми — Дирака $F_{1/2}$ от уровня Ферми.

где

$$N_v = 2 \left(\frac{2\pi m_p^* kT}{h^2} \right)^{3/2} \quad (4-49)$$

есть эффективная плотность состояний в валентной зоне;

$$F_{1/2}(-\eta - \varepsilon_i) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon} e^{\eta + \varepsilon_i} + 1} \quad (4-50)$$

— интеграл Ферми—Дирака для валентной зоны.

Если положить $m_n^* = m_p^* = m_0$, то при $T = 300$ К $N_c \approx N_v = 2,5 \cdot 10^{19} \text{ см}^{-3}$. Это составляет приблизительно 1/2000 числа атомов в 1 см^3 .

Если ввести обозначение $\frac{E_c - E_d}{kT} = \varepsilon_d$, то на основании (4-36) концентрация электронов на уровне донорной примеси будет равна:

$$n_d = f_0(E_d) N_d = \frac{N_d}{g^{-1} e^{-\varepsilon_d - \eta} + 1}, \quad (4-51)$$

а количество положительных ионов донорной примеси согласно выражению (4-37) составит:

$$p_d = f_{0p}(E_d) N_d = \frac{N_d}{g e^{\varepsilon_d + \eta} + 1}. \quad (4-52)$$

Найдем концентрацию дырок p_a и концентрацию электронов n_a на акцепторной примеси. Если обозначить $\frac{E_c - E_a}{kT} = \varepsilon_a$, то они на основании (4-38) и (4-39) будут выражаться соотношениями вида

$$p_a = \frac{N_a}{g^{-1} e^{\varepsilon_a + \eta} + 1}; \quad (4-53)$$

$$n_a = \frac{N_a}{g e^{-\varepsilon_a - \eta} + 1}. \quad (4-54)$$

4-5. ПРИМЕСНЫЙ ПОЛУПРОВОДНИК

Для того чтобы подсчитать равновесные концентрации электронов n_0 и дырок p_0 , необходимо вычислить значения интегралов Ферми—Дирака (4-44) и (4-50), что во многих конкретных случаях представляет значительные трудности. Но для интеграла Ферми—Дирака существуют приближенные аналитические выражения и составлены таблицы.

В зависимости от значения приведенного уровня Ферми интегралы Ферми—Дирака (4-44) и (4-50), определяющие равновес-

ную концентрацию электронов и равновесную концентрацию дырок, можно выразить следующим образом.

1. При $-\infty < \eta < -1$

$$F_{1,2}(\eta) = e^\eta, \quad (4-55)$$

а при $-\infty < -\eta - \varepsilon_i < -1$

$$F_{1,2}(-\eta - \varepsilon_i) = e^{-\eta - \varepsilon_i}. \quad (4-56)$$

Такие приближения соответствуют статистике Больцмана, а полупроводник является невырожденным. Это классическое приближение, для которого равновесные концентрации электронов и дырок соответственно равны:

$$n_0 = N_c e^\eta; \quad (4-57)$$

$$p_0 = N_v e^{-\eta - \varepsilon_i}. \quad (4-58)$$

На рис. 4-5 классическое приближение изображено кривой 1.

2. При $5 < \eta < \infty$

$$F_{1,2}(\eta) = \frac{4}{3\sqrt{\pi}} \eta^{3,2}, \quad (4-59)$$

а при $5 < -\eta - \varepsilon_i < \infty$

$$F_{1,2}(-\eta - \varepsilon_i) = \frac{4}{3\sqrt{\pi}} (-\eta - \varepsilon_i)^{3,2}. \quad (4-60)$$

Эти приближения справедливы для случая полного вырождения донорного и акцепторного полупроводников соответственно. На рис. 4-5 приближение полного вырождения донорного полупроводника дано кривой 2.

3. При $-1 < \eta < 5$ интеграл Ферми—Дирака для зоны проводимости можно приближенно записать в виде

$$F_{1,2}(\eta) = \frac{1}{0,27 + e^\eta}, \quad (4-61)$$

а при $-1 < -\eta - \varepsilon_i < 5$ интеграл Ферми—Дирака для валентной зоны равен:

$$F_{1,2}(-\eta - \varepsilon_i) = \frac{1}{0,27 + e^{-\eta - \varepsilon_i}}. \quad (4-62)$$

При этих условиях свойства полупроводника являются переходными от невырожденного к полностью вырожденному. Расчет интегралов Ферми—Дирака по формулам (4-61) и (4-62) дает значительную погрешность, а поэтому целесообразно пользоваться специальными таблицами интегралов Ферми—Дирака.

На рис. 4-5 кривая n_0 представляет собой истинное значение концентрации равновесных электронов проводимости, подсчитанное по соотношению (4-42).

Рассмотрим более подробно крайние случаи.

Невырожденный примесный полупроводник. Согласно рис. 4-5 можно считать, что донорный полупроводник невырожден, если

$$\eta = \frac{F - E_c}{kT} < -1$$

или

$$F < E_c - kT. \quad (4-63)$$

Это означает, что в невырожденном донорном полупроводнике уровень Ферми лежит в запрещенной зоне ниже дна зоны проводимости по крайней мере на kT . В этом случае для электронов, находящихся в зоне проводимости и обладающих энергией $E \gg E_c$, в функции

Рис. 4-5. Зависимость концентрации электронов проводимости n_0 от положения приведенного уровня Ферми.

распределения Ферми—Дирака экспоненциальный член будет намного больше единицы, а поэтому ею можно пренебречь, т. е.

$$f_0 = \frac{1}{e^{(E-F)/kT} + 1} \approx e^{-(E-F)/kT}. \quad (4-64)$$

Из соотношения (4-64) следует, что в невырожденном донорном полупроводнике носители заряда подчиняются статистике Больцмана. В соответствии с этим интеграл Ферми—Дирака запишется в виде

$$F_{1,2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{e^{-\eta} + 1}} = \frac{2}{\sqrt{\pi}} e^{\eta} \int_0^{\infty} e^{-\varepsilon^{1/2}} d\varepsilon = e^{\eta} = e^{-(E_c - F)/kT}. \quad (4-65)$$

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

$$n_0 = N_c F_{1,2}(\eta) = N_c e^{\eta} = N_c e^{-(E_c - F)/kT}; \quad (4-66)$$

$$p_0 = N_v F_{1,2}(-\eta - \varepsilon_i) = N_v e^{-\eta - \varepsilon_i} = N_v e^{(E_v - F)/kT}. \quad (4-67)$$

Анализ соотношений (4-66) и (4-67) показывает, что для полностью невырожденного распределения носителей заряда значение интеграла Ферми совпадает с выражением вероятности за-

полнения уровня энергии E_c . Следовательно, как вытекает из равенств (4-66) и (4-67), в невырожденном примесном полупроводнике концентрация электронов в зоне проводимости такая, какая была бы на N_c уровнях с энергией E_c , а полное число свободных дырок такое же, как если бы вместо валентной зоны было бы N_v уровней и все с энергией E_v .

Акцепторный полупроводник не вырожден при $-\eta - \varepsilon_i < -1$ или $F > E_v + kT$. Следовательно, дырочный полупроводник не вырожден, если уровень Ферми лежит в запрещенной зоне выше потолка валентной зоны не менее чем на kT , вследствие чего равновесные концентрации электронов и дырок подчиняются статистике Максвелла—Больцмана и определяются равенствами (4-66) и (4-67).

В собственном полупроводнике свободные носители заряда возникают в результате разрыва валентных связей, поэтому концентрации свободных электронов и дырок равны: $n_0 = p_0 = n_i$. Величина n_i носит название собственной концентрации носителей заряда. Если для собственного полупроводника обозначить приведенный уровень Ферми через η_i , то на основании равенств (4-66) и (4-67) следует:

$$n_i = N_c e^{\eta_i} = N_v e^{-\eta_i - \varepsilon_i}. \quad (4-68)$$

Выразим теперь равновесные концентрации электронов n_0 (4-66) и дырок p_0 (4-67) для невырожденного примесного полупроводника через n_i . Тогда

$$\left. \begin{aligned} n_0 &= N_c e^{\eta} = n_i e^{\eta - \eta_i}; \\ p_0 &= N_v e^{-\eta - \varepsilon_i} = n_i e^{\eta_i - \eta}. \end{aligned} \right\} \quad (4-69)$$

Отсюда следует:

$$n_0 p_0 = n_i^2. \quad (4-70)$$

Таким образом, в невырожденном примесном полупроводнике произведение концентраций свободных электронов и дырок при тепловом равновесии есть постоянная величина, равная квадрату концентрации носителей заряда в собственном полупроводнике при той же температуре.

Зонная диаграмма, плотность состояний $N(E)$, распределение Ферми—Дирака $f_0(E)$ и концентрация носителей заряда в собственном (а), электронном (б) и дырочном (в) полупроводниках при термодинамическом равновесии показаны на рис. 4-6.

Если возбуждение носителя заряда происходит за счет тепловой энергии, то вероятность такого процесса пропорциональна $e^{-\omega/kT}$. Величина ω называется энергией активации. Поэтому концентрацию электронов n_0 можно выразить и через энергию активации. Она будет равна:

$$n_0 = C e^{-\omega/kT}, \quad (4-71)$$

где C — постоянная величина.

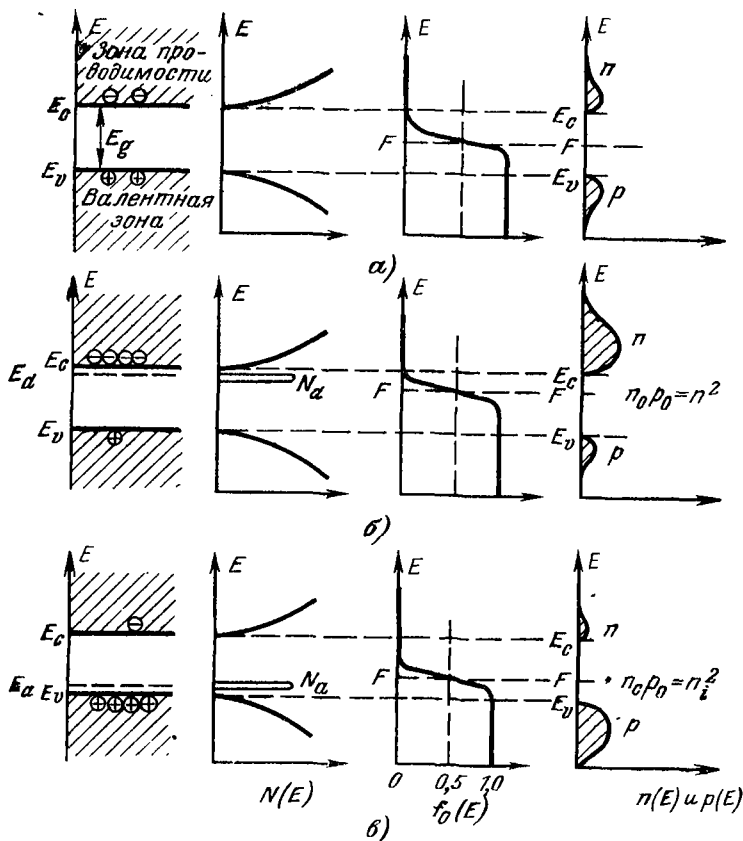


Рис. 4-6. Зонная диаграмма, плотность состояний, распределение Ферми—Дирака и концентрация носителей заряда в собственном (а), донорном (б) и акцепторном (в) полупроводниках при термодинамическом равновесии.

Соотношение (4-66) может быть записано как

$$n_0 T^{-3/2} = 2 \left(\frac{2\pi m_n^* k}{h^2} \right)^{3/2} e^{-(E_c - F)/kT} = C e^{-(E_c - F)/kT}.$$

Из сравнения этой формулы и (4-71) следует, что

$$\omega = -k \frac{d}{d\left(\frac{1}{T}\right)} \ln(n_0 T^{-3/2}) = E_c - F + T \frac{dF}{dT}, \quad (4-72)$$

т. е. энергия активации зависит от положения уровня Ферми.

Полностью вырожденный примесный полупроводник. Согласно (4-59) можно считать, что донорный

полупроводник полностью вырожден при $\eta = \frac{F - E_c}{kT} < 5$, т. е. когда уровень Ферми расположен в зоне проводимости выше дна зоны проводимости E_c по крайней мере на $5kT$.

Если в зоне проводимости при $T = 0$ К имеются свободные электроны, то полупроводник, для которого приведенный уровень Ферми $\eta = \frac{F - E_c}{kT} \rightarrow \infty$, будет полностью вырожденным. В этом случае в выражении (4-42) знаменатель равен единице и равновесная концентрация электронов в полностью вырожденном донорном полупроводнике будет равна:

$$n_0 = N_c \frac{2}{\sqrt{\pi}} \int_0^{\eta_0} e^{-\epsilon^2} d\epsilon = \frac{4}{3\sqrt{\pi}} N_c \eta_0^{3/2} = \frac{4}{3\sqrt{\pi}} N_c \left(\frac{F_0 - E_c}{kT} \right)^{3/2} =$$

$$= \frac{8\pi}{3} \left(\frac{2m_n^*}{h^2} \right)^{3/2} (F_0 - E_c)^{3/2}, \quad (4-73)$$

а концентрация равновесных дырок p_0 определяется соотношением (4-67).

Уровень Ферми при температуре абсолютного нуля, как следует из соотношения (4-73), определяется равенством вида

$$F_0 = E_c + \frac{h^2}{2m_n^*} \left(\frac{3n_0}{8\pi} \right)^{2/3}. \quad (4-74)$$

Акцепторный полупроводник является полностью вырожденным, если $-\eta - \epsilon_i > 5$ и $F < E_v - 5kT$, т. е. если уровень Ферми лежит в валентной зоне ниже E_v на $5kT$. В полностью вырожденном акцепторном полупроводнике равновесная концентрация свободных дырок в соответствии с выражениями (4-48) и (4-60) равна:

$$p_0 = N_v \frac{4}{3\sqrt{\pi}} (-\eta_0 - \epsilon_i)^{3/2} =$$

$$= \frac{8\pi}{3} \left(\frac{2m_p^*}{h^2} \right)^{3/2} (E_v - F_0)^{3/2}, \quad (4-75)$$

а концентрация электронов проводимости n_0 определяется соотношением (4-66).

Из равенств (4-73) и (4-75) следует, что концентрация основных носителей заряда в полностью вырожденном примесном полупроводнике не зависит от температуры. Она определяется положением уровня Ферми F_0 и величиной эффективной массы носителей заряда. При этом чем меньше эффективная масса носителей заряда, тем соответственно и меньше их концентрация, при которой наступает вырождение.

Условие полного вырождения с учетом (4-74) можно записать следующим образом:

$$\eta = \frac{F - E_c}{kT} = \frac{\hbar^2}{2m_n^* kT} \left(\frac{3n_0}{8\pi} \right)^{2/3} \gg 1. \quad (4-76)$$

Здесь мы заменили η_0 на η , что, как будет показано ниже, для случая сильного вырождения не вносит существенной ошибки. Из анализа условия (4-76) следует, что *вырождение наступает тем раньше, чем больше концентрация свободных носителей заряда, чем ниже температура (при данной концентрации) и чем меньше эффективная масса носителей заряда.*

Граница перехода от невырожденного полупроводника к вырожденному определяется концентрацией носителей заряда вырождения n_v при температуре вырождения T_v . Поскольку n_v является минимальной концентрацией носителей заряда, соответствующей началу вырождения при данной температуре, на основании (4-73) можно записать:

$$\begin{aligned} n_v &= \frac{4}{3\sqrt{\pi}} N_c = 3,78 \cdot 10^{15} \left(\frac{m_n^*}{m_0} \right)^{3/2} T_v^{3/2} = \\ &= 2 \cdot 10^{19} \left(\frac{m_n^*}{m_0} \right)^{3/2} \left(\frac{T_v}{300} \right)^{3/2} \end{aligned} \quad (4-77)$$

или

$$T_v = 4,1 \cdot 10^{-11} \left(\frac{m_0}{m_n^*} \right) n_v^{2/3}. \quad (4-78)$$

Эти выводы подтверждаются экспериментальными данными. На рис. 4-7 представлены кривые температурной зависимости концентрации дырок для кремния с разным содержанием бора.

Анализ этих зависимостей показывает, что с увеличением степени легирования концентрация дырок перестает зависеть от температуры, т. е. полупроводник становится полностью вырожденным. Область вырождения с увеличением температуры сдвигается к более высоким концентрациям носителей заряда. Если при температуре жидкого азота вырождение наступает при $p_0 = 2,5 \cdot 10^{18} \text{ см}^{-3}$, то при комнатной температуре оно имеет место при $p_0 = 2 \cdot 10^{19} \text{ см}^{-3}$. Антимонид индия, обладающий очень малой эффективной массой электронов ($m_n^* = 0,013 m_0$), при $T = 300 \text{ К}$ оказывается вырожденным уже при $n_0 = 2 \cdot 10^{16} \text{ см}^{-3}$.

Найдем зависимость уровня Ферми для вырожденного донорного полупроводника от температуры. Если $\eta \gg 1$, то $e^{\eta} \ll 1$ и интеграл Ферми можно разложить в степенной ряд, а поэтому рав-

новесная концентрация электронов n_0 может быть представлена формулой

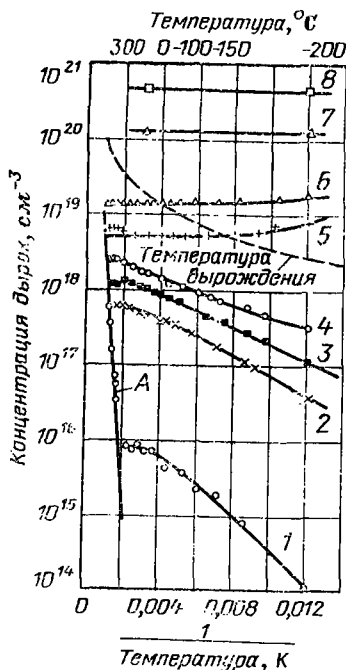
$$n_0 = N_c \frac{4}{3\sqrt{\pi}} \eta^{3/2} \left(1 + \frac{\pi^2}{8\eta^2} - \frac{7\pi^4}{640\eta^4} + \frac{31\pi^6}{3072\eta^6} - \dots \right). \quad (4-79)$$

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

$$n_0 \approx N_c \frac{4}{3\sqrt{\pi}} \eta^{3/2} \left(1 + \frac{\pi^2}{8\eta_0^2} \right). \quad (4-80)$$

Рис. 4-7. Температурная зависимость концентрации дырок для кремния, легированного бором.

A — собственная концентрация. Концентрация атомов бора для различных кривых равна: 1 — не легированный; 2 — $6,7 \cdot 10^{17} \text{ см}^{-3}$; 3 — $1,3 \cdot 10^{18} \text{ см}^{-3}$; 4 — $2,7 \cdot 10^{18} \text{ см}^{-3}$; 5 — $6,7 \cdot 10^{18} \text{ см}^{-3}$; 6 — $1,3 \cdot 10^{19} \text{ см}^{-3}$; 7 — $1,3 \cdot 10^{20} \text{ см}^{-3}$; 8 — $1,3 \cdot 10^{21} \text{ см}^{-3}$. Пунктирная кривая показывает связь концентрации дырок с температурой, при которой наступает вырождение.



Из соотношений (4-80) и (4-73) имеем:

$$\frac{3\sqrt{\pi}}{4} \frac{n_0}{N_c} = \eta^{3/2} \left(1 + \frac{\pi^2}{8\eta_0^2} \right) = \eta_0^{3/2}, \quad (4-81)$$

откуда

$$\eta = \eta_0 \left(1 + \frac{\pi^2}{8\eta_0^2} \right)^{-2/3}. \quad (4-81a)$$

При $\eta_0 > 5$ второй член в скобке мал. Разложив выражение (4-81a) в ряд и ограничившись двумя членами, найдем окончательно:

$$\eta = \eta_0 \left(1 - \frac{\pi^2}{12\eta_0^2} \right) \quad (4-82)$$

или

$$F = F_0 \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{F_0 - E_c} \right)^2 \right]. \quad (4-83)$$

Как следует из соотношения (4-83), при достаточно низкой температуре уровень Ферми F в вырожденном полупроводнике очень близок к F_0 .

4-6. СОБСТВЕННЫЙ ПОЛУПРОВОДНИК

Полупроводник будет являться собственным, если влияние примесей для него пренебрежимо мало. В нем свободные носители заряда возникают только за счет разрыва валентных связей (рис. 4-8), поэтому

$$n_0 = p_0. \quad (4-84)$$

Соотношение (4-84) выражает собой условие электронейтральности собственного полупроводника, по которому суммарный заряд всех заряженных частиц кристалла должен быть равен нулю.

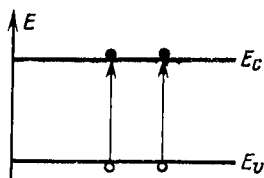


Рис. 4-8. Тепловая генерация носителей заряда в собственном полупроводнике.

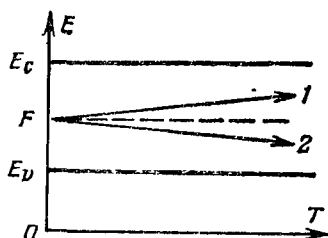


Рис. 4-9. Изменение уровня Ферми с температурой в собственном полупроводнике.

Это требование определяет положение уровня Ферми. Его можно найти, решив уравнение

$$N_c F_{1/2}(\eta) = N_v F_{1/2}(-\eta - \epsilon_i). \quad (4-85)$$

Невырожденный собственный полупроводник. Для него зонная диаграмма, плотность состояний, распределение Ферми—Дирака и концентрация носителей заряда при термодинамическом равновесии представлены на рис. 4-6, а. Для невырожденного полупроводника интеграл Ферми—Дирака можно аппроксимировать экспонентой (4-65). При этом условии уравнение (4-85) примет вид:

$$N_c e^{\eta} = N_v e^{-\eta - \epsilon_i}, \quad (4-86)$$

откуда легко получить для уровня Ферми

$$F = E_c - \frac{1}{2} E_g + \frac{kT}{2} \ln \frac{N_v}{N_c} = \frac{E_c + E_v}{2} + \frac{kT}{2} \ln \frac{N_v}{N_c}, \quad (4-87)$$

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

Если подставить значения N_c из выражения (4-43) и N_v из выражения (4-49), то получим:

$$F = \frac{E_c + E_v}{2} + \frac{3}{4} kT \ln \frac{m_p^*}{m_n^*}. \quad (4-88)$$

Из соотношения (4-88) следует, что при температуре абсолютного нуля уровень Ферми для собственного полупроводника располагается посередине между дном зоны проводимости и потолком валентной зоны, т. е. $F = \frac{1}{2} (E_c + E_v)$. В случае $m_p^* = m_n^*$ его положение не зависит от температуры и F лежит в середине запрещенной зоны. При $m_p^* > m_n^*$ уровень Ферми с повышением температуры линейно смещается к дну зоны проводимости, а при $m_p^* < m_n^*$ — к потолку валентной зоны (рис. 4-9, кривые 1 и 2 соответственно).

Определим собственную концентрацию носителей заряда:

$$\begin{aligned} n_i &= (n_0 p_0)^{1/2} = (N_c N_v)^{1/2} e^{-\frac{E_c - E_v}{2kT}} = 2 \left(\frac{2\pi k}{h^2} \right)^{3/2} (m_n^* m_p^*)^{3/4} T^{3/2} e^{-\frac{E_g}{2kT}} = \\ &= 4,9 \cdot 10^{15} \left(\frac{m_n^* m_p^*}{m_0^2} \right)^{3/4} T^{3/2} e^{-\frac{E_g}{2kT}}. \end{aligned} \quad (4-89)$$

Таким образом, концентрация носителей заряда в невырожденном собственном полупроводнике зависит от температуры T , ширины запрещенной зоны E_g , значений эффективных масс электронов m_n^* и дырок m_p^* и не зависит от положения уровня Ферми F . Температурная зависимость собственной концентрации носителей заряда при $E_g \gg kT$ определяется в основном экспоненциальным членом уравнения (4-89). Доказательством этого являются данные рис. 4-10, на котором представлена зависимость собственной концентрации носителей заряда в германии, кремнии и арсениде галлия от обратной температуры.

Энергия активации для собственного полупроводника, как следует из соотношений (3-72) и (3-87),

$$\omega = \frac{1}{2} E_g, \quad (4-90)$$

т. е. равна половине ширины запрещенной зоны.

Оценим собственную концентрацию носителей заряда в кремнии, германии и антимониде индия. Для кремния $m_{dn}^* = 1,08 m_0$, $m_{dp}^* = 0,56 m_0$ при $T = 300 \text{ K}$, $E_g = 1,11 \text{ эВ}$, поэтому $n_i = 1,9 \cdot 10^{10} \text{ см}^{-3}$. Для германия $m_{dn}^* = 0,56 m_0$, $m_{dp}^* = 0,35 m_0$, $E_g = 0,67 \text{ эВ}$ и $n_i = 2 \cdot 10^{13} \text{ см}^{-3}$. Для антимонида индия $m_n^* = 0,013 m_0$, $m_{dp}^* = 0,6 m_0$, $E_g = 0,18 \text{ эВ}$ и $n_i = 2 \cdot 10^{16} \text{ см}^{-3}$.

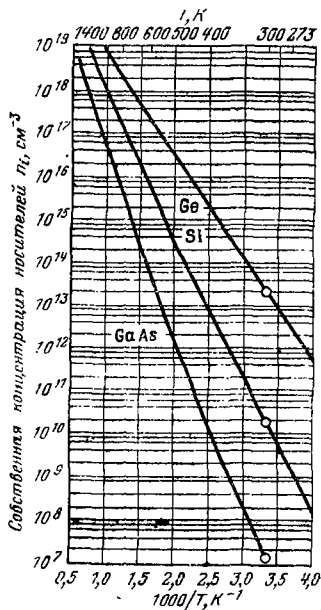


Рис. 4-10. Зависимость собственной концентрации носителей заряда в Ge, Si и GaAs от обратной температуры.

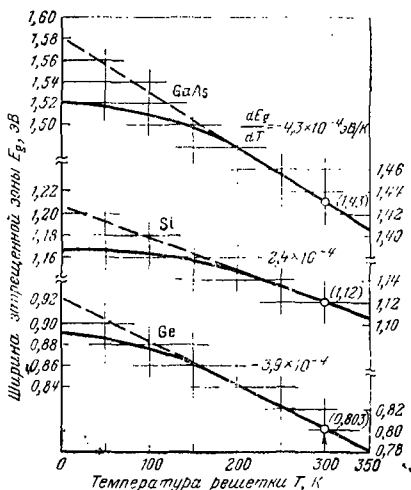


Рис. 4-11. Зависимость ширины запрещенной зоны от температуры для Ge, Si и GaAs.

Согласно уравнению (4-89) графическая зависимость $\ln(n_i T^{-3/2})$ от $1/T$ должна выражаться прямой линией:

$$\ln(n_i T^{-3/2}) = \text{const} - \frac{E_g}{2k} \cdot \frac{1}{T}. \quad (4-91)$$

Угол наклона этой прямой $\text{tg } \theta = E_g/2k$. Однако вычисленная по данным рис. 4-10 величина E_g не соответствует истинной ширине запрещенной зоны германия. Причина этого расхождения заключается в том, что сама ширина запрещенной зоны изменяется с температурой. Рост амплитуды тепловых колебаний атомов решетки приводит к ее уменьшению. Кроме того, с увеличением температуры изменяются межатомные расстояния, что также оказывает влияние на ширину запрещенной зоны. В результате зависимость $E_g(T)$ может иметь сложный характер. В качестве примера на рис. 4-11

показаны изменения ширины запрещенной зоны в зависимости от температуры для германия, кремния и арсенида галлия. Для этих полупроводников значения ширины запрещенной зоны при 0 К составляют 0,89; 1,16 и 1,52 эВ соответственно. У них, как видно из рис. 4-11, в диапазоне температур 175—350 К ширина запрещенной зоны меняется линейно с температурой. При этом температурный коэффициент изменения ширины запрещенной зоны $\alpha = dE/dT$ является отрицательным и равен $3,9 \cdot 10^{-4}$; $2,4 \cdot 10^{-4}$; $2,9 \cdot 10^{-4}$ и $4,3 \cdot 10^{-4}$ эВ/К соответственно для германия, кремния, антимонида индия и арсенида галлия. У PbS температурный коэффициент положителен и ширина запрещенной зоны E_g возрастает от 0,34 эВ при 0 К до 0,41 эВ при 300 К. В этих случаях зависимость подчиняется линейному закону вида

$$E_g = E_{g0} - \alpha T, \quad (4-92)$$

где E_{g0} — ширина запрещенной зоны при 0 К.

Подставив выражение (4-92) в уравнение (4-89), получим:

$$n_i T^{-3/2} = 2 \left(\frac{2\pi^2 m_n^* m_p^* k^2}{h^2} \right)^{3/2} e^{-\alpha/2k} e^{-E_{g0}/2kT}, \quad (4-93)$$

откуда найдем, что

$$\ln(n_i T^{-3/2}) = \text{const} - \frac{\alpha}{2k} - \frac{E_{g0}}{2k} \frac{1}{T}. \quad (4-94)$$

Таким образом, если ширина запрещенной зоны полупроводника линейно зависит от температуры, график зависимости $\ln(n_i T^{-3/2})$ от $1/T$ также представляет собой прямую линию, наклон которой характеризуется величиной E_{g0} , которая является шириной запрещенной зоны при 0 К. Истинная ширина запрещенной зоны полупроводника при данной температуре определяется по формуле (4-92).

Вырожденный собственный полупроводник. У собственного полупроводника согласно соотношению (4-88) скорость изменения уровня Ферми с температурой пропорциональна отношению эффективных масс дырок и электронов. В результате этого с повышением температуры уровень Ферми отдалится от зоны с тяжелыми носителями заряда, приближаясь к зоне с легкими носителями заряда. И если расстояние от уровня Ферми до этой зоны становится соизмеримо с величиной kT , то в ней наступает вырождение и соответствующий интеграл Ферми—Дирака уже не может быть заменен экспонентой. При этом чем сильнее различаются эффективные массы электронов и дырок, тем раньше наступает вырождение. Если, например, вырождение наступило в зоне проводимости, то в валентной зоне оно отсутствует и выражение для собственной концентрации носителей будет иметь вид:

$$n_i = N_c F_{1/2}(\eta) = N_c e^{-\eta - \epsilon_i}. \quad (4-95)$$

Отсюда получается следующая формула, определяющая положение уровня Ферми:

$$e^{\eta} F_{1,2}(\eta) = \left(\frac{m_p^*}{m_n^*} \right)^{3/2} e^{-\varepsilon_i}. \quad (4-96)$$

Классическим примером полупроводника с сильно различающимися эффективными массами является антимонид индия, для которого $m_p^* \approx 10 m_n^*$. На рис. 4-12 показано изменение положения

приведенного уровня Ферми в InSb в зависимости от температуры. При этом учитывается зависимость $E_g(T)$, которая согласно экспериментальным данным имеет вид:

$$E_g = (0,26 - 2,7 \cdot 10^{-4} T) \text{ эВ.}$$

Как видно из рис. 4-12, уже при $T = 200$ К уровень Ферми лежит ниже E_c всего на $4kT$, а около 450 К пересекает уровень E_c и входит в зону проводимости, что приводит к сильному вырождению электронного газа. Это необходимо учитывать, и при расчете величины n_i следует использовать равенство (4-95).

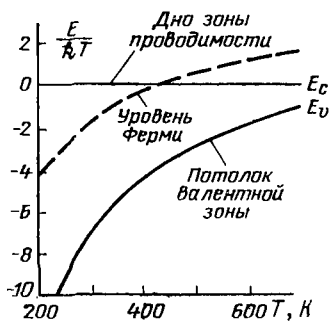


Рис. 4-12. Положение уровня Ферми (в единицах kT) в зависимости от температуры для антимонида индия.

4-7. ЗАВИСИМОСТЬ ПОЛОЖЕНИЯ УРОВНЯ ФЕРМИ ОТ КОНЦЕНТРАЦИИ ПРИМЕСИ И ТЕМПЕРАТУРЫ ДЛЯ НЕВЫРОЖДЕННОГО ПОЛУПРОВОДНИКА

Полупроводник может содержать как донорную, так и акцепторную примеси. Пусть концентрации этих примесей соответственно равны N_d и N_a . Энергетическая схема такого полупроводника представлена на рис. 4-13. Электроны в зоне проводимости и дырки в валентной зоне возникают в результате теплового возбуждения электронов из валентной зоны и атомов примеси. Эти переходы на рисунке показаны стрелками.

Для определения положения уровня Ферми в полупроводнике используется условие электронейтральности, по которому в любой точке кристалла суммарный заряд всех заряженных частиц должен быть равен нулю. Для полупроводника с двумя видами примеси уравнение электронейтральности запишется в виде

$$n_0 + n_a = p_0 + p_d, \quad (4-97)$$

где $n_a = N_a - p_a$ — количество электронов, связанных с акцепторной примесью; $p_d = N_d - n_d$ — число ионизированных атомов

донорной примеси. С учетом выражений для n_a и p_d условие электронейтральности (4-97) можно представить в виде

$$(n_0 + n_d) - (p_0 + p_a) = N_d - N_a. \quad (4-98)$$

Выразив концентрации носителей заряда в соответствии с формулами (4-42), (4-48), (4-51) и (4-53), получим:

$$N_c F_{1/2}(\eta) + \frac{N_d}{g^{-1} e^{-\varepsilon_d - \eta}} - N_v F_{1/2}(-\eta - \varepsilon_i) - \frac{N_a}{g^{-1} e^{\varepsilon_a + \eta} + 1} = N_d - N_a. \quad (4-99)$$

Для невырожденного полупроводника равенство (4-99) значительно упрощается:

$$N_c e^{\eta} + \frac{N_d}{g^{-1} e^{-\varepsilon_d - \eta} + 1} - N_v e^{-\eta - \varepsilon_i} - \frac{N_a}{g^{-1} e^{\varepsilon_a + \eta} + 1} = N_d - N_a, \quad (4-100)$$

однако и в этом случае определение уровня Ферми встречает большие трудности. Поэтому в дальнейшем 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 (4-98) запишется при этом так:

$$n_0 + n_d = N_d \quad (4-101)$$

или

$$n_0 = p_d,$$

откуда согласно (4-52) будем иметь:

$$n_0 = \frac{N_d}{g e^{\varepsilon_d + \eta} + 1}. \quad (4-102)$$

Считаем, что вырождение отсутствует, следовательно, $n_0 = N_c e^{\eta}$. Из этого равенства найдем, что

$$e^{\eta} = n_0 / N_c. \quad (4-103)$$

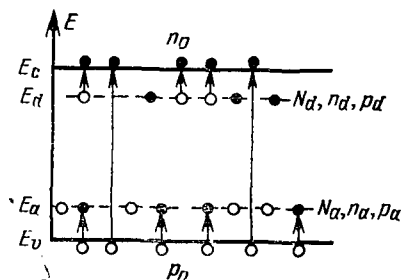


Рис. 4-13. Полупроводник, имеющий примеси донорного и акцепторного типа.

Подставив (4-103) в формулу (4-102), получим уравнение для концентрации электронов в зоне проводимости:

$$n_0 = \frac{N_c N_d}{N_c + g e^{\varepsilon_d} n_0}, \quad (4-104)$$

которое легко преобразуется к виду

$$n_0^2 + g^{-1} e^{-\varepsilon_d} N_c n_0 - g^{-1} e^{-\varepsilon_d} N_c N_d = 0. \quad (4-105)$$

Решая полученное квадратное уравнение относительно n_0 , находим:

$$n_0 = \frac{g^{-1} e^{-\varepsilon_d} N_c}{2} \left(\sqrt{1 + 4 g e^{\varepsilon_d} N_c^{-1} N_d} - 1 \right). \quad (4-106)$$

Минус перед корнем отброшен, так как подкоренное выражение больше единицы, а $n_0 > 0$. Это соотношение можно свести к следующему выражению:

$$n_0 = \frac{2 N_d}{\sqrt{1 + 4 g e^{\varepsilon_d} N_c^{-1} N_d} + 1}. \quad (4-107)$$

Положение уровня Ферми определится из соотношения (4-103)

$$e^{\eta} = \frac{n_0}{N_c} = \frac{2 N_d / N_c}{\sqrt{1 + 4 g e^{\varepsilon_d} N_c^{-1} N_d} + 1}. \quad (4-108)$$

Прологарифмировав равенство (4-108), получим:

$$\eta = -\ln \left[\frac{N_c}{2 N_d} + \sqrt{\left(\frac{N_c}{2 N_d} \right)^2 + g e^{\varepsilon_d} N_c N_d^{-1}} \right] \quad (4-109)$$

или

$$F = E_c - kT \ln \left[\frac{N_c}{2 N_d} + \sqrt{\left(\frac{N_c}{2 N_d} \right)^2 + g e^{\varepsilon_d} N_c N_d^{-1}} \right]. \quad (4-110)$$

Проведем анализ найденного выражения для уровня Ферми. В области низких температур, где справедливо условие $e^{-(E_c - E_d)/kT} \gg \frac{N_c}{2 N_d}$, формула (4-110) упрощается и положение уровня Ферми будет определяться соотношением вида

$$F = E_c - \frac{E_c - E_d}{2} - \frac{kT}{2} \ln \frac{g N_c}{N_d} = \frac{E_c + E_d}{2} - \frac{kT}{2} \ln \frac{g N_c}{N_d}, \quad (4-111)$$

а энергия активации согласно формуле (4-72) будет:

$$\omega = \frac{1}{2} (E_c - E_d),$$

т. е. равна половине энергии ионизации донорной примеси.

Как следует из соотношения (4-111), в донорном полупроводнике при температуре абсолютного нуля уровень Ферми располагается посередине между дном зоны проводимости и уровнем донорной примеси. В области достаточно низких температур (несколько градусов по шкале Кельвина), когда $gN_c < N_d$, уровень Ферми вначале повышается до некоторого максимального значения, а затем начинает снижаться и при $gN_c = N_d$ снова имеем $F = 1/2(E_c + E_d)$, как и для случая $T = 0$. Дальнейшее повышение температуры сопровождается ростом N_c и в области температуры, когда $gN_c > N_d$, уровень Ферми продолжает снижаться. Такому перемещению уровня Ферми соответствует экспоненциальная температурная зависимость концентрации электронов

$$n_0 = N_c e^{\eta} = N_c e^{\left(-\frac{E_c - E_d}{2kT} - \frac{1}{2} \ln \frac{gN_c}{N_d}\right)} = (g^{-1} N_c N_d)^{1/2} e^{-\frac{E_c - E_d}{2kT}}. \quad (4-112)$$

Эта область изменения уровня Ферми с температурой, которая описывается формулой (4-111), является областью слабой ионизации примеси (или областью вымороза). Она обозначена цифрой 1 на рис. 4-14, на котором проиллюстрировано изменение уровня Ферми (а) и концентрации электронов (б) в зависимости от температуры для донорного полупроводника.

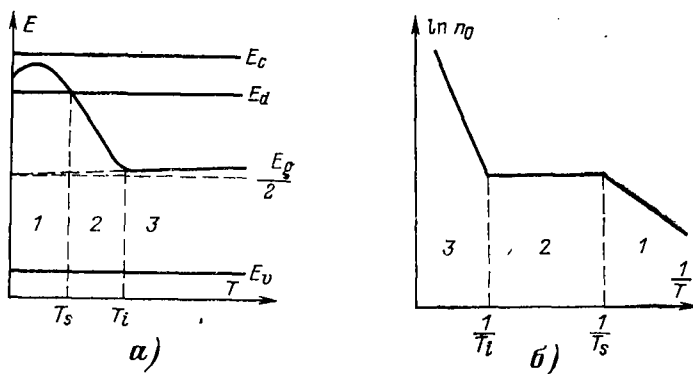


Рис. 4-14. Изменение положения уровня Ферми (а) и концентрации электронов (б) с температурой для донорного полупроводника.

При дальнейшем повышении температуры концентрация электронов в зоне проводимости становится сравнимой с концентрацией примеси и выражения (4-111) и (4-112) в этом случае неприменимы. Однако теперь можно рассматривать другой крайний случай, когда температура достаточно высока и выполняется неравенство $e^{-(E_c - E_d)/kT} \ll \frac{N_c}{2N_d}$.

При этом формула (4-110) аппроксимируется выражением

$$F = E_c - kT \ln \frac{N_c}{N_d}, \quad (4-113)$$

которому соответствует:

$$n_0 = N_d. \quad (4-114)$$

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. 4-14 отмечена цифрой 2.

Условие полной ионизации донорной примеси, когда $n_0 = N_d$, соответствует положению уровня Ферми на несколько kT ниже уровня примеси E_d . Это значит, что при повышении температуры уровень Ферми, понижаясь, пересекает уровень E_d и уходит вниз. Температура, при которой $F = E_d$, носит название температуры насыщения T_s , ее можно определить из условия

$$F = E_c - kT_s \ln \frac{N_c}{N_d} = E_d. \quad (4-115)$$

Отсюда

$$T_s = \frac{E_c - E_d}{k \ln \frac{N_c}{N_d}}. \quad (4-116)$$

Как следует из (4-116), температура насыщения тем ниже, чем меньше энергия ионизации ($E_c - E_d$) и концентрация донорной примеси N_d и чем больше эффективная масса электронов, определяющая величину N_c . При малых значениях ($E_c - E_d$) насыщение наступает при очень низких температурах. Например, в электронном германии, легированном сурьмой в количестве $N_d = 10^{16} \text{ см}^{-3}$, для которой энергия ионизации равна 0,0096 эВ, насыщение наступает уже при 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 (4-88), (4-115) и (4-89). На рис. 4-14 область 3 соответствует области собственной проводимости. В этом случае $F = F_i$, и из равенств (4-88) и (4-113) следует, что

$$F_i = E_c - \frac{1}{2} E_g - \frac{kT_i}{2} \ln \frac{N_c}{N_v} = E_c - kT_i \ln \frac{N_c}{N_d}. \quad (4-117)$$

Отсюда получаем:

$$T_i = \frac{E_g}{k \ln \left(\frac{N_c N_v}{N_d^2} \right)}. \quad (4-118)$$

Анализ этого выражения показывает, что температура T_i , при которой наступает собственная проводимость у донорного полупроводника, тем ниже, чем меньше ширина запрещенной зоны и концентрация примеси и чем больше значение эффективных масс носителей заряда.

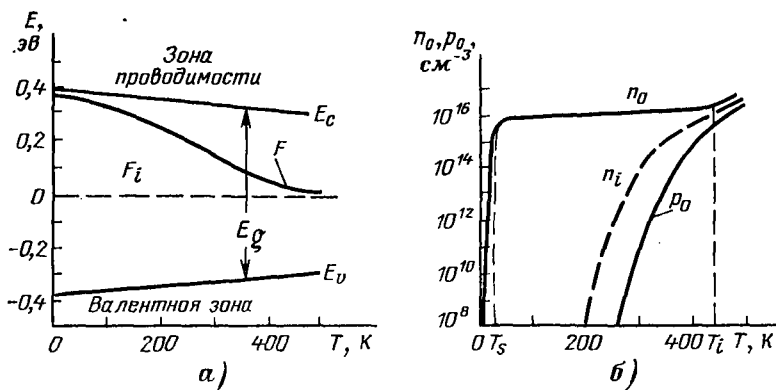


Рис. 4-15. Температурная зависимость уровня Ферми (а) и концентрации (б) носителей заряда для германия, легированного сурьмой.

Таким образом, используя описанные приближения, можно проследить изменение концентрации электронов и положения уровня Ферми в запрещенной зоне электронного полупроводника во всей области изменения температуры.

В качестве примера на рис. 4-15, а приведены температурные зависимости уровня Ферми и на рис. 4-15, б — концентрация равновесных электронов n_0 и дырок p_0 для германия, легированного сурьмой в количестве $N_d \approx 10^{16} \text{ см}^{-3}$. Кроме того, на этих кривых пунктиром показан ход F_i и n_i в собственном 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 опускается и приближается к уровню примеси E_d . При температуре насыщения T_s на донорной примеси электроны находятся в количестве, равном:

$$n_d = \frac{N_d}{\frac{1}{2} e^{(E_d - F)/kT} + 1} = \frac{2}{3} N_d,$$

а в зоне проводимости соответственно $1/3 N_d$ электронов. С дальнейшим ростом температуры уровень Ферми продолжает опускаться и наступает область истощения; вся примесь ионизована и концентрация электронов проводимости остается постоянной и равной $n_0 = N_d$. В этой температурной области имеет место уже ионизация атомов основного вещества и появляются неосновные носители заряда — дырки. Их концентрация резко возрастает с ростом температуры согласно соотношению

$$p_0 = \frac{n_i^2}{n_0} = \frac{N_c N_v}{N_d} e^{-E_g/kT}. \quad (4-119)$$

Когда уровень Ферми достигает середины запрещенной зоны, то $n_0 = p_0 = n_i$ и полупроводник от примесного переходит к собственному. При дальнейшем повышении температуры уровень Ферми приближается к той зоне, которая имеет меньшую эффективную плотность состояний.

2. Акцепторный полупроводник ($N_d=0$)

Решение задачи для полупроводника, легированного акцепторной примесью, полностью аналогично случаю донорного полупроводника. Условие электронейтральности теперь будет иметь вид:

$$p_0 = n_a = N_a - p_a. \quad (4-120)$$

Отсюда для невырожденного полупроводника с учетом формул (4-67) и (4-54) легко получить соотношение для

$$p_0 = \frac{N_a}{ge^{-\varepsilon_a} \eta + 1} = \frac{N_a N_v}{N_v + ge^{\varepsilon_i - \varepsilon_a} p_0}. \quad (4-121)$$

После несложных алгебраических вычислений находим:

$$p_0 = \frac{2N_a}{\sqrt{1 + 4ge^{\varepsilon_i - \varepsilon_a} N_a N_v^{-1}} + 1}. \quad (4-122)$$

Положение уровня Ферми будет определяться равенством

$$e^{\eta + \varepsilon_i} = \frac{N_v}{p_0} = \frac{N_v}{2N_a} \left[\sqrt{1 + 4ge^{\varepsilon_i - \varepsilon_a} N_a N_v^{-1}} + 1 \right]. \quad (4-123)$$

После логарифмирования получим выражение для уровня Ферми:

$$F = E_c - E_g + kT \ln \left[\frac{N_v}{2N_a} + \sqrt{\left(\frac{N_v}{2N_a} \right)^2 + g N_v N_a^{-1}} e^{(E_a - E_v)/kT} \right]. \quad (4-124)$$

В области низких температур, когда выполняется условие $e^{(E_a - E_v)/kT} \gg \frac{N_v}{2N_a}$, имеем:

$$F = E_c - E_g + \frac{1}{2}(E_a - E_v) + \frac{kT}{2} \ln \frac{gN_v}{N_a} = \frac{E_v + E_a}{2} + \frac{kT}{2} \ln \frac{gN_v}{N_a}. \quad (4-125)$$

Таким образом, в дырочном полупроводнике при температуре абсолютного нуля уровень Ферми лежит посередине между потолком валентной зоны и уровнем акцепторной примеси, т. е. $F = \frac{1}{2}(E_v + E_a)$. С увеличением температуры уровень Ферми вначале опускается, когда $N_a > gN_v$, а затем поднимается до уровня E_a . Концентрация свободных дырок в этой области определяется соотношением

$$p_0 = N_v e^{-\eta - \varepsilon_i} = (g^{-1} N_a N_v)^{1/2} e^{(E_v - E_a)/2kT}. \quad (4-126)$$

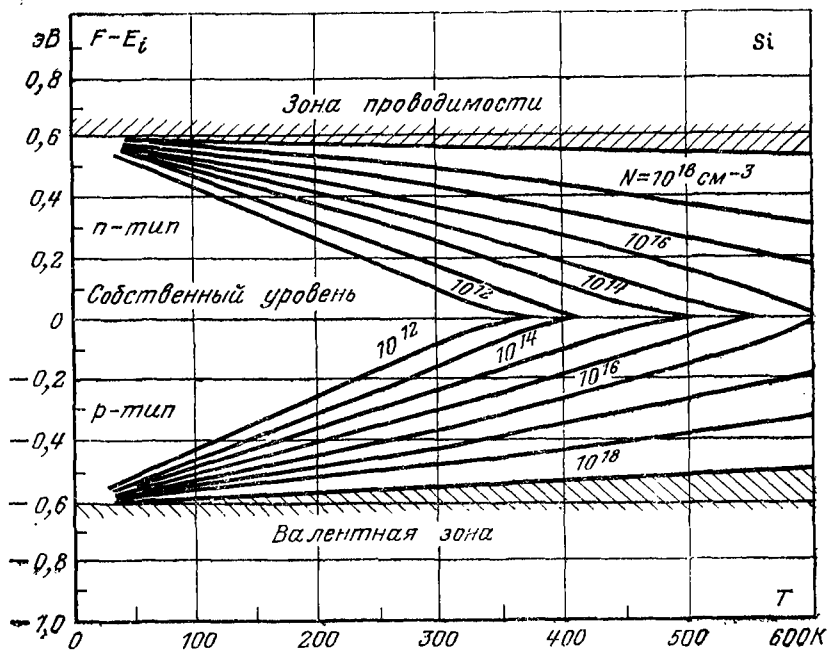


Рис. 4-16. Зависимость уровня Ферми в кремнии от температуры и концентрации примесей.

При высоких температурах, когда $e^{(E_a - E_v)/kT} \ll \frac{N_v}{2N_a}$, из выражения (4-124) получаем:

$$F = E_c - E_g + kT \ln \frac{N_v}{N_a}, \quad (4-127)$$

а концентрация свободных дырок

$$p_0 = N_a e^{-\ln(N_v/N_a)} = N_a. \quad (4-128)$$

В акцепторном полупроводнике, как и в случае донорной примеси, при высоких температурах наступает область истощения, характеризующаяся полной ионизацией атомов акцепторной примеси. С дальнейшим ростом температуры уровень Ферми поднимается к середине запрещенной зоны и полупроводник ведет себя как собственный.

Уровень Ферми для кремния в зависимости от концентрации примесей и температуры показан на рис. 4-16. Здесь же приведена и зависимость ширины запрещенной зоны от температуры.

4-8. ЗАВИСИМОСТЬ ПОЛОЖЕНИЯ УРОВНЯ ФЕРМИ ОТ ТЕМПЕРАТУРЫ ДЛЯ НЕВЫРОЖДЕННОГО ПОЛУПРОВОДНИКА С ЧАСТИЧНО КОМПЕНСИРОВАННОЙ ПРИМЕСЬЮ

Рассмотренные случаи чисто донорного и чисто акцепторного полупроводника обычно не реализуются на практике. Как бы тщательно ни проводилась очистка материала, в нем обычно остаются примеси нескольких видов, причем некоторые из примесных центров могут отдавать или присоединять к себе более одного электрона.

Остановимся более подробно на случае, когда наряду с одноэлектронной донорной примесью в полупроводнике имеется некоторое количество акцепторных центров согласно схеме, изображенной на рис. 4-13.

Рассмотрим случай, когда доноры и акцепторы введены в полупроводник в равных количествах, 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 **компенсированным**. В нем имеет место взаимная компенсация примесей, вследствие чего последние не могут являться поставщиками элект-

тронов и дырок. С ростом температуры свободные носители заряда возникают только благодаря ионизации атомов основного вещества и полупроводник ведет себя, по существу, как собственный. Однако в компенсированном полупроводнике из-за наличия примесей периодичность поля решетки нарушена значительно сильнее, чем в истинно собственном. Это будет сказываться на эффектах, связанных с рассеянием носителей заряда, в частности на их подвижности.

Будем теперь считать, что $N_d > N_a$ и во всей рассматриваемой области температур уровень Ферми лежит значительно выше E_a , т. е. лежит выше уровня энергии акцепторной примеси. При этом условии все акцепторные состояния заполнены электронами, перешедшими с уровней донорной примеси. В результате этого общее количество электронов на донорных уровнях и в зоне проводимости составляет $N_d - N_a$. Пока не наступит собственная проводимость, свободные дырки в валентной зоне отсутствуют и полупроводник ведет себя как чисто донорный. Такой же полупроводник называется частично компенсированным. Степень компенсации определяется отношением N_a/N_d . В таком полупроводнике перенос заряда осуществляется только электронами, и постоянная Холла имеет отрицательный знак.

В случае, когда $N_a > N_d$ и уровень Ферми проходит значительно ниже E_d , донорная примесь полностью 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 атомам акцепторной примеси, полупроводник будет вести себя как дырочный, имеющий положительный коэффициент Холла. С наступлением собственной проводимости знак постоянной Холла изменится, так как подвижность электронов больше подвижности дырок.

Найдем зависимость концентрации свободных электронов и уровня Ферми от температуры в полупроводнике с частично компенсированной примесью для случая $N_d > N_a$. Условие электронейтральности при этом имеет вид:

$$n_0 + n_d = N_d - N_a \quad (4-129)$$

или

$$n_0 + N_a = p_d. \quad (4-130)$$

Будем считать, что полупроводник не вырожден. Используя соотношение $n_0 = N_c e^n$, получаем из (4-130) и (4-52):

$$n_0 + N_a = \frac{N_d N_c}{N_c + g e^{\epsilon_d} n_0}. \quad (4-131)$$

Решив это уравнение, найдем:

$$n_0 = \frac{2(N_d - N_a)}{1 + \frac{gN_a}{N_c} e^{\varepsilon_d} + \left[\left(1 + \frac{gN_a}{N_c} e^{\varepsilon_d} \right)^2 + \frac{4g(N_d - N_a) e^{\varepsilon_d}}{N_c} \right]^{1/2}} \quad (4-132)$$

Из соотношения $e^{\eta} = n_0/N_c$ определим энергию Ферми:

$$F = E_c + kT \times \ln \frac{2(N_d - N_a) N_c^{-1}}{1 + \frac{gN_a}{N_c} e^{\frac{E_c - E_d}{kT}} + \left[\left(1 + \frac{gN_a}{N_c} e^{\frac{E_c - E_d}{kT}} \right)^2 + \frac{4g(N_d - N_a) e^{\frac{E_c - E_d}{kT}}}{N_c} \right]^{1/2}} \quad (4-133)$$

Нетрудно видеть, что при $N_a = 0$ найденные выражения совпадают с формулами (4-107) и (4-110) для невырожденного донорного полупроводника.

В области высоких температур, когда донорные атомы полностью ионизированы, справедливы неравенства

$$\frac{gN_a}{N_c} e^{\varepsilon_d} \ll 1 \quad \text{и} \quad \frac{4g(N_d - N_a) e^{\varepsilon_d}}{N_c} \ll 1. \quad (4-134)$$

Учет этих неравенств дает для n_0 и F следующие соотношения:

$$n_0 = N_d - N_a; \quad (4-135)$$

$$F = E_c + kT \ln \frac{N_d - N_a}{N_c}. \quad (4-136)$$

Как и следовало ожидать, концентрация свободных электронов постоянна и определяется числом нескомпенсированных донорных атомов. Уровень Ферми с ростом температуры понижается и при температурах, соответствующих началу собственной проводимости, приближается к середине запрещенной зоны.

В области низких температур вид выражений для n_0 и F зависит от соотношения концентраций доноров и акцепторов. Если степень компенсации мала, т. е. $N_a \ll N_d$, то можно найти область температур, для которой выполняется соотношение $N_a \ll n_0 \ll N_d$, что эквивалентно условию

$$\frac{gN_a}{N_c} e^{\varepsilon_d} \ll 1 \quad \text{и} \quad \frac{4g(N_d - N_a) e^{\varepsilon_d}}{N_c} \gg 1. \quad (4-137)$$

В этой области температурные зависимости концентрации и уровня Ферми на основании соотношений (4-132) и (4-133) приближенно описываются формулами

$$n_0 = (g^{-1} N_c N_d)^{1/2} e^{-(E_c - E_d)/2kT}; \quad (4-138)$$

$$F = \frac{E_c + E_d}{2} + \frac{kT}{2} \ln \frac{N_d}{gN_c}, \quad (4-139)$$

совпадающими с выражениями (4-112) и (4-111) для донорного полупроводника. Поютно, что при $N_d \gg N_a$ концентрация свободных электронов n_0 не зависит от количества акцепторных атомов, как и следует из равенства (4-138). При $T = 0$ уровень Ферми лежит в середине между E_c и E_d . С ростом температуры он понижается (кривая 1 на рис. 4-17, а).

Однако при более низких температурах, когда $n_0 \ll N_a$, влиянием акцепторов уже нельзя пренебречь. Положение уровня Ферми в этой области существенно отличается от того, что имеет место в чисто донорном полупроводнике [соотношение (4-111)].

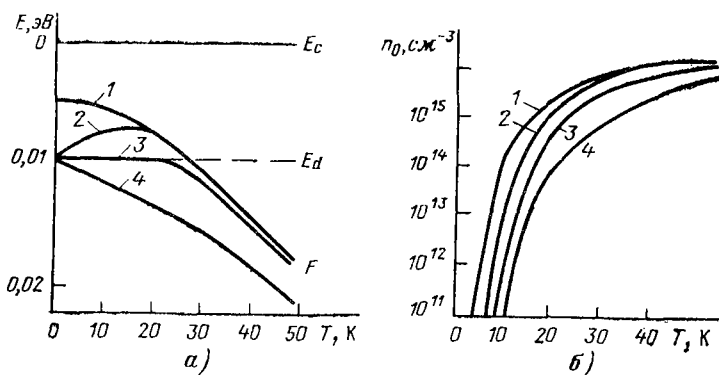


Рис. 4-17. Зависимость положения уровня Ферми (а) и концентрации электронов (б) от температуры при различной степени компенсации.

В случае сильной компенсации справедливо неравенство $n_0 \ll N_a < N_d$, поэтому $\frac{gN_a}{N_c} e^{e_d} \gg 1$ и $\frac{gN_a}{N_c} e^{e_d} > \frac{4g(N_d - N_a) e^{e_d}}{N_c}$ и из выражений (4-132) и (4-133) следует:

$$n_0 \approx N_c \frac{N_d - N_a}{gN_a} e^{-(E_c - E_d)/kT}; \quad (4-140)$$

$$F = E_d + kT \ln \frac{N_d - N_a}{gN_a} = E_d + kT \ln \frac{N_d - N_a}{2N_a}, \quad (4-141)$$

так как $g = 2$.

В этом случае энергия активации равна энергии ионизации донорной примеси $w = E_c - E_d$, т. е. в два раза больше, чем у чисто донорного полупроводника.

На основании соотношения (4-141) на рис. 4-17, а построена зависимость $F(T)$, а на рис. 4-17, б изображена зависимость $n_0(T)$ для различной степени компенсации. Видно, что при повышении температуры, начиная от абсолютного нуля, если $\frac{N_d - N_a}{2N_a} > 1$, уровень Ферми вначале повышается от уровня донорной примеси, а

затем понижается (кривая 2). В противоположном случае, когда $\frac{N_d - N_a}{2N_a} < 1$, с ростом температуры уровень Ферми снижается от донорного уровня (кривая 4), а при $\frac{N_d - N_a}{2N_a} = 1$ он остается неизменным, пока выполняется условие $n_0 \ll N_a$, а затем понижается (кривая 3). Из сравнения кривых, представленных на рис. 4-17, б, следует, что температура, при которой наступает насыщение зависимости $n_0(T)$, растет с увеличением степени компенсации донорной примеси.

4.9. ПРИМЕСНЫЕ ЗОНЫ

Отметим еще одну особенность примесных полупроводников. У донорного полупроводника при температурах, недостаточных для ионизации основного вещества, наличие электронов в зоне проводимости определяется процессом теплового возбуждения атомов донорной примеси. При этом количество электронов проводимости равно количеству образовавшихся положительных ионов примеси, т. е.

$$n_0 = p_d = \frac{N_d}{g e^{(F - E_d)/kT} + 1}. \quad (4-142)$$

В таком полупроводнике при температуре, когда kT меньше энергии ионизации донора или акцептора, удельное сопротивление $\rho = (en_0\mu)^{-1}$ и коэффициент Холла ($R = -1/en_0$) примесного полупроводника согласно формуле (4-142) должны возрастать экспоненциально при понижении температуры. Однако, как показывают исследования, у полупроводников с концентрацией примеси свыше 10^{16} см^{-3} коэффициент Холла при низких температурах проходит через максимум. Положение этого максимума, как видно из рис. 4-18, зависит от концентрации примеси. Температурная зависимость удельного сопротивления этих образцов приведена на рис. 4-19. У них при температуре, когда имеет место максимум коэффициента Холла, удельное сопротивление достигает насыщения, возрастая при дальнейшем понижении температуры значительно медленнее. У образцов с содержанием примеси свыше 10^{19} см^{-3} удельное сопротивление не зависит от температуры.

Кроме того, у полупроводников при увеличении концентрации примеси наблюдается уменьшение энергии ионизации примеси. При достаточно больших концентрациях атомов примеси энергия ионизации обращается в нуль. В германии и кремнии, легированных атомами пятой группы, это происходит при $3 \cdot 10^{17} \text{ см}^{-3}$ и $3 \cdot 10^{18} \text{ см}^{-3}$ соответственно. В качестве примера эта зависимость представлена на рис. 4-20 для кремния, легированного бором.

Из этих экспериментальных фактов следует, что в полупроводнике с большим содержанием примеси при низких температурах

имеет место явление, получившее название проводимости по примесной зоне. При этом механизм примесной проводимости различен в зависимости от концентрации введенной примеси.

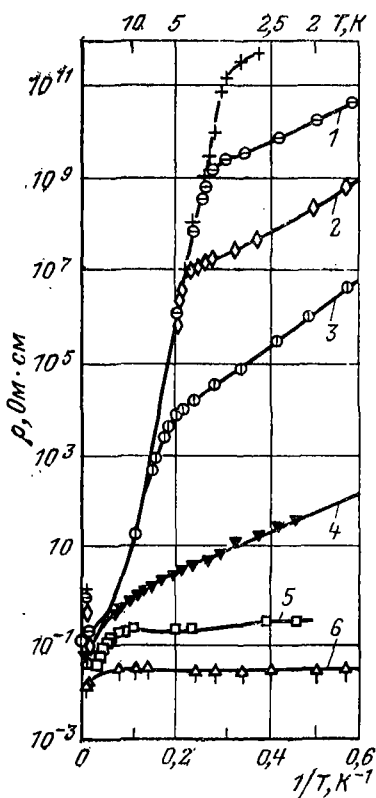


Рис. 4-18. Удельное сопротивление дырочных образцов германия как функция температуры при низких ее значениях.

Концентрация бора: 1 — $6,7 \cdot 10^{17}$ см $^{-3}$; 2 — $1,3 \cdot 10^{18}$ см $^{-3}$; 3 — $2,7 \cdot 10^{18}$ см $^{-3}$; 4 — $6,7 \cdot 10^{18}$ см $^{-3}$; 5 — $1,3 \times 10^{19}$ см $^{-3}$; 6 — $2,7 \cdot 10^{19}$ см $^{-3}$.

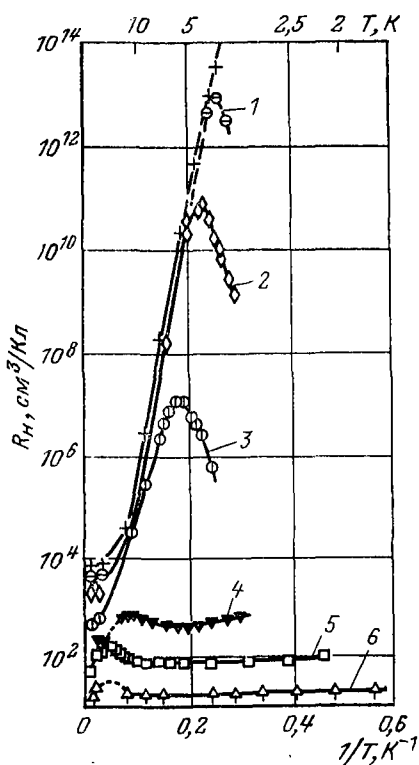


Рис. 4-19. Коэффициент Холла дырочных образцов германия как функция температуры при низких ее значениях. Концентрации примеси те же, что и на рис. 4-18.

В реальном донорном полупроводнике всегда имеется некоторое количество акцепторных центров N_a , поэтому часть электронов с уровней донорной примеси N_d перейдем на уровни акцепторов. В результате такой компенсации даже при очень низкой температуре доноры в количестве N_a оказываются ионизированными. Поэтому электрон, локализованный на одном из доноров, может «перескочить» на ближайший ионизированный донор, ионизировав при этом первый и нейтрализовав последний. Такой перескок элек-

трона происходит случайным образом. При наличии внешнего электрического поля прыжки электронов приобретут направленный характер, в результате чего возникнет поток заряда. Так как доноры расположены на больших расстояниях друг от друга, то подвижность носителей заряда, перемещающихся за счет перескокового механизма, значительно меньше подвижности электронов в зоне проводимости. В области же высоких температур будет преобладать обычная проводимость по зонам энергии. Поэтому для всей области изменения температуры проводимость σ и коэффициент Холла R такого образца будут определяться выражениями

$$\sigma = en_1\mu_1 + en_2\mu_2; \quad (4-143)$$

$$R = \frac{r_1 n_1 \mu_1^2 + r_2 n_2 \mu_2^2}{e (n_1 \mu_1 + n_2 \mu_2)^2} = \frac{n_1 \mu_1 \mu_{1H} + n_2 \mu_2 \mu_{2H}}{e (n_1 \mu_1 + n_2 \mu_2)^2}, \quad (4-144)$$

где согласно соотношению (6-115) $r_1 = \frac{\mu_{1H}}{\mu_1}$; $r_2 = \frac{\mu_{2H}}{\mu_2}$, а n_1 , μ_1 , μ_{1H} и n_2 , μ_2 , μ_{2H} — концентрации, дрейфовые и Холловы подвижности электронов в зоне проводимости и в примесной зоне соответственно.

Водородоподобная модель примесных атомов не учитывает взаимодействия их между собой. Однако при достаточно высоких концентрациях атомов примеси учет их взаимодействия становится необходимым, поскольку боровский радиус для электрона, связанного на водородоподобном примесном центре, согласно выражению (2-175) пропорционален диэлектрической проницаемости и обратно пропорционален эффективной массе электрона.

В силу этого эффективный радиус орбиты для электрона примеси в германии примерно в 60 раз больше, чем радиус Бора у электрона в атоме водорода. Поэтому

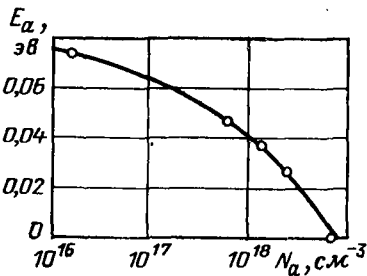


Рис. 4-20. Зависимость энергии ионизации атомов бора от его концентрации в кремнии.

уже при концентрации примеси примерно 10^{16} см^{-3} волновые функции электронов, связанных на нейтральной донорной примеси, значительно перекрываются. В этом случае примесный уровень, расплываясь, превращается в примесную зону, расположенную вблизи дна зоны проводимости. Ширина этой энергетической полосы зависит от степени перекрытия волновых функций, а следовательно, зависит от концентрации донорной примеси. Появление примесной зоны будет сопровождаться уменьшением энергии ионизации примесных атомов (рис. 4-20). В этом случае возникает проводимость по примесной зоне, обусловленная электронами донорной примеси, которые ведут себя подобно электронному газу.

При больших концентрациях донорной примеси примесная зона, расширяясь, может перекрываться с зоной проводимости (или с валентной зоной в случае акцепторного полупроводника). Таким образом, в кристалле образуется единая разрешенная зона, что ведет к полному исчезновению энергии ионизации примеси (см. рис. 4-20).

Наличие примесной зоны влияет на плотность состояний в зоне проводимости вблизи уровня E_c (или соответственно вблизи уровня E_v). При концентрации примеси порядка 10^{19} см⁻³ перекрытие зон 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 на рис. 4-21. У такого полупроводника вследствие высокой степени легирования образуется так называемый «хвост» плотности состояний и ширина запрещенной зоны E_g уменьшается до E_g' .

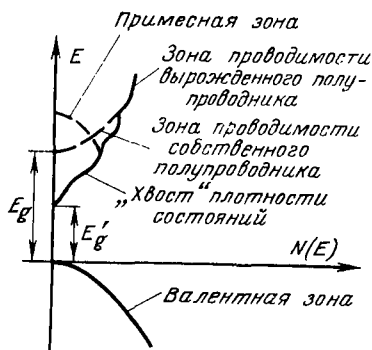


Рис. 4-21. Плотность состояний для вырожденного полупроводника.

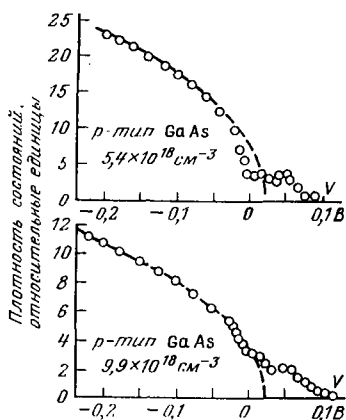


Рис. 4-22. Экспериментальные данные значения плотности состояний в арсениде галлия p -типа; пунктиром показана наилучшая аппроксимация к параболической зоне.

У вырожденного германия со степенью легирования 10^{19} см⁻³ сужение ширины запрещенной зоны составляет несколько сотых долей электронвольта. С увеличением степени легирования эффект «хвоста» плотности состояний возрастает. Это видно из рис. 4-22, на котором даны результаты исследования по определению плотности состояний в арсениде галлия p -типа, полученные при изучении туннелирования в контакте металл—полупроводник.

Несмотря на то что вырожденные полупроводники уже нашли весьма широкое применение в некоторых полупроводниковых приборах (например, туннельные диоды, квантовые генераторы), строгая теория сильно легированных полупроводников является математически сложной задачей.

**МЕХАНИЗМЫ РАССЕЯНИЯ НОСИТЕЛЕЙ ЗАРЯДА
В ПОЛУПРОВОДНИКАХ**

5-1. КИНЕТИЧЕСКОЕ УРАВНЕНИЕ БОЛЬЦМАНА

До сих пор мы рассматривали совокупность носителей заряда в кристалле как систему частиц, находящихся в термодинамическом равновесии. В этом случае движение носителей заряда хаотично, их поведение описывается равновесной функцией распределения, зависящей в общем случае от энергии частиц:

$$f_0 = \frac{1}{e^{(E-F)/kT} + 1}. \quad (5-1)$$

Если наложить на кристалл внешнее постоянное поле, то движение носителей заряда приобретает упорядоченный характер. *Явления, обусловленные движением носителей заряда под действием электрического и магнитного полей, градиента температуры и градиента концентрации носителей заряда, называются явлениями переноса или кинетическими явлениями.* К явлениям переноса относятся электропроводность, электронная теплопроводность, гальваномагнитные, термоэлектрические и термомагнитные эффекты.

При наличии внешних сил носители заряда паходятся в неравновесных условиях. Естественно, что в неравновесном состоянии система носителей заряда описывается другой неравновесной функцией распределения, которая зависит как от энергии частиц, т. е. от \mathbf{k} , так и от их координаты \mathbf{r} и времени t , следовательно, функцией $f(\mathbf{k}, \mathbf{r}, t)$.

Для того чтобы определить концентрацию электронов в неравновесном состоянии, поступим следующим образом. В фазовом пространстве кристалла единичного объема, т. е. в пространстве координат x, y, z и квазиимпульсов p_x, p_y, p_z , выделим элемент объема $d\Gamma$. Это будет пространство, построенное вокруг точки с координатами (x, y, z, p_x, p_y, p_z) :

$$d\Gamma = dx dy dz dp_x dp_y dp_z = d\tau_r d\tau_p = \hbar^3 d\tau_k d\tau_r, \quad (5-2)$$

где

$$d\tau_r = dx dy dz; \quad d\tau_p = dp_x dp_y dp_z = \hbar^3 dk_x dk_y dk_z = \hbar^3 d\tau_k.$$

В элементе объема фазового пространства кристалла единичного объема имеется $d\Gamma/\hbar^3$ фазовых ячеек, в каждой из которых могут быть два электрона с противоположно направленными спинами. Следовательно, в элементе объема $d\Gamma$ фазового пространства содержится $2 \frac{d\Gamma}{\hbar^3}$ квантовых состояний. Если $f(\mathbf{k}, \mathbf{r}, t)$ есть вероятность нахождения электрона в этих состояниях, то число элек-

тронов в элементе фазового пространства в момент времени t будет равно:

$$dn = 2 \frac{d\Gamma}{h^3} f(\mathbf{k}, \mathbf{r}, t) = \frac{1}{4\pi^3} f(\mathbf{k}, \mathbf{r}, t) d\tau_{\mathbf{k}} d\tau_{\mathbf{r}}. \quad (5-3)$$

Проинтегрировав выражение (5-3) по всему фазовому пространству V_{Γ} , получим полное число электронов:

$$n = \frac{1}{4\pi^3} \int_{(V_{\Gamma})} f(\mathbf{k}, \mathbf{r}, t) d\tau_{\mathbf{k}} d\tau_{\mathbf{r}}. \quad (5-4)$$

Таким образом, при рассмотрении кинетических явлений необходимо знание неравновесной функции распределения. Выведем уравнение, которому удовлетворяет неравновесная функция распределения $f(\mathbf{k}, \mathbf{r}, t)$. Для этого рассмотрим движение электронов в обычном пространстве и пространстве волновых векторов.

Сначала рассмотрим систему электронов в обычном, геометрическом пространстве. Пусть внешнее поле приложено таким образом, что электроны движутся со скоростью v_x вдоль положительного направления оси x . Определим изменение числа электронов за время dt внутри элемента объема $d\tau_{\mathbf{r}}$.

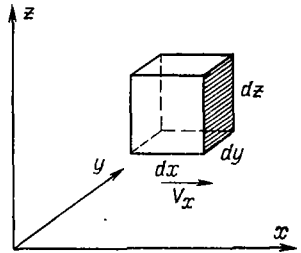


Рис. 5-1. Элементарный объем в геометрическом пространстве.

В элемент объема $d\tau_{\mathbf{r}}$, который изображен на рис. 5-1, через его левую грань за время dt согласно (5-2) войдет $dN_{\text{л}}$ электронов, движущихся со скоростью v_x :

$$dN_{\text{л}} = f(\mathbf{k}, x, y, z, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3} v_x dy dz dt. \quad (5-5)$$

За это же время число электронов $dN_{\text{п}}$, ушедших из объема через правую грань, составит:

$$dN_{\text{п}} = -f(\mathbf{k}, x + dx, y, z, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3} v_x dy dz dt. \quad (5-6)$$

За время dt количество электронов в элементе объема изменится на величину

$$\begin{aligned} dN_{\text{л}} + dN_{\text{п}} &= [f(\mathbf{k}, x, y, z, t) - f(\mathbf{k}, x + dx, y, z, t)] \frac{d\tau_{\mathbf{k}}}{4\pi^3} v_x dy dz dt = \\ &= -v_x \frac{\partial f}{\partial x} \frac{d\tau_{\mathbf{k}}}{4\pi^3} dx dy dz dt = -v_x \frac{\partial f}{\partial x} \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt. \end{aligned} \quad (5-7)$$

В общем случае произвольного направления движения носителей заряда со скоростью \mathbf{v} с составляющими v_x, v_y, v_z изменение

числа электронов с данным \mathbf{k} в элементе объема фазового пространства $d\Gamma$ за время dt составит величину, равную:

$$-\left[v_x \frac{\partial f}{\partial x} + v_y \frac{\partial f}{\partial y} + v_z \frac{\partial f}{\partial z} \right] \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt = -(\mathbf{v} \nabla_{\mathbf{r}} f) \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt. \quad (5-8)$$

Это изменение количества электронов обусловлено процессами переноса, происходящими в результате диффузии носителей заряда при наличии градиента их концентрации или градиента температуры.

Рассмотрим теперь движение электронов в пространстве волновых векторов. Подсчитаем изменение числа электронов за время dt за счет прихода и ухода их через соответствующие грани элемента объема $d\tau_{\mathbf{k}}$. Для этого нужно, так же как и в случае движения электронов в обычном пространстве, подсчитать количество электронов, пришедших и ушедших через все грани элемента объема $d\tau_{\mathbf{k}}$. В результате подсчета получим, что при движении электронов в пространстве волновых векторов их количество в элементе объема фазового пространства $d\Gamma$ за время dt изменится на величину:

$$\begin{aligned} & - \left[\frac{\partial k_x}{\partial t} \frac{\partial f}{\partial k_x} + \frac{\partial k_y}{\partial t} \frac{\partial f}{\partial k_y} + \frac{\partial k_z}{\partial t} \frac{\partial f}{\partial k_z} \right] \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt = \\ & = - \left[\frac{d\mathbf{k}}{dt} \nabla_{\mathbf{k}} f \right] \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt = - \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt. \quad (5-9) \end{aligned}$$

Здесь

$$\mathbf{F}(\mathbf{r}, t) = \frac{d\mathbf{p}}{dt} = \hbar \frac{d\mathbf{k}}{dt} \quad (5-10)$$

есть внешняя сила, действующая на электрон в точке \mathbf{r} в момент времени t .

Соотношение (5-9) выражает изменение количества электронов в результате действия внешней силы, обусловленной электрическим и магнитным полями.

Кроме того, число электронов в элементе объема фазового пространства $d\Gamma$ может меняться в результате взаимодействия электронов с локальными нарушениями периодичности поля кристаллической решетки. Можно считать, что область действия локального возмущения распространяется на расстояние порядка 10^{-7} см, которое охватывает несколько периодов решетки. Так как тепловая скорость электрона в полупроводнике при комнатной температуре порядка 10^7 см/с, то время взаимодействия электрона со структурными нарушениями кристалла составляет 10^{-14} с. Такое кратковременное взаимодействие не может привести к заметному изменению координаты, но зато сильно изменит скорость и квазиимпульс 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, поэтому процессы столкновения называются также процессами рассеяния.

Предположим, что при рассеянии носители заряда переходят из состояний (\mathbf{r}, \mathbf{k}) в состояния $(\mathbf{r}', \mathbf{k}')$. Поскольку при столкновении имеет место кратковременное взаимодействие, которое не вызывает изменения координаты, то вероятность перехода не будет зависеть от \mathbf{r} и \mathbf{r}' . Обозначим вероятность такого перехода в единицу времени как $W(\mathbf{k}, \mathbf{k}')$. В элементе объема $d\tau_{\mathbf{k}}$, число занятых состояний составляет величину $f(\mathbf{k}, \mathbf{r}, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3}$, а число свободных состояний в этом объеме соответственно равно $[1 - f(\mathbf{k}, \mathbf{r}, t)] \frac{d\tau_{\mathbf{k}}}{4\pi^3}$.

Каждое столкновение сопровождается переходом электрона из состояния \mathbf{k} в состояние \mathbf{k}' . Но чтобы такой переход состоялся, состояние \mathbf{k}' должно быть свободным. Если $f(\mathbf{k}', \mathbf{r}, t) \frac{d\tau_{\mathbf{k}'}}{4\pi^3}$ — число занятых состояний в объеме элемента $d\tau_{\mathbf{k}'}$, то $[1 - f(\mathbf{k}', \mathbf{r}, t)] \frac{d\tau_{\mathbf{k}'}}{4\pi^3}$ есть число свободных состояний. Поэтому число электронов, способных за время dt в результате рассеяния перейти из состояния \mathbf{k} в состояние \mathbf{k}' , приводящего к уменьшению количества электронов в элементе объема фазового пространства $d\Gamma$ за время dt , определяется вероятностью перехода $W(\mathbf{k}, \mathbf{k}')$, числом занятых исходных состояний и числом свободных конечных состояний:

$$-W(\mathbf{k}, \mathbf{k}') f(\mathbf{k}, \mathbf{r}, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3} [1 - f(\mathbf{k}', \mathbf{r}, t)] \frac{d\tau_{\mathbf{k}'}}{4\pi^3} d\tau_{\mathbf{r}} dt. \quad (5-11)$$

При столкновении носителей заряда с дефектами решетки возможен и обратный переход электронов из состояния \mathbf{k}' в состояние \mathbf{k} . Пусть вероятность такого перехода в единицу времени, $W(\mathbf{k}', \mathbf{k})$. Тогда число электронов, совершивших за время dt обратный переход, равно:

$$W(\mathbf{k}', \mathbf{k}) f(\mathbf{k}', \mathbf{r}, t) \frac{d\tau_{\mathbf{k}'}}{4\pi^3} [1 - f(\mathbf{k}, \mathbf{r}, t)] \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_{\mathbf{r}} dt. \quad (5-12)$$

Поэтому количество электронов в элементе объема фазового пространства в $d\Gamma$ в результате рассеяния, сопровождающегося прямыми переходами из \mathbf{k} в \mathbf{k}' и обратными из \mathbf{k}' в \mathbf{k} , изменится. Если ради сокращения записи f представить как функцию волнового вектора, то это изменение количества электронов за время dt составит величину, равную

$$\{W(\mathbf{k}', \mathbf{k}) f(\mathbf{k}') [1 - f(\mathbf{k})] - W(\mathbf{k}, \mathbf{k}') f(\mathbf{k}) [1 - f(\mathbf{k}')] \} \times \\ \times \frac{d\tau_{\mathbf{k}}}{4\pi^3} \frac{d\tau_{\mathbf{k}'}}{4\pi^3} d\tau_{\mathbf{r}} dt. \quad (5-13)$$

Чтобы найти полное число изменений состояний, занятых электронами в элементе объема фазового пространства за время dt в результате рассеяния, необходимо учесть все возможные \mathbf{k}' , т. е. необходимо выражение (5-13) проинтегрировать по объему зоны Бриллюэна V_B :

$$\frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r dt \int_{(V_B)} \{W(\mathbf{k}', \mathbf{k}) f(\mathbf{k}') [1 - f(\mathbf{k})] - W(\mathbf{k}, \mathbf{k}') f(\mathbf{k}) [1 - f(\mathbf{k}')] \} \frac{d\tau_{\mathbf{k}'}}{4\pi^3}. \quad (5-14)$$

Но изменение числа электронов в элементе объема фазового пространства $d\Gamma$, происходящее за время от t до $t + dt$, составляет величину

$$f(\mathbf{k}, \mathbf{r}, t + dt) \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r - f(\mathbf{k}, \mathbf{r}, t) \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r = \frac{\partial f}{\partial t} \frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r dt. \quad (5-15)$$

Это изменение числа носителей заряда обусловлено движением электронов за счет диффузии, действия на них внешней силы и наличия процессов рассеяния. Следовательно, выражение (5-15) равно сумме выражений (5-8), (5-9) и (5-14). После сокращения $\frac{d\tau_{\mathbf{k}}}{4\pi^3} d\tau_r dt$ получим:

$$\frac{\partial f}{\partial t} = -(\mathbf{v}\nabla_{\mathbf{r}}f) - \frac{1}{\hbar} (\mathbf{F}\nabla_{\mathbf{k}}f) + \int_{(V_B)} \{W(\mathbf{k}', \mathbf{k}) f(\mathbf{k}') [1 - f(\mathbf{k})] - W(\mathbf{k}, \mathbf{k}') f(\mathbf{k}) [1 - f(\mathbf{k}')] \} \frac{d\tau_{\mathbf{k}'}}{4\pi^3}. \quad (5-16)$$

Согласно принципу микроскопической обратимости вероятности прямых и обратных переходов равны, т. е.

$$W(\mathbf{k}, \mathbf{k}') = W(\mathbf{k}', \mathbf{k}). \quad (5-17)$$

С учетом (5-17) соотношение (5-16) упрощается и принимает вид:

$$\frac{\partial f}{\partial t} = -(\mathbf{v}\nabla_{\mathbf{r}}f) - \frac{1}{\hbar} (\mathbf{F}\nabla_{\mathbf{k}}f) + \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3}. \quad (5-18)$$

Уравнение (5-18) называется кинетическим уравнением Больцмана. Это уравнение является интегро-дифференциальным, так как под интегралом стоит некая функция.

Легко видеть физический смысл уравнения (5-18). Левая часть его отражает изменение функции распределения во времени. Первый и второй члены правой части уравнения показывают изменения функции распределения под действием внешних полей, приводящих к движению электронов в обычном пространстве и в пространстве волновых векторов. Изменение функции распределения

электронов по состояниям в результате их рассеяния (столкновения) учтено в уравнении последним членом. Таким образом, на посетители заряда с одной стороны действуют внешние силы, стремящиеся упорядочить их движение, а с другой — столкновения, стремящиеся разупорядочить их движение. Поэтому уравнение (5-18) можно записать в виде

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t} \right)_{\text{пол}} + \left(\frac{\partial f}{\partial t} \right)_{\text{ст}} \quad (5-19)$$

Здесь член

$$\left(\frac{\partial f}{\partial t} \right)_{\text{пол}} = -(\mathbf{v} \nabla_{\mathbf{r}} f) - \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) \quad (5-20)$$

носит название полевого члена уравнения Больцмана, а величина

$$\left(\frac{\partial f}{\partial t} \right)_{\text{ст}} = \int_{(V_{\text{Б}})} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3} \quad (5-21)$$

называется интегралом столкновений (или соударений).

В случае стационарных процессов, когда величины, описывающие явления переноса, не зависят от времени, функция распределения удовлетворяет равенству

$$\frac{\partial f}{\partial t} = 0 \quad (5-22)$$

и кинетическое уравнение Больцмана (5-19) для стационарного состояния имеет вид:

$$\left(\frac{\partial f}{\partial t} \right)_{\text{пол}} = - \left(\frac{\partial f}{\partial t} \right)_{\text{ст}} \quad (5-23)$$

или

$$(\mathbf{v} \nabla_{\mathbf{r}} f) + \frac{1}{\hbar} (\mathbf{F} \nabla_{\mathbf{k}} f) = \int_{(V_{\text{Б}})} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3} \quad (5-24)$$

Из выражений (5-23) и (5-24) следует, что в стационарном состоянии изменения функции распределения, вызванные движением носителей заряда и действием на них внешней силы, компенсируются соударениями носителей заряда с локальными нарушениями периодичности поля решетки.

Необходимо указать, что кинетическое уравнение Больцмана по существу является квазиклассическим, поскольку в функции распределения одновременно задаются координата и волновой вектор. В силу этого оно применимо к процессам, для которых имеет смысл понятие длины свободного пробега. Это означает, что средняя длина свободного пробега носителей заряда должна быть больше длины

волны де Бройля. Кроме того, это уравнение можно использовать для описания явления переноса в магнитном поле, если это поле не меняет спектр энергий носителей заряда в кристалле, т. е. не становятся существенными квантовые эффекты. Это условие накладывает ограничение на величину магнитного поля.

5-2. РАВНОВЕСНОЕ СОСТОЯНИЕ

При термодинамическом равновесии, при котором состояние электронов описывается равновесной функцией распределения (4-28), движение электронов в кристалле обуславливается тепловой энергией, поэтому для равновесного случая левая часть уравнения (5-24) равна нулю. Это значит, что

$$\int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f_0(\mathbf{k}') - f_0(\mathbf{k})] d\tau_{\mathbf{k}'} = 0. \quad (5-25)$$

Из этого соотношения следует, что в системе, находящейся в равновесном состоянии, имеет место равенство потоков носителей заряда для прямого и обратного процессов, что согласуется с принципом детального равновесия.

На основании (5-25) получаем:

$$f_0(\mathbf{k}') = f_0(\mathbf{k}). \quad (5-26)$$

Подставляя в (5-26) выражение для равновесной функции распределения (4-28), будем иметь:

$$E_1 - F_1 = E_2 - F_2, \quad (5-27)$$

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

$$F_1 = F_2. \quad (5-28)$$

Это выражение показывает, что в равновесном состоянии во всех частях системы, между которыми может происходить переход электронов, положение уровня Ферми одинаково.

5-3. ВРЕМЯ РЕЛАКСАЦИИ

Решение кинетического уравнения (5-18) и даже для стационарного случая (5-24) является весьма сложной задачей. Но она значительно упрощается для тех случаев, когда можно ввести время релаксации. Как оказывается, описание процессов рассеяния при помощи времени релаксации возможно, если столкновения частиц упругие, т. е. такие, при которых энергия носителя заряда мало изменяется, и если процессы рассеяния приводят к случайному рас-

пределимо носителей заряда по скоростям, т. е. имеет место равновероятное рассеяние носителей заряда по всем направлениям.

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 процессов соударений система придет в равновесное состояние, описываемое равновесной функцией распределения f_0 . Это значит, что после выключения внешнего поля функция распределения изменяется благодаря наличию соударений электронов с дефектами решетки:

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t} \right)_{\text{ст}}. \quad (5-29)$$

В том случае, когда отклонение распределения носителей заряда от равновесного состояния невелико, можно положить, что в отсутствие внешних полей скорость изменения функции распределения вследствие соударений пропорциональна величине отклонения функции от равновесия, т. е. пропорциональна $f - f_0$:

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t} \right)_{\text{ст}} = -\frac{f - f_0}{\tau(\mathbf{k})}, \quad (5-30)$$

где $1/\tau(\mathbf{k})$ — коэффициент пропорциональности, зависящий от \mathbf{k} . Решая уравнения (5-30), получаем:

$$f - f_0 = (f - f_0)_{t=0} e^{-t/\tau}. \quad (5-31)$$

Из (5-31) 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 примерно равно времени свободного пробега электрона. С учетом этого можно считать, что длина свободного пробега l определяется скоростью движения электрона v и временем релаксации $\tau(\mathbf{k})$.

Используя выражение (5-30), интеграл столкновений (5-21) можно представить в виде

$$\left(\frac{\partial f}{\partial t} \right)_{\text{ст}} = \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f(\mathbf{k}') - f(\mathbf{k})] \frac{d\tau_{\mathbf{k}'}}{4\pi^3} = -\frac{f(\mathbf{k}) - f_0(\mathbf{k})}{\tau(\mathbf{k})}. \quad (5-32)$$

Поскольку внешние силы в соответствии с равенством (5-30) не приводят к большим отклонениям распределения носителей заряда от равновесного, то неравновесную функцию распределения можно представить как

$$f(\mathbf{k}) = f_0 + f_1(\mathbf{k}), \quad (5-33)$$

где $f_1(\mathbf{k})$ — поправка к равновесной функции распределения, появившаяся в результате внешнего воздействия на систему частиц. Запишем ее в виде

$$f_1(\mathbf{k}) = -\frac{\partial f_0}{\partial E} \mathbf{k}\chi(E), \quad (5-34)$$

где $\chi(E)$ — некоторый вектор, зависящий лишь от энергии носителя заряда.

Будем считать, что время релаксации не зависит от внешних полей. Это значит, что $\tau(\mathbf{k})$ определяет и процессы рассеяния при наличии внешних полей. В силу этого для уравнения (5-24) можно использовать интеграл столкновения (5-32) и кинетическое уравнение для стационарного случая примет вид:

$$(\mathbf{v}\nabla_{\mathbf{r}}f) + \frac{1}{\hbar} (\mathbf{F}\nabla_{\mathbf{k}}f) = -\frac{f-f_0}{\tau(\mathbf{k})} = -\frac{f_1(\mathbf{k})}{\tau(\mathbf{k})}. \quad (5-35)$$

Запишем интеграл столкновения (5-32), используя соотношение (5-33):

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') [f_1(\mathbf{k}') - f_1(\mathbf{k})] d\tau_{\mathbf{k}'} = -\frac{f_1(\mathbf{k})}{\tau(\mathbf{k})}. \quad (5-36)$$

Из равенства (5-36) можно определить время релаксации. Если учесть соотношение (5-34), то $\tau(\mathbf{k})$ будет определяться равенством следующего вида:

$$\begin{aligned} \frac{1}{\tau(\mathbf{k})} &= -\frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \frac{f_1(\mathbf{k}') - f_1(\mathbf{k})}{f_1(\mathbf{k})} d\tau_{\mathbf{k}'} = \\ &= \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left[1 - \frac{\frac{\partial f_0}{\partial E'} \mathbf{k}'\chi(E')}{\frac{\partial f_0}{\partial E} \mathbf{k}\chi(E)} \right] d\tau_{\mathbf{k}'}. \quad (5-37) \end{aligned}$$

Сударения, как известно, могут быть упругими и неупругими. Сударения носят название упругих, если кинетическая энергия сталкивающихся частиц сохраняется. Неупругими соударениями являются такие, в результате которых энергия частиц меняется и после удара становится больше или меньше, чем до удара.

Будем считать, что соударения электронов со структурными нарушениями являются упругими. В этом случае у электрона изменяется лишь направление волнового вектора и, следовательно, направление скорости движения, а его абсолютное значение остается неизменным. Поэтому при упругом соударении можно считать, что

$$\mathbf{k}' = \mathbf{k}; \quad v' = v \quad \text{и} \quad E' = E, \quad (5-38)$$

если зоны энергии сферические.

В силу этого время релаксации для упругого соударения с учетом (5-38) будет равно:

$$\frac{1}{\tau(\mathbf{k})} = \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left[1 - \frac{\mathbf{k}'\chi(E)}{k\chi(E)} \right] d\tau_{\mathbf{k}'}. \quad (5-39)$$

Пусть векторы χ , \mathbf{k} , \mathbf{k}' направлены так, как это изображено на рис. 5-2. В этом случае проекции векторов \mathbf{k}' и \mathbf{k} на вектор χ соответственно равны:

$$k_\chi = k; \quad k'_\chi = k' \cos \theta = k \cos \theta, \quad (5-40)$$

где θ — угол отклонения электрона от первоначального направления движения в результате столкновения. С учетом (5-40) равенство (5-39) запишется в виде

$$\begin{aligned} \frac{1}{\tau(\mathbf{k})} &= \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left[1 - \frac{k'_\chi}{k_\chi} \right] d\tau_{\mathbf{k}'} = \\ &= \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') (1 - \cos \theta) d\tau_{\mathbf{k}'}. \end{aligned} \quad (5-41)$$

Таким образом, при рассмотрении процессов переноса основной задачей является вычисление неравновесной функции распределения, с помощью которой можно рассчитать все величины, характеризующие кинетические эффекты. Но как следует из равенства (5-35), эта задача решается при условии, если известно время релаксации (5-41), а следовательно, необходимо знание механизма рассеяния носителей заряда.

Рассмотрим теперь рассеяние, приводящее к случайному распределению носителей заряда по скоростям, т. е. к такому, при котором электрон, рассеиваясь, попадает в состоянии \mathbf{k}' , находящееся возле того же минимума энергии, что и \mathbf{k} , но с противоположно направленной скоростью. Такое рассеяние возможно в том случае, если поверхности постоянной энергии являются сферами. Поскольку $W(\mathbf{k}, \mathbf{k}') = W(\mathbf{k}', \mathbf{k})$ — четная функция, а $f_1(\mathbf{k}') = -\frac{\partial f_0}{\partial E}(\mathbf{k}'\chi)$ — нечетная функция относительно инверсии $\mathbf{k} \rightarrow -\mathbf{k}$, то

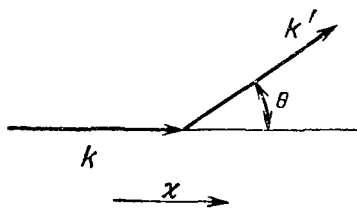


Рис. 5-2. Изменение волнового вектора электрона при рассеянии.

$$\int_{(V_B)} W(\mathbf{k}, \mathbf{k}') f_1(\mathbf{k}') d\tau_{\mathbf{k}'} = - \int_{(V_B)} \frac{\partial f_0}{\partial E} W(\mathbf{k}, \mathbf{k}') (\mathbf{k}'\chi) d\tau_{\mathbf{k}'} = 0, \quad (5-42)$$

так как нечетная функция при интегрировании в симметричных пределах дает нуль. С учетом выражения (5-42) интеграл столкновения (5-36) будет равен:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{ст}} = -\frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') f_1(\mathbf{k}) d\tau_{\mathbf{k}} = -\frac{f_1(\mathbf{k})}{\tau(\mathbf{k})}, \quad (5-43)$$

откуда следует, что время релаксации при рассеянии, приводящем к случайному распределению носителей заряда по скоростям, определяется выражением

$$\frac{1}{\tau(\mathbf{k})} = \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') d\tau_{\mathbf{k}}. \quad (5-44)$$

5-4. ЭФФЕКТИВНОЕ СЕЧЕНИЕ РАССЕЯНИЯ. ТИПЫ ЦЕНТРОВ РАССЕЯНИЯ

Любая неоднородность кристалла, искажающая периодичность поля решетки, является центром рассеяния. Поскольку природа дефектов решетки различна, то и рассеяние носителей заряда на них будет различно. Для количественной оценки процесса рассеяния вводится величина, называемая эффективным сечением и обозначаемая σ .

Предположим, что имеется n свободных электронов, которые со средней тепловой скоростью v_0 движутся в данном направлении. Тогда nv_0 есть плотность потока электронов, т. е. 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 σ , количеством центров рассеяния N и плотностью падающего потока электронов nv_0 , т. е.

$$n_1 = \sigma N n v_0. \quad (5-45)$$

С другой стороны, если W — вероятность рассеяния одной частицы в единицу времени, то количество рассеянных электронов за 1 с

$$n_1 = W n. \quad (5-46)$$

Тогда на основании выражений (5-45) и (5-46) можно записать:

$$\sigma = \frac{n_1}{N} = \frac{W}{n v_0}. \quad (5-47)$$

Таким образом, эффективное сечение рассеяния σ есть отношение числа электронов, удаленных из пучка в результате рассеяния

на одном центре в единицу времени, к плотности падающего пучка частиц. Эффективное сечение рассеяния имеет размерность площади

$$[\sigma] = \frac{[W]}{[Nv_0]} = \frac{T^-}{L^{-3}LT^{-1}} = L^2.$$

Из формулы (5-47) найдем, что вероятность рассеяния

$$W = \sigma N v_0. \quad (5-48)$$

Следовательно, вероятность рассеяния или вероятность столкновения определяется эффективным сечением, количеством центров рассеяния и скоростью движения носителя заряда.

В то же время вероятность столкновения обратно пропорциональна времени свободного пробега:

$$W = \frac{1}{\tau}. \quad (5-49)$$

Поэтому время свободного пробега и длину свободного пробега можно выразить через эффективное сечение. Из равенств (5-48) и (5-49) следует, что

$$\tau = \frac{1}{\sigma N v_0} = \frac{l}{v_0} \quad (5-50)$$

или

$$l = \frac{1}{\sigma N}. \quad (5-51)$$

Величина $l^{-1} = \sigma N$ есть вероятность рассеяния на единичном интервале пути.

Рассмотрим случай, когда имеются различные центры рассеяния. Пусть i -й центр характеризуется эффективным сечением σ_i и их число равно N_i . Механизм рассеяния на этих центрах определяет длину свободного пробега l_i . Если все процессы рассеяния возможны, то согласно теории вероятности полная вероятность рассеяния в единицу времени будет определяться суммой отдельных вероятностей рассеяния:

$$W = \sum_i W_i. \quad (5-52)$$

Так как $\sum_i \sigma_i N_i = \sum_i l_i^{-1}$, то полная длина свободного пробега может быть определена из соотношения

$$l^{-1} = \sum_i l_i^{-1}. \quad (5-53)$$

Согласно выражению (5-53) полная длина свободного пробега всегда меньше самой малой парциальной длины свободного пробега.

В полупроводниках возможны 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, 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 $N_A = 10^{16} \text{ см}^{-3}$, то длина свободного пробега при рассеянии носителей заряда на атомах примеси будет составлять: $l_A = (3 \cdot 10^{-15} \cdot 10^{16})^{-1} \approx 3 \cdot 10^{-2} \text{ см}$.

Для иона примеси будем считать, что его диаметр в 10 раз больше диаметра примесного центра, т. е. $\sigma_I = (5 \cdot 10^{-8} \cdot 10)^2 \approx 3 \cdot 10^{-13} \text{ см}^2$. В случае, если $N_I = 10^{16} \text{ см}^{-3}$, то $l_I = 3 \cdot 10^{-4} \text{ см} = 3 \text{ мкм}$.

Дислокации являются линейными дефектами, простирающимися на большие области кристалла. Предположим, что линейная дислокация имеет длину 1 мм, а ее диаметр измеряется сотней периодов решетки. В этом случае 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 $N_D = 10^8 \text{ см}^{-3}$ длина свободного пробега будет порядка $l_D = 2 \cdot 10^{-2} \text{ см}$.

Границы двойников и кристаллитов, а также дефекты упаковки представляют собой двумерные переупорядоченности. Они имеют место в блочных кристаллах и поликристаллических образцах.

Эффективное сечение рассеяния на тепловых колебаниях решетки определяется площадью сечения области, которую занимает колеблющийся атом за вычетом площади сечения самого атома. Если считать, что амплитуда колебаний порядка $r = 0,05 \text{ нм} = 5 \cdot 10^{-9} \text{ см}$, а диаметр атома $d = 10^{-8} \text{ см}$, то $\sigma_T = (d+r)^2 - d^2 \approx 2rd = 10^{-16} \text{ см}^2$. Это значение много меньше, чем для других центров рассеяния, но число колеблющихся атомов велико: $N_T = 10^{22} \text{ см}^{-3}$, поэтому $l_T = 10^{-6} \text{ см} = 10 \text{ нм}$.

5.5. РАССЕЯНИЕ НА ИОНАХ ПРИМЕСИ

В примесном полупроводнике каждый ион примеси вокруг себя создает электрическое поле. Под воздействием этого поля движущийся носитель заряда отклоняется от своего первоначального направления (рис. 5-3). При этом носитель заряда отклоняется тем сильнее, чем медленнее и ближе к иону примеси он движется.

Задача сводится фактически к нахождению траектории движения заряженной частицы в центральном поле иона. Она может быть решена чисто классическими методами. Для этого будем считать носитель заряда классической частицей, движущейся в кулоновском поле положительного иона примеси с зарядом Ze . В этом случае потенциальную энергию взаимодействия можно представить в виде

$$U(r) = \pm \frac{Ze^2}{4\pi\epsilon_0 r}, \quad (5-54)$$

где плюс соответствует дырке, минус — электрону.

Как показывает расчет¹, в кулоновском поле иона носитель заряда движется по гиперболе. На рис. 5-4 изображены траектории электрона и дырки в поле положительного иона. Оба носителя заряда отклоняются одинаково, хотя в одном случае действует

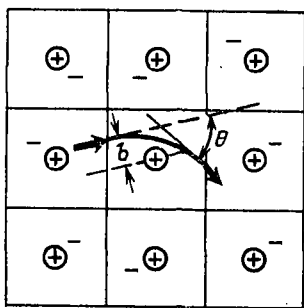


Рис. 5-3. Искривление траектории движения электрона в поле положительного иона примеси.

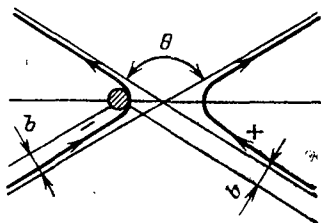


Рис. 5-4. Рассеяние электрона и дырки на положительном ионе примеси.

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 заряда и его движением после рассеяния. При этом угол рассеяния зависит от прицельного расстояния следующим образом:

$$b = \frac{Ze^2}{\epsilon m^* v^2} \operatorname{ctg} \frac{\theta}{2}, \quad (5-55)$$

где m^* и v — эффективная масса и скорость движения носителя заряда.

Процесс рассеяния является случайным процессом, поэтому различные носители заряда будут отклоняться на различные углы

¹ Шпольский Э. В. Атомная физика. М., «Наука», 1974, т. 1, с. 101.

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. 5-5. Здесь через $d\Omega$ обозначен телесный угол, заключенный между этими конусами, он равен:

$$d\Omega = 2\pi \sin \theta d\theta. \quad (5-56)$$

Процессы рассеяния в данном случае характеризуются дифференциальным сечением рассеяния $\sigma(\theta)$. Согласно (5-47) $\sigma(\theta)$ равно отношению числа частиц, отклоненных одним центром в 1 с на угол $d\theta$ в единичный телесный угол, к потоку падающих частиц на 1 см² в 1 с.

Пусть имеется n электронов, движущихся хаотически со скоростью v . Если выбрать перпендикулярно к каждому направлению движения электронов площадку в 1 см², то на нее за 1 с упадет nv электронов. Полное число частиц, отклоненных за 1 с на угол θ в пределах телесного угла $d\Omega$, будет равно потоку частиц, падающих на кольцо (рис. 5-5) площадью

Рис. 5-5. Рассеяние электронов точечным несовершенством.

$dS = 2\pi b |db|$. Поэтому на основании определения и равенства (5-56) можно записать:

$$\sigma(\theta) = \frac{nv2\pi b |db|}{nv d\Omega} = \frac{b |db|}{\sin \theta d\theta}. \quad (5-57)$$

Так как из формулы (5-55) имеем:

$$|db| = \frac{Ze^2}{2\epsilon m^* v^2} \frac{1}{\sin^2 \frac{\theta}{2}} d\theta, \quad (5-58)$$

то, подставляя выражения (5-55) и (5-58) в (5-57), получаем:

$$\begin{aligned} \sigma(\theta) &= \frac{Ze^2}{\epsilon m^* v^2} \operatorname{ctg} \frac{\theta}{2} \frac{Ze^2}{2\epsilon m^* v^2} \frac{1}{\sin^2 \frac{\theta}{2}} d\theta \frac{1}{2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} d\theta} = \\ &= \left(\frac{Ze^2}{2\epsilon m^* v^2} \right)^2 \frac{1}{\sin^4 \frac{\theta}{2}}. \end{aligned} \quad (5-59)$$

Формула (5-59) была получена Резерфордом при изучении рассеяния α -частиц на ядрах элементов.

Для подсчета времени релаксации необходимо выразить вероятность перехода $W(\mathbf{k}, \mathbf{k}')$ через дифференциальное сечение рассеяния $\sigma(\theta)$. Предположим, что в кристалле единичного объема имеется N_I ионов примеси. Тогда число рассеянных носителей заряда всеми центрами в 1 с в телесный угол $d\Omega$ составит:

$$W(\mathbf{k}, \mathbf{k}') n d\Omega = \sigma(\theta) N_I n v d\Omega, \quad (5-60)$$

откуда

$$W(\mathbf{k}, \mathbf{k}') = \sigma(\theta) N_I v. \quad (5-61)$$

Таким образом, *вероятность перехода из состояния \mathbf{k} в состояние \mathbf{k}' при рассеянии электронов на ионах примеси определяется дифференциальным эффективным сечением, количеством центров рассеяния и скоростью движения носителей заряда.*

Соударения носителей заряда с ионами примеси можно считать упругими по двум причинам. Во-первых, масса электрона в 10^4 раз меньше массы иона, являющегося рассеивающим центром. Во-вторых, ионы решетки гораздо менее подвижны относительно своего положения равновесия, чем свободные носители заряда, поэтому при соударении носителя заряда с ионом примеси энергия носителя заряда изменяется очень незначительно.

Если при упругом соударении носители заряда рассеиваются в телесный угол $d\Omega$, то время релаксации, используя (5-41), можно записать в виде

$$\frac{1}{\tau(\mathbf{k})} = \int_{(\theta)} W(\mathbf{k}, \mathbf{k}') (1 - \cos \theta) d\Omega. \quad (5-62)$$

Подставляя формулы (5-56), (5-61) в (5-62), получаем:

$$\frac{1}{\tau(\mathbf{k})} = 2\pi N_I v(\mathbf{k}) \int_{(\theta)} \sigma(\theta) (1 - \cos \theta) \sin \theta d\theta = N_I v(\mathbf{k}) \sigma_c. \quad (5-63)$$

Величина

$$\sigma_c = 2\pi \int_{(\theta)} \sigma(\theta) (1 - \cos \theta) \sin \theta d\theta \quad (5-64)$$

есть усредненное дифференциальное эффективное сечение рассеяния; усреднение по углам θ производится с весовой функцией, $(1 - \cos \theta)$. σ_c называется эффективным сечением проводимости (или подвижности), а также транспортным эффективным сечением.

Время релаксации на основании (5-63) равно:

$$\tau(\mathbf{k}) = \frac{1}{N_I \sigma_c v(\mathbf{k})}, \quad (5-65)$$

а длина свободного пробега

$$l(\mathbf{k}) = \tau(\mathbf{k}) v(\mathbf{k}) = \frac{1}{\sigma_c N_I} \quad (5-66)$$

Из формулы (5-59) следует, что рассеяние носителей заряда на ионах примеси в достаточной степени анизотропно. Рассеяние под малыми углами или при больших значениях прицельного расстояния является преобладающим. Если в формулу (5-64) подставить значение $\sigma(\theta)$ из формулы (5-59), то интеграл на нижнем пределе при $\theta = 0$ будет расходящимся, а сечение рассеяния — бесконечным. Но рассеяние носителей заряда на ионизованной примеси обладает конечным сечением, поэтому при определении времени релаксации интегрирование выражения (5-63) проводится от $\theta_{\text{мин}}$ до π . Определить значение $b_{\text{макс}}$, соответствующее $\theta_{\text{мин}}$, 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 $b_{\text{макс}} = \frac{1}{2} N_I^{-1/3}$. Он соответствует наименьшему значению угла рассеяния $\theta_{\text{мин}}$, который можно определить из формулы (5-55):

$$\text{ctg} \frac{\theta_{\text{мин}}}{2} = \frac{\epsilon m^* v^2}{Ze^2} b_{\text{макс}} = \frac{\epsilon m^* v^2}{2Ze^2 N_I^{1/3}} \quad (5-67)$$

Учитывая, что $1 - \cos \theta = 2 \sin^2 \frac{\theta}{2}$; $\sin \theta = 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}$, а

$$\int_{\theta_{\text{мин}}}^{\pi} \frac{(1 - \cos \theta) \sin \theta d\theta}{\sin^4 \frac{\theta}{2}} = 8 \int_{\theta_{\text{мин}}}^{\pi} \frac{\cos \frac{\theta}{2} d\left(\frac{\theta}{2}\right)}{\sin \frac{\theta}{2}} = -4 \ln \sin^2 \frac{\theta_{\text{мин}}}{2},$$

и воспользовавшись выражениями (5-67) для $\text{ctg} \frac{\theta_{\text{мин}}}{2}$ и (5-59) для $\sigma(0)$, найдем эффективное сечение проводимости:

$$\begin{aligned} \sigma_c &= 2\pi \int_{\theta_{\text{мин}}}^{\pi} \sigma(0) (1 - \cos \theta) \sin \theta d\theta = 8\pi \left(\frac{Ze^2}{2\epsilon m^* v^2} \right)^2 \ln \frac{1}{\sin^2 \frac{\theta_{\text{мин}}}{2}} = \\ &= 2\pi \left(\frac{Ze^2}{\epsilon m^* v^2} \right)^2 \ln \left[1 + \text{ctg}^2 \frac{\theta_{\text{мин}}}{2} \right] = \\ &= 2\pi \left(\frac{Ze^2}{\epsilon m^* v^2} \right)^2 \ln \left[1 + \left(\frac{\epsilon m^* v^2}{2Ze^2 N_I^{1/3}} \right)^2 \right]. \end{aligned} \quad (5-68)$$

Время релаксации при рассеянии носителей заряда на ионизованных атомах примеси на основании (5-65) равно:

$$\tau_I(\mathbf{k}) = \frac{1}{\sigma_c N_I v(\mathbf{k})} = \frac{e^2 m^* v^3(\mathbf{k})}{2\pi Z^2 e^4 N_I \ln \left[1 + \left(\frac{\epsilon m^* v^2}{2Ze^2 N_I^{1/3}} \right)^2 \right]}. \quad (5-69)$$

Это соотношение для τ называется формулой Конуэлл—Вайскопфа.

Если выразить скорость через энергию $v = (2E/m^*)^{1/2}$, то время релаксации при рассеянии носителей заряда на ионах примеси примет вид:

$$\tau_I = \tau_{0I} E^{3/2}, \quad (5-70)$$

где

$$\tau_{0I} = \frac{(2m^*)^{1.2} \epsilon^2}{\pi Z^2 e^4 N_I \ln \left[1 + \left(\frac{\epsilon E}{Ze^2 N_I^{1/3}} \right)^2 \right]}, \quad (5-71)$$

а длина свободного пробега равна:

$$l_I(\mathbf{k}) = \tau_I(\mathbf{k}) v(\mathbf{k}) = \left(\frac{2}{m^*} \right)^{1/2} \tau_{0I} E^2. \quad (5-72)$$

Таким образом, при рассеянии носителей заряда на ионах примеси время релаксации зависит от энергии в степени $3/2$ и пропорционально эффективной массе в степени $1/2$.

5-6. РАССЕЯНИЕ НА АТОМАХ ПРИМЕСИ И ДИСЛОКАЦИЯХ

Рассеяние носителей заряда на нейтральных атомах примеси проявляется гораздо слабее, чем рассеяние на ионах примеси. Однако при низких температурах, когда концентрация ионизованных атомов примеси меньше концентрации нейтральных атомов примеси, этот механизм рассеяния играет существенную роль. В этом случае рассеяние носителей заряда может осуществляться двумя путями: при помощи прямого упругого соударения и обменом падающего электрона с электроном примесного атома. Определение времени релаксации для этих двух механизмов рассеяния достаточно сложно. Однако был проведен элементарный расчет, причем рассеяние носителей заряда на нейтральных атомах примеси рассматривалось как рассеяние медленных электронов на атоме водорода, находящемся в среде с диэлектрической проницаемостью ϵ . При этом получается формула

$$\tau_A = \frac{e^2 m^*}{20\epsilon h^3} \frac{1}{N_A}, \quad (5-73)$$

где N_A — концентрация нейтральных атомов примеси при данной температуре.

Как следует из (5-73), время релаксации при рассеянии носителей заряда на нейтральной примеси не зависит ни от температуры, ни от энергии рассеиваемого носителя заряда, но определяется концентрацией атомов примеси. Поэтому этот механизм рассеяния особенно важен при очень низких температурах, когда примесь не ионизована.

Рассеяние носителей заряда может происходить на нарушениях решетки, вносимых дислокациями. Вокруг дислокации существует область, в которой кристалл деформирован, поэтому в окрестности дислокации зависимость энергии от волнового вектора отличается от закона дисперсии в идеальном кристалле. Это приводит к рассеянию свободных носителей заряда. В полупроводниках, которые кристаллизуются в структуре типа алмаза, дислокации действуют как цепочки акцепторов, которые захватывают свободные электроны. Поэтому они представляют собой положительный цилиндр с отрицательно заряженной линией в середине, т. е. дислокацию можно уподобить линейному отрицательному заряду в виде бесконечно длинного цилиндра радиуса R , вокруг которого имеется положительный пространственный заряд. Электроны, движущиеся перпендикулярно дислокациям, не могут свободно перемещаться в направлении поля, так как должны обойти дислокации, и рассеиваются на них, испытывая отталкивание. Расчет показывает, что при рассеянии носителей заряда на дислокациях время релаксации не зависит от температуры кристалла, а определяется плотностью дислокаций на единицу поверхности N_D и скоростью рассеиваемого электрона v согласно выражению вида

$$\tau_D = \frac{3}{8Rv} \frac{1}{N_D}. \quad (5-74)$$

Если принять $v = 10^7$ см/с; $R = 3 \cdot 10^{-5}$ см; $N_D = 10^6$ см $^{-2}$, то время релаксации $\tau_D = 1,25 \cdot 10^{-9}$ с. Эта величина на 4—5 порядков больше, чем время релаксации для рассеяния на тепловых колебаниях решетки при комнатной температуре. Поэтому это рассеяние незначительно при комнатных температурах, но зато оно может быть существенно при низких температурах.

5-7. РАССЕЯНИЕ НА ТЕПЛОВЫХ КОЛЕБАНИЯХ РЕШЕТКИ

Взаимодействие электрона (или дырки) с колебаниями решетки осуществляется двояким образом: 1) электрон передает часть своей энергии решетке, так что определенное нормальное колебание с частотой ω_q увеличивает свое квантовое число на единицу. В результате такого процесса рассеяния образуется фонон с энергией $\hbar\omega_q$ и квазиимпульсом $Q = \hbar q$ и число фононов возрастает на еди-

ницу; $\mathbf{Q} = \hbar\mathbf{q}$ есть импульс, который приобретает решетка вследствие соударения с электронами; 2) электрон при взаимодействии получает часть энергии от решетки, так что квантовое число определенного свободного колебания с частотой ω_q уменьшается на единицу. В этом случае исчезает фонон с энергией $\hbar\omega_q$ и квазиимпульсом $\mathbf{Q} = \hbar\mathbf{q}$ и число фононов уменьшится 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 рассеяние электронов на тепловых колебаниях решетки должно зависеть от температуры. В кристалле возможен и многофононный механизм рассеяния, при котором имеет место поглощение или испускание более чем одного фонона. Однако вероятность такого процесса мала по сравнению с вероятностью однофононного рассеяния.

Пусть до столкновения электрон имел волновой вектор \mathbf{k} и энергию $E(\mathbf{k})$, а после столкновения его волновой вектор стал \mathbf{k}' и энергия $E(\mathbf{k}')$. Фонон, возникающий или исчезающий при столкновении, обладает энергией $\hbar\omega_q$ и волновым вектором \mathbf{q} . При взаимодействии электрона с фононом должны выполняться два закона: закон сохранения энергии и закон сохранения квазиимпульса. Поэтому при поглощении электроном фонона

$$E'(\mathbf{k}') = E(\mathbf{k}) + \hbar\omega_q; \quad (5-75)$$

$$\mathbf{k}' = \mathbf{k} + \mathbf{q}, \quad (5-76)$$

а число фононов, обладающих волновым числом q , в конечном состоянии уменьшится на единицу, т. е.

$$N'_q = N_q - 1. \quad (5-77)$$

При испускании электроном фонона

$$E(\mathbf{k}') = E(\mathbf{k}) - \hbar\omega_q; \quad (5-78)$$

$$\mathbf{k}' = \mathbf{k} - \mathbf{q}, \quad (5-79)$$

при этом

$$N'_q = N_q + 1. \quad (5-80)$$

Рассеяние носителя заряда может быть упругим и неупругим. Для характеристики рассеяния необходимо знание относительного изменения энергии δ электрона за одно столкновение

$$\delta = \frac{E' - E}{E} = \frac{\Delta E}{E} \quad (5-81)$$

или за единицу времени

$$A = \frac{\Delta E}{\tau} = \frac{E\delta}{\tau}. \quad (5-82)$$

В случае изотропного и упругого рассеяния $\delta \ll 1$ и среднее время рассеяния τ имеет смысл времени релаксации системы по импульсу, а $\tau_E = \tau/\delta$ — времени релаксации по энергии.

Рассмотрим процесс рассеяния электронов в атомном полупроводнике кубической структуры, энергетическая схема которого представлена на рис. 5-6, где a — постоянная решетки в отсутствие деформации. Пусть зона проводимости такого кристалла простая, так что энергия электрона в зоне проводимости равна:

$$E = \frac{\hbar^2 k^2}{2m^*}. \quad (5-83)$$

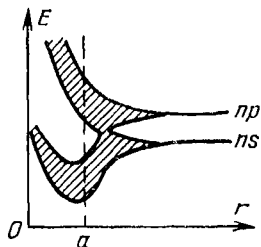


Рис. 5-6. Энергетическая схема атомного полупроводника.

В твердом теле поперечные волны представляют собой волны деформации сдвига и в кубическом кристалле не приводят к изменению объема. Волны продольные, являющиеся волнами деформации сжатия и растяжения, вызывают изменение объема кристалла. Как следует из рис. 5-6, 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 , ширина запрещенной зоны уменьшится. Следовательно, в таком кристалле *локальная деформация, создаваемая продольной акустической волной, приводит к волнообразному смещению дна зоны проводимости и потолка валентной зоны*, как это схематически представлено на рис. 5-7. Поэтому движущийся электрон, сталкиваясь с волной смещения, обусловленной тепловыми колебаниями решетки, будет рассеиваться только на продольных колебаниях.



Рис. 5-7. Изменение энергии дна зоны проводимости и потолка валентной зоны атомного полупроводника под воздействием продольных акустических колебаний решетки.

Покажем, что это рассеяние упругое и происходит на длинноволновых фонах. Закон сохранения энергии согласно (5-75), (5-79) и (5-83) запишется в виде

$$\frac{\hbar^2 k'^2}{2m^*} = \frac{\hbar^2 (\mathbf{k} \pm \mathbf{q})^2}{2m^*} = \frac{\hbar^2 k^2}{2m^*} \pm \hbar \omega_q, \quad (5-84)$$

откуда

$$q^2 \pm 2kq \cos \vartheta \mp \frac{2m^*}{\hbar} \omega_q = 0, \quad (5-85)$$

где ϑ — угол между направлениями векторов \mathbf{k} и \mathbf{q} . В равенстве (5-84) знак плюс относится к случаю поглощения, а минус — к случаю испускания фона. Если электрон взаимодействует только с длинноволновыми фонами, то $\omega_q = v_{зв} q$, где $v_{зв}$ — скорость распространения продольной звуковой волны. Учитывая это, из (5-85) найдем:

$$q = \mp 2k \cos \vartheta \pm \frac{2m^* v_{зв}}{\hbar}. \quad (5-86)$$

Так как $\frac{\hbar^2 k^2}{2m^*} = \langle E \rangle = kT$, то $k = \frac{1}{\hbar} \sqrt{2m^* kT}$

и

$$\frac{m^* v_{зв}}{\hbar k} = \sqrt{\frac{m^* v_{зв}^2}{2kT}} = \sqrt{\frac{T_{кр}}{T}}, \quad (5-87)$$

где $T_{кр} = \frac{m^* v_{зв}^2}{2k}$. Если принять $v_{зв} = 3 \cdot 10^5$ см/с, $m^* = 10^{-27}$ г, а $k = 10^{-16}$ эрг/град, то $T_{кр} = 1$ К. Поэтому при температурах, намного превышающих 1 К, можно пренебречь вторым слагаемым в (5-86) по сравнению с первым и записать:

$$q = \mp 2k \cos \vartheta. \quad (5-88)$$

В зависимости от ϑ волновое число q может меняться от $q_{мин} = 0$ до $q_{макс} = 2k$, т. е. в среднем на k . При $T = 300$ К $k = 10^7$ см $^{-1}$ и q может принимать значения от 0 до $2 \cdot 10^7$ см $^{-1}$, что соответствует энергии фононов от 0 до $\hbar \omega_q = \hbar v_{зв} q = \hbar v_{зв} k = 10^{-15}$ эрг = $6 \cdot 10^{-4}$ эВ. Но при комнатной температуре энергия электронов $E = \frac{\hbar^2 k^2}{2m^*} = 2,5 \cdot 10^{-2}$ эВ, поэтому

$$\frac{\hbar \omega_q}{E} \approx 2 \cdot 10^{-2} \ll 1. \quad (5-89)$$

Максимальное значение волнового вектора q_0 по теории Дебая согласно (3-96) равно: $q_0 = \left(\frac{6\pi^2 N}{V}\right)^{1/3} = \left(\frac{6\pi^2 N}{a^3 N}\right)^{1/3} \approx 10^8$ см $^{-1}$, следовательно, $q_{макс} \ll q_0$. Отсюда следует, что рассеяние электронов на продольных акустических колебаниях является упругим; при этом электрон рассеивается на длинноволновых фонах, поглощая или излучая фотон с $q = k$. Так как энергия фононов значительно меньше энергии электронов, то $k' = k$ и $E' = E$, поскольку в (5-84) можно не учитывать энергию фотона.

При низких температурах средняя скорость движения электрона приближается к скорости распространения упругих волн в кристалле и энергией фонона в (5-84) пренебречь нельзя. В этом случае имеет место неупругое столкновение электрона с фононом.

Поскольку электроны рассеиваются на продольных акустических фононах упруго, то для определения времени релаксации можно воспользоваться формулой (5-39):

$$\begin{aligned} \frac{1}{\tau} &= \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left[1 - \frac{k'_x}{k_x} \right] d\tau_{\mathbf{k}'} = \\ &= \frac{1}{4\pi^3} \int_{(V_B)} W(\mathbf{k}, \mathbf{k}') \left(\mp \frac{q_x}{k_x} \right) d\tau_{\mathbf{k}'} = \\ &= -\frac{1}{4\pi^3} \int_{(V_B)} W^+(\mathbf{k}, \mathbf{k}') \frac{q_x}{k_x} d\tau_{\mathbf{k}'} + \frac{1}{4\pi^3} \int_{(V_B)} W^-(\mathbf{k}, \mathbf{k}') \frac{q_x}{k_x} d\tau_{\mathbf{k}'} \quad (5-90) \end{aligned}$$

Здесь первый член правой части равенства учитывает поглощение, а второй — испускание фононов при рассеянии электронов.

Вероятность перехода в 1 с электрона из состояния \mathbf{k} в состояние \mathbf{k}' , как показывает расчет*, равна в случае поглощения фонона:

$$W^+(\mathbf{k}, \mathbf{k}') = W^+(\mathbf{k}, \mathbf{k} + \mathbf{q}) = \omega(q) N_q \delta(E' - E - \hbar\omega_q), \quad (5-91)$$

а в случае испускания фонона

$$W^-(\mathbf{k}, \mathbf{k}') = W^-(\mathbf{k}, \mathbf{k} - \mathbf{q}) = \omega(q) (N_q + 1) \delta(E' - E + \hbar\omega_q), \quad (5-92)$$

где $\delta(E' - E \pm \hbar\omega_q)**$ — дельта-функция, она выражает законы сохранения энергии и волнового вектора;

$$\omega(q) = \frac{4\pi}{9N} \frac{C^2 q^2}{M \omega_q}; \quad (5-93)$$

C — постоянная, имеющая размерность энергии; N — число элементарных ячеек в кристалле; M — масса атома.

В электронных процессах переноса функция распределения фононов N_q мало отличается от равновесной $\langle n \rangle$ (3-75), поэтому можно считать, что

$$N_q = \langle n \rangle = \frac{1}{e^{\hbar\omega_q/kT} - 1}. \quad (5-94)$$

* Ансельм А. И. Введение в теорию полупроводников. М., Физматгиз, 1962, с. 278.

** Ансельм А. И. Введение в теорию полупроводников. М., Физматгиз, 1962, с. 280.

Согласно (5-87) $\frac{\hbar\omega_q}{kT} = \frac{\hbar v_{зв}q}{kT} = \frac{\hbar v_{зв}k}{kT} = \sqrt{\frac{2T_{кр}}{T}} \ll 1$, поэтому выражение (5-94) можно разложить в ряд, в результате будем иметь:

$$\langle n \rangle = \langle n + 1 \rangle = \frac{kT}{\hbar v_{зв}q}. \quad (5-95)$$

Если в (5-90) подставить (5-91) и (5-92) с учетом (5-94) и (5-95), затем перейти к сферической системе координат, то получим:

$$\tau_l = \frac{9\pi}{4\sqrt{2}} \frac{\hbar^4 v_{зв}^2 M}{C^2 a^3 k m^{*3/2}} T^{-1} E^{-1/2} = \frac{\tau_{0l}}{m^{*3/2}} T^{-1} E^{-1/2}, \quad (5-96)$$

где $\tau_{0l} = \frac{9\pi \hbar^4 v_{зв}^2 M}{4\sqrt{2} a^3 k C^2}$; a — параметр решетки.

Длина свободного пробега

$$l = \tau_l v = \frac{\sqrt{2} \tau_{0l}}{m^{*2}} T^{-1}. \quad (5-97)$$

Таким образом, время релаксации при рассеянии электронов на длинноволновых акустических колебаниях решетки в атомных полупроводниках обратно пропорционально температуре в первой степени и энергии в степени 1/2, а длина свободного пробега не зависит от энергии носителей заряда.

Оценим время релаксации для германия в случае рассеяния электронов на колебаниях решетки. В чистом германии при комнатной температуре подвижность электронов $\mu_n = 0,39 \text{ м}^2/(\text{В}\cdot\text{с})$. Если принять $m_n^* = 0,3 m_0$, то $c/m_n^* = 6 \cdot 10^{11} \text{ Кл/кг}$, следовательно,

$$\tau_l = \frac{m_n^* \mu_n}{e} = \frac{0,39}{6 \cdot 10^{11}} = 6,5 \cdot 10^{-13} \text{ с}.$$

Из сравнения формул (5-96) и (5-70) 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 колебания. Каждое столкновение электрона с оптическим фононом приводит к возникновению или исчезновению фонона с энергией $\hbar\omega_{\text{оп}}$. Частота оптического фонона слабо зависит от квазиимпульса q , поэтому при соударении энергия электрона либо увеличивается, либо уменьшается на одинаковую величину, равную:

$$\pm \hbar\omega_{\text{оп}} = \pm \hbar\omega_0, \quad (5-98)$$

где ω_0 — максимальная частота продольной оптической ветви.

Энергия, передаваемая носителем заряда решетке за одно столкновение, определяется энергией возникшего фонона, умноженной на отношение разности вероятностей испускания и поглощения к сумме этих вероятностей. Но вероятность поглощения фонона согласно (5-91) пропорциональна концентрации фононов N_q , а вероятность испускания фонона согласно (5-92) пропорциональна $N_q + 1$. Следовательно, энергия, передаваемая электроном фонону, равна:

$$\begin{aligned} \Delta E &= \hbar\omega_0 \frac{W^-(\mathbf{k}, \mathbf{k}') - W^+(\mathbf{k}, \mathbf{k}')}{W^-(\mathbf{k}, \mathbf{k}') + W^+(\mathbf{k}, \mathbf{k}')} = \hbar\omega_0 \frac{(N_q + 1) - N_q}{(N_q + 1) + N_q} = \\ &= \hbar\omega_0 \frac{1}{2N_q + 1} = \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2kT}. \end{aligned} \quad (5-99)$$

Из формулы (5-99) следует, что на одно столкновение, сопровождающееся передачей энергии фонону, приходится $(2N_q + 1)$ столкновений.

При высоких температурах, когда $T \gg \theta_0$ или $kT \gg \hbar\omega_0$, т. е. когда энергия оптического фонона $\hbar\omega_0$ много меньше энергии электрона kT , тогда

$$\delta_{\text{оп}} = \frac{\Delta E}{E} = \frac{(\hbar\omega_0)^2}{2(kT)^2} \ll 1 \quad (5-100)$$

и рассеяние можно считать упругим. В этом случае у большинства электронов изменение энергии при поглощении или испускании оптического фонона будет незначительно. Поэтому можно ввести время релаксации τ и определить его по формуле (5-90). Для случая рассеяния на оптических фононах, как показывает расчет*,

* Ансельм А. И. Введение в теорию полупроводников. М., Физматгиз, 1962, § 6.

время релаксации равно:

$$\tau_{\text{оп}} = \frac{\sqrt{2}}{4\pi} \frac{Ma^3 (\hbar\omega_0)^2 E^{1.2}}{Z^2 e^4 m^{*1.2} kT}, \quad (5-101)$$

а длина свободного пробега

$$l_{\text{оп}} = \tau_{\text{оп}} v = \frac{aM}{2\pi m^*} \left(\frac{\hbar\omega_0}{Ze^2/a} \right)^2 \frac{E}{kT}, \quad (5-102)$$

где $M = \frac{m_1 m_2}{m_1 + m_2}$ — приведенная масса ионов; Ze — заряд иона; a — расстояние между ближайшими разноименными ионами.

При низких температурах, когда $T \ll \theta_0$ или $kT \ll \hbar\omega_0$, электроны могут только поглощать оптические фононы, в результате чего энергия их сильно увеличивается. Поскольку вероятность испускания фонона (5-92) в $(N_q + 1)/N_q$ раз превышает вероятность поглощения фонона (5-91), время пребывания электрона в состоянии с большой энергией будет очень мало. Можно считать, что электрон сразу после поглощения фонона испускает фонон той же энергии. Фактически рассеяние можно рассматривать как процесс обмена фононами, а поэтому энергия электрона практически остается неизменной, но направление его квазиимпульса сильно изменится. Это позволяет считать, что и при низких температурах, значительно меньших характеристической температуры $\theta_0 = \hbar\omega_0/k$, электрон, взаимодействуя с оптическими колебаниями решетки, рассеивается упруго.

Если учесть, что в случае низких температур

$$N_q = \langle n \rangle = \frac{1}{e^{\hbar\omega_0/kT} - 1} \approx e^{-\frac{\hbar\omega_0}{kT}}, \quad (5-103)$$

расчет для времени релаксации по формуле (5-90), в которой учитывается только поглощение фонона, дает выражение вида

$$\tau_{\text{оп}} = \frac{3\sqrt{2}Ma^3 (\hbar\omega_0)^{3.2}}{4\pi Z^2 e^4 m^{*1.2}} e^{\frac{\hbar\omega_0}{kT}}, \quad (5-104)$$

которое не зависит от энергии электрона, но экспоненциально зависит от температуры.

Длина свободного пробега электронов в случае рассеяния на оптических фононах при низких температурах определяется соотношением вида

$$l_{\text{оп}} = \frac{3a^3 M}{2m^*} \left(\frac{\hbar\omega_0}{Ze^2} \right)^2 e^{\hbar\omega_0/kT} \sqrt{\frac{E}{kT}}, \quad (5-105)$$

т. е. зависит от энергии носителей заряда в степени 1/2.

КИНЕТИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

6-1. НЕРАВНОВЕСНАЯ ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

При описании явлений переноса с помощью кинетического уравнения Больцмана необходимо знание неравновесной функции распределения для электронов и дырок. Ее определение проведем для невырожденного полупроводника, находящегося в электрическом и магнитном полях и имеющего градиент температуры. Для полупроводника со сферическими изоэнергетическими поверхностями энергия электронов

$$E = \frac{\hbar^2 k^2}{2m_n^*}, \quad (6-1)$$

а их скорость

$$v = \frac{1}{\hbar} \frac{\partial E}{\partial k} = \frac{\hbar}{m_n^*} k. \quad (6-2)$$

При наличии внешних электрического поля напряженности \mathcal{E} и магнитного поля с индукцией \mathbf{B} на электрон действует сила

$$\mathbf{F} = -e (\mathcal{E} + [\mathbf{vB}]) \quad (6-3)$$

и кинетическое уравнение (5-31) для стационарного случая примет вид:

$$(\mathbf{v} \nabla_{\mathbf{r}} f) - \frac{e}{\hbar} (\mathcal{E} + [\mathbf{vB}]) \nabla_{\mathbf{k}} f = -\frac{f_1(\mathbf{k})}{\tau_e}, \quad (6-4)$$

где τ_e — время релаксации электрона; $\mathcal{E} = -\nabla \varphi$, а φ — электростатический потенциал.

Будем считать, что неравновесная функция распределения $f(\mathbf{k}) = f_0 + f_1(\mathbf{k})$ мало отличается от равновесной f_0 . Тогда в уравнении (6-4) можно заменить f на f_0 везде, кроме члена, содержащего магнитное поле. Для этого случая уравнение (6-4) запишется в виде

$$(\mathbf{v} \nabla_{\mathbf{r}} f_0) - \frac{e}{\hbar} (\mathcal{E} \nabla_{\mathbf{k}} f_0) - \frac{e}{\hbar} [\mathbf{vB}] \nabla_{\mathbf{k}} f = -\frac{f_1(\mathbf{k})}{\tau_e}. \quad (6-5)$$

Для невырожденного электронного газа

$$\frac{\partial f_0}{\partial E} = -\frac{1}{kT} e^{-(E-F)/kT}; \quad (6-6)$$

$$\frac{\partial f_0}{\partial T} = \frac{E-F}{kT^2} e^{-(E-F)/kT} = \frac{F-E}{T} \frac{\partial f_0}{\partial E}; \quad (6-7)$$

$$\frac{\partial f_0}{\partial F} = \frac{1}{kT} e^{-(E-F)/kT} = -\frac{\partial f_0}{\partial E}, \quad (6-8)$$

поэтому

$$\nabla_{\mathbf{r}} f_0 = \frac{\partial f_0}{\partial T} \nabla T + \frac{\partial f_0}{\partial F} \nabla F = \frac{\partial f_0}{\partial E} \left(\frac{F-E}{T} \nabla T - \nabla F \right), \quad (6-9)$$

а

$$\nabla_{\mathbf{k}} f_0 = \frac{\partial f_0}{\partial E} \frac{\hbar^2}{m_n^*} \mathbf{k} = \hbar \frac{\partial f_0}{\partial E} \mathbf{v} \quad (6-10)$$

и

$$\begin{aligned} [\mathbf{vB}] \nabla_{\mathbf{k}} f &= [\mathbf{vB}] \nabla_{\mathbf{k}} (f_0 + f_1) = \\ &= \hbar \frac{\partial f_0}{\partial E} [\mathbf{vB}] \mathbf{v} + [\mathbf{vB}] \nabla_{\mathbf{k}} f_1. \end{aligned} \quad (6-11)$$

Но $[\mathbf{vB}] \mathbf{v} = [\mathbf{vv}] \mathbf{B} = 0$, а поскольку

$$f_1(\mathbf{k}) = -\frac{\partial f_0}{\partial E} \chi_n \mathbf{k} = -\frac{m_n^*}{\hbar} \frac{\partial f_0}{\partial E} (\chi_n \mathbf{v}), \quad (6-12)$$

то

$$\nabla_{\mathbf{k}} f_1 = -\nabla_{\mathbf{k}} \left\{ \frac{\partial f_0}{\partial E} (\chi_n \mathbf{k}) \right\} = -\frac{\partial f_0}{\partial E} \chi_n - \left\{ \hbar \mathbf{k} \frac{\partial}{\partial E} \left(\frac{\partial f_0}{\partial E} \chi_n \right) \right\} \mathbf{v}. \quad (6-13)$$

Тогда

$$[\mathbf{vB}] \nabla_{\mathbf{k}} f = [\mathbf{vB}] \nabla_{\mathbf{k}} f_1 = -\frac{\partial f_0}{\partial E} ([\mathbf{vB}] \chi_n) = -\frac{\partial f_0}{\partial E} [\mathbf{B} \chi_n] \mathbf{v}, \quad (6-14)$$

так как $[\mathbf{vB}] \chi_n = [\mathbf{B} \chi_n] \mathbf{v}$.

Далее, используя равенства (6-9), (6-10), (6-14) и (6-8), из уравнения (6-5) находим:

$$\begin{aligned} f_1(\mathbf{k}) &= -\frac{m_n^*}{\hbar} \frac{\partial f_0}{\partial E} (\chi_n \mathbf{v}) = \\ &= -\tau_e \frac{\partial f_0}{\partial E} \left\{ \frac{F-E}{T} \nabla T - \nabla (F - e\varphi) + \frac{e}{\hbar} [\mathbf{B} \chi_n] \right\} \mathbf{v}. \end{aligned} \quad (6-15)$$

Следовательно, на основании (6-12)

$$\chi_n = -\frac{\hbar}{m_n^*} \tau_e \left\{ \frac{E-F}{T} \nabla T + \nabla (F - e\varphi) - \frac{e}{\hbar} [\mathbf{B} \chi_n] \right\}. \quad (6-16)$$

Поскольку энергия E' и уровень Ферми F' дырок равны:

$$E' = -\frac{\hbar^2 \mathbf{k}'^2}{2m_p^*}; \quad (6-17)$$

$$F' = -E_g - F, \quad (6-18)$$

где E_g — ширина запрещенной зоны, то

$$\chi_p = -\frac{\hbar}{m_p^*} \tau_h \left\{ \frac{E' + E_g + F}{T} \nabla T - \nabla (F - e\varphi) + \frac{e}{\hbar} [\mathbf{B} \chi_p] \right\}. \quad (6-19)$$

Здесь τ_h — время релаксации дырки.

Определение χ_n и χ_p из уравнений (6-16) и (6-19) сводится к решению векторного уравнения вида

$$\mathbf{x} = \mathbf{a} + [\mathbf{b}\mathbf{x}]. \quad (6-20)$$

Умножим уравнение (6-20) скалярно на вектор \mathbf{b} :

$$(\mathbf{b}\mathbf{x}) = (\mathbf{b}\mathbf{a}) + \mathbf{b}[\mathbf{b}\mathbf{x}] = (\mathbf{b}\mathbf{a}). \quad (6-21)$$

Умножим теперь уравнение (6-20) векторно на вектор \mathbf{b} :

$$\begin{aligned} [\mathbf{b}\mathbf{x}] &= [\mathbf{b}\mathbf{a}] + [\mathbf{b}[\mathbf{b}\mathbf{x}]] = [\mathbf{b}\mathbf{a}] + \mathbf{b}(\mathbf{b}\mathbf{x}) - \mathbf{x}(\mathbf{b}\mathbf{b}) = \\ &= [\mathbf{b}\mathbf{a}] + \mathbf{b}(\mathbf{b}\mathbf{a}) - \mathbf{x}b^2. \end{aligned} \quad (6-22)$$

Тогда

$$\mathbf{x} = \mathbf{a} + [\mathbf{b}\mathbf{a}] + \mathbf{b}(\mathbf{b}\mathbf{a}) - \mathbf{x}b^2, \quad (6-23)$$

откуда

$$\mathbf{x} = \frac{\mathbf{a} + [\mathbf{b}\mathbf{a}] + \mathbf{b}(\mathbf{b}\mathbf{a})}{1 + b^2}. \quad (6-24)$$

Если воспользоваться уравнением (6-24), то выражения (6-16) и (6-19) запишутся в виде

$$\begin{aligned} \chi_n &= -\frac{\hbar\tau_e}{m_n^*} \times \\ &\times \left[\frac{\left\{ \frac{E-F}{T} \nabla T + \nabla(F-e\varphi) + \frac{e\tau_e}{m_n^*} \mathbf{B} \left\{ \frac{E-F}{T} \nabla T + \nabla(F-e\varphi) \right\} \right\}}{1 + \left(\frac{e\tau_e}{m_n^*} B \right)^2} + \right. \\ &\left. + \frac{\left(\frac{e\tau_e}{m_n^*} \right)^2 \left(\mathbf{B} \left\{ \frac{E-F}{T} \nabla T + \nabla(F-e\varphi) \right\} \right) \mathbf{B}}{1 + \left(\frac{e\tau_e}{m_n^*} B \right)^2} \right]; \end{aligned} \quad (6-25)$$

$$\chi_p = -\frac{\hbar\tau_h}{m_p^*} \times$$

$$\begin{aligned} &\times \left[\frac{\left\{ \frac{E'+E_g+F}{T} \nabla T - \nabla(F-e\varphi) - \frac{e\tau_h}{m_p^*} \mathbf{B} \left\{ \frac{E'+E_g+F}{T} \nabla T - \nabla(F-e\varphi) \right\} \right\}}{1 + \left(\frac{e\tau_h}{m_p^*} B \right)^2} + \right. \\ &\left. + \frac{\left(\frac{e\tau_h}{m_p^*} \right)^2 \left(\mathbf{B} \left\{ \frac{E'+E_g+F}{T} \nabla T - \nabla(F-e\varphi) \right\} \right) \mathbf{B}}{1 + \left(\frac{e\tau_h}{m_p^*} B \right)^2} \right]. \end{aligned} \quad (6-26)$$

6.2. УДЕЛЬНАЯ ПРОВОДИМОСТЬ ПОЛУПРОВОДНИКОВ

Рассчитаем удельную проводимость однородного невырожденного полупроводника при отсутствии градиента температуры, ($\nabla T = 0$) и магнитного поля ($\mathbf{B} = 0$), изоэнергетические поверхности которого представляют сферы. Пусть проводник находится в постоянном однородном электрическом поле напряженности \mathcal{E} .

В элементе объема $d\tau_{\mathbf{k}}$ кристалла единичного объема согласно (5-3) количество электронов

$$dn = f(\mathbf{k}) \frac{d\tau_{\mathbf{k}}}{4\pi^3}. \quad (6-27)$$

Эти электроны, движущиеся под действием электрического поля со скоростью \mathbf{v} , создают элементарную плотность тока

$$d\mathbf{j} = -e\mathbf{v}f(\mathbf{k}) \frac{d\tau_{\mathbf{k}}}{4\pi^3}. \quad (6-28)$$

Если в полупроводнике имеется два сорта носителей заряда, то полная плотность тока

$$\mathbf{j} = \mathbf{j}_n + \mathbf{j}_p = -\frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_n f_n(\mathbf{k}) d\tau_{\mathbf{k}} + \frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_p f_p(\mathbf{k}') d\tau_{\mathbf{k}'}, \quad (6-29)$$

где \mathbf{v}_n , $f_n(\mathbf{k})$ и \mathbf{v}_p , $f_p(\mathbf{k}')$ — скорость и неравновесная функция распределения электронов и дырок соответственно. Интегрирование нужно проводить по всей зоне Бриллюэна.

Если в (6-29) поставить равновесную функцию распределения f_0 , то

$$\mathbf{j} = \int_{(V_B)} \mathbf{v} f_0(\mathbf{k}) d\tau_{\mathbf{k}} = 0, \quad (6-30)$$

так как $f_0(\mathbf{k})$ — четная функция, а $\mathbf{v} f_0(\mathbf{k}) = \frac{\hbar}{m^*} \mathbf{k} f_0(\mathbf{k})$ является нечетной функцией и дает нуль при интегрировании в симметричных пределах. Выражение (6-30) означает, что *при термодинамическом равновесии в веществе нет тока*.

Если неравновесная функция

$$f_n(\mathbf{k}) = f_0 + f_1^{(n)}(\mathbf{k}) = f_0 - \frac{\partial f_0}{\partial E}(\chi_n \mathbf{k}), \quad (6-31)$$

то с учетом (6-30) вместо выражения (6-29) будем иметь:

$$\begin{aligned} \mathbf{j} &= -\frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_n f_1^{(n)}(\mathbf{k}) d\tau_{\mathbf{k}} + \frac{e}{4\pi^3} \int_{(V_B)} \mathbf{v}_p f_1^{(p)}(\mathbf{k}') d\tau_{\mathbf{k}'}, \\ &= \frac{e\hbar}{4\pi^3 m_n^*} \int_{(V_B)} \frac{\partial f_0}{\partial E}(\chi_n \mathbf{k}) \mathbf{k} d\tau_{\mathbf{k}} - \frac{e\hbar}{4\pi^3 m_p^*} \int_{(V_B)} \frac{\partial f_0}{\partial E'}(\chi_p \mathbf{k}') \mathbf{k}' d\tau_{\mathbf{k}'}. \end{aligned} \quad (6-32)$$

Поскольку полупроводник однородный, то $\nabla_{\mathbf{r}} f = 0$ и при отсутствии градиента температуры ($\nabla T = 0$) на основании равенства (6-9)

$$\nabla F = 0. \quad (6-33)$$

При наличии электрического поля $\mathcal{E} = -\nabla\varphi$ и отсутствии магнитного поля ($\mathbf{B} = 0$), учитывая равенство (6-33), из уравнений (6-25) и (6-26) получаем:

$$\chi_n = -\frac{e\hbar\tau_e}{m_n^*} \mathcal{E}; \quad (6-34)$$

$$\chi_p = \frac{e\hbar\tau_h}{m_p^*} \mathcal{E}. \quad (6-35)$$

Используя равенство (6-34), выражение для плотности электронного тока запишем в виде

$$\mathbf{j}_n = -\frac{e^2\hbar^2}{4\pi^3 m_n^{*2}} \int_{(V_B)} \tau_e \frac{\partial f_0}{\partial E} (\mathbf{k}\mathcal{E}) \mathbf{k} d\tau_{\mathbf{k}}. \quad (6-36)$$

Для невырожденного полупроводника при отсчете энергии от дна зоны проводимости количество электронов в зоне проводимости в соответствии с (4-66) определяется соотношением

$$n = N_c e^{-(E_c - F)/kT} = N_c e^{F/kT}. \quad (6-37)$$

Поскольку для невырожденного состояния $f_0 = e^{-(E-F)/kT}$, то

$$\frac{\partial f_0}{\partial E} = -\frac{1}{kT} e^{-(E-F)/kT} = -\frac{n}{kTN_c} e^{-E/kT}. \quad (6-38)$$

Подставляя равенство (6-38) в (6-36), получаем:

$$\mathbf{j}_n = \frac{e^2\hbar^2 n}{4\pi^3 m_n^{*2} kTN_c} \int_{(V_B)} \tau_e e^{-E/kT} (\mathbf{k}\mathcal{E}) \mathbf{k} d\tau_{\mathbf{k}}. \quad (6-39)$$

Если воспользоваться соотношением¹

$$\int F(\mathbf{k}) (\mathbf{a}\mathbf{k}) \mathbf{k} d\tau_{\mathbf{k}} = \frac{4\pi}{3} \mathbf{a} \int_0^\infty F(\mathbf{k}) k^4 dk, \quad (6-40)$$

то выражение (6-39) запишется в виде

$$\mathbf{j}_n = \frac{e^2\hbar^2 n \mathcal{E}}{4\pi^3 m_n^{*2} kTN_c} \frac{4\pi}{3} \int_0^\infty \tau_e e^{-E/kT} k^4 dk. \quad (6-41)$$

Полагая $\alpha = E/kT$, будем иметь:

$$k^4 dk = \frac{(2m_n^* kT)^{5/2}}{2\hbar^5} \alpha^{3/2} d\alpha, \quad (6-42)$$

¹ Ансельм А. И. Введение в теорию полупроводников. М., Физматгиз, 1962, с. 416.

а подставляя выражение (6-42) в (6-41) и учитывая, что $N_c = 2 \left(\frac{2\pi m_n^* kT}{\hbar^2} \right)^{3/2}$, получаем:

$$\mathbf{j}_n = \frac{e^2 n}{m_n^*} \mathcal{E} \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha. \quad (6-43)$$

Введем обозначение

$$\langle \tau_e \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha. \quad (6-44)$$

Тогда

$$\mathbf{j}_n = \frac{en \langle \tau_e \rangle}{m_n^*} \mathcal{E} = en\mu_n \mathcal{E} = \sigma_n \mathcal{E}, \quad (6-45)$$

где

$$\mu_n = \frac{e \langle \tau_e \rangle}{m_n^*} \quad (6-46)$$

есть дрейфовая подвижность, а

$$\sigma_n = en\mu_n \quad (6-47)$$

— удельная проводимость, обусловленная электронами.

Соответственно дырочная составляющая тока

$$\mathbf{j}_p = \frac{ep \langle \tau_h \rangle}{m_p^*} \mathcal{E} = ep\mu_p \mathcal{E} = \sigma_p \mathcal{E}. \quad (6-48)$$

Здесь среднее время релаксации дырок

$$\langle \tau_h \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_h e^{-\alpha} \alpha^{3/2} d\alpha, \quad (6-49)$$

дрейфовая подвижность дырок

$$\mu_p = \frac{e \langle \tau_h \rangle}{m_p^*} \quad (6-50)$$

и удельная проводимость, обусловленная дырками,

$$\sigma_p = ep\mu_p. \quad (6-51)$$

Следовательно, полная плотность тока в полупроводнике, имеющем два типа носителей заряда, равна:

$$\mathbf{j} = \mathbf{j}_n + \mathbf{j}_p = (\sigma_n + \sigma_p) \mathcal{E} = \sigma \mathcal{E}, \quad (6-52)$$

где $\sigma = \sigma_n + \sigma_p$ — удельная проводимость полупроводника.

6-3. ЗАВИСИМОСТЬ ПОДВИЖНОСТИ НОСИТЕЛЕЙ ЗАРЯДА ОТ ТЕМПЕРАТУРЫ

Как следует из (6-46) и (6-50), для вычисления подвижности носителей заряда необходимо знать среднее время релаксации. Однако на время релаксации существенным образом влияет механизм рассеяния, поэтому при столкновении носителей заряда с различными по природе дефектами кристаллической решетки подвижность по-разному будет зависеть от температуры.

А. Решеточное рассеяние

В атомных полупроводниках при рассеянии носителей заряда на акустических колебаниях решетки время релаксации согласно (5-96) равно:

$$\tau_l = \frac{\tau_0 l}{m^{*3/2}} T^{-1} E^{-1/2}. \quad (6-53)$$

На основании (6-44) найдем:

$$\langle \tau \rangle = \frac{4}{3\sqrt{\pi}} \frac{\tau_0 l}{m^{*3/2} k^{1/2}} T^{-3/2} \quad (6-54)$$

и

$$\mu_l = \frac{4e}{3\sqrt{\pi}} \frac{\tau_0 l}{k^{1/2} m^{*5/2}} T^{-3/2}. \quad (6-55)$$

Подвижность можно выразить через длину свободного пробега. Используя формулу (5-97), получаем:

$$\mu_l = \frac{4cl}{3(2\pi m^* kT)^{1/2}}. \quad (6-56)$$

Из (6-55) следует, что в атомных полупроводниках подвижность носителей заряда при рассеянии их на тепловых колебаниях решетки уменьшается с ростом температуры. Кроме того, подвижность обратно пропорциональна эффективной массе носителей заряда в степени 5/2. Действительно, как показывает опыт, подвижность электронов больше подвижности дырок, а у полупроводников с малой эффективной массой носителей заряда подвижность очень большая (см. табл. 6-1).

Б. Рассеяние ионизированными примесями

При рассеянии на ионах примеси время релаксации в зависимости от энергии согласно (5-69) имеет вид:

$$\tau_l = \frac{l \sqrt{2} \varepsilon^2 m^{*1/2} E^{3/2}}{\pi Z^2 e^4 N_I \ln \left[1 + \left(\frac{\varepsilon E}{Z e^2 N_I^{1/2}} \right)^2 \right]}. \quad (6-57)$$

При усреднении этого выражения логарифмический член, являющийся медленно меняющейся функцией, можно вынести за знак интеграла, если в этот член подставить такое значение энергии, при котором остальная часть подынтегрального выражения дости-

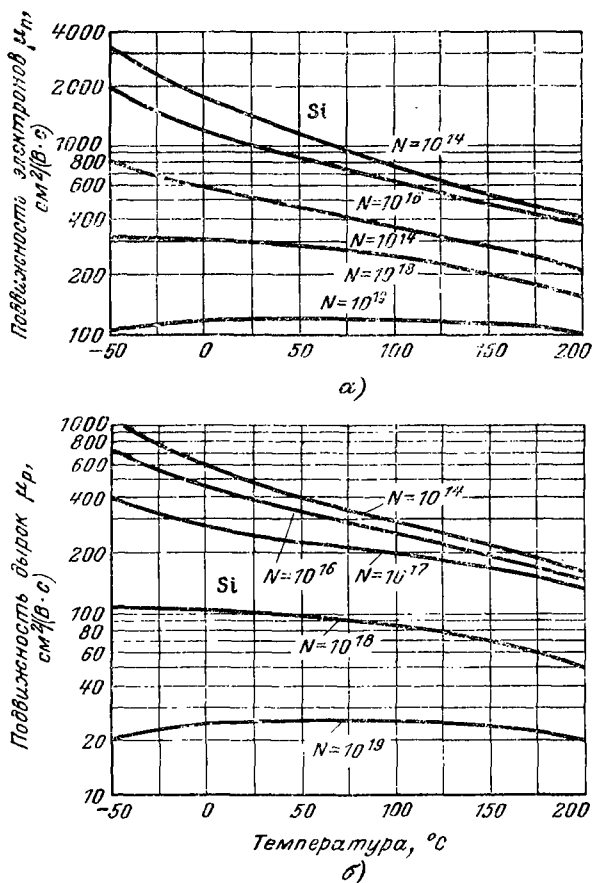


Рис. 6-1. Зависимость дрейфовой подвижности электронов (а) и дырок (б) от температуры в образцах кремния, имеющих разную концентрацию примеси.

гает максимума. Для этого необходимо положить $E = 3kT$. Поэтому среднее время релаксации будет равно:

$$\langle \tau_I \rangle = \frac{8\sqrt{2} \epsilon^2 k^3 m^{*1/2} T^{1/2}}{\pi^3 Z^2 e^4 N_I \ln \left[1 + \left(\frac{3\epsilon kT}{Ze^2 N_I} \right)^2 \right]}, \quad (6-58)$$

а подвижность при рассеянии ионами примеси запишется в виде

$$\mu_I = \frac{8\sqrt{2}e^2k^{3/2}T^{3/2}}{\pi^{3/2}Z^2e^3N_I m^{*1/2} \ln \left[1 + \left(\frac{3ekT}{Ze^2N_I^{1/3}} \right)^2 \right]} \quad (6-59)$$

При достаточно больших температурах можно считать, что

$$\mu_I \approx \mu_{0I} T^{3/2} \quad (6-60)$$

Из формулы (6-60) следует, что подвижность носителей заряда, обусловленная рассеянием на ионах примеси, уменьшается при уменьшении температуры. Это происходит потому, что с увеличением температуры кристалла возрастает тепловая скорость носителей заряда, а это приводит к ослаблению взаимодействия носителей заряда с 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. 6-1 представлена зависимость подвижности электронов и дырок от температуры в образцах кремния, имеющих разную концентрацию примесей.

В. Рассеяние нейтральными примесями

При рассеянии носителей заряда на нейтральной примеси время релаксации согласно выражению (5-73) не зависит от температуры и энергии носителя заряда. В силу этого подвижность носителей заряда, обусловленная рассеянием на нейтральных атомах примеси, в явном виде не зависит от температуры и имеет вид:

$$\mu_A = \frac{e^2 m^*}{20e\hbar^3} \frac{1}{N_A} \quad (6-61)$$

Г. Рассеяние на дислокациях

При рассеянии носителей заряда на дислокациях подвижность

$$\mu_D = \frac{eT^{-1/2}}{2^{3/2}\pi^{1/2}k^{1/2}m^{*1/2}RN_D} \quad (6-62)$$

Д. Сложный механизм рассеяния

В реальных полупроводниках строгая периодичность поля решетки нарушена в результате наличия в ней разных дефектов. В силу этого механизм рассеяния носит сложный характер. Если

считать, что все механизмы рассеяния независимы, то полная вероятность рассеяния равна сумме вероятностей рассеяния на каждом типе рассеивающих центров, т. е.

$$W(\mathbf{k}, \mathbf{k}') = \sum_i W_i(\mathbf{k}, \mathbf{k}'). \quad (6-63)$$

В этом случае полное время релаксации τ согласно выражению (6-40) будет равно:

$$\tau^{-1} = \frac{1}{4\pi^3} \int_{(V_B)} \sum_i W_i(\mathbf{k}, \mathbf{k}') (1 - \cos \theta) d\tau_{\mathbf{k}'} = \sum_i \tau_i^{-1} \quad (6-64)$$

или

$$\tau = \left(\sum_i \tau_i^{-1} \right)^{-1}. \quad (6-65)$$

Естественно, что вклад различных механизмов рассеяния в суммарное время релаксации зависит от температуры кристалла. По мере изменения энергии носителей заряда роль одних механизмов рассеяния уменьшается, а роль других возрастает, поэтому характер движения носителей заряда под воздействием внешнего поля зависит от того, какой из механизмов рассеяния является доминирующим в данном интервале температур. При очень низких температурах в выражении (6-65) можно учитывать только рассеяние на атомах примеси и дислокациях. С повышением температуры роль этих механизмов уменьшается по сравнению с рассеянием на ионах примеси. При высоких температурах доминирующим становится рассеяние на фононах. Поэтому можно считать, что комбинированная подвижность определяется рассеянием на ионах примеси и на фононах и согласно (6-65) равна $\mu = (\mu_i^{-1} + \mu_f^{-1})^{-1}$, а в зависимости от температуры с учетом формул (6-55) и (6-59) будет определяться соотношением вида

$$\frac{1}{\mu} = aT^{-3/2} + bT^{1/2}, \quad (6-66)$$

где a и b — постоянные величины.

Как следует из формулы (6-66), в атомных полупроводниках с ростом температуры подвижность носителей заряда растет пропорционально $T^{3/2}$, если рассеяние происходит только на ионах при-

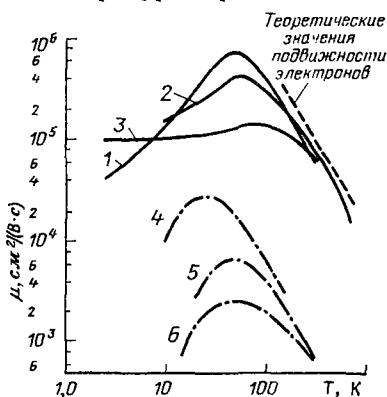


Рис. 6-2. Температурные зависимости подвижности электронов и дырок в антимониде индия.

1 — $n = 2 \cdot 10^{11}$; 2 — $n = 10^{15}$; 3 — $n = 10^{16}$; 4 — $p = 3 \cdot 10^{11}$; 5 — $p = 4 \cdot 10^{15}$; 6 — $p = 3 \cdot 10^{16}$ см⁻³.

меси, затем она 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. 6-2 представлены температурные зависимости подвижностей электронов и дырок в антимониде индия.

Таблица 6-1

Экспериментальные данные по подвижности электронов и дырок некоторых полупроводников

Полупроводник	Подвижность электронов, см ² /(В·с)		Подвижность дырок, см ² /(В·с)		Показатель степени температурной зависимости решеточной подвижности ρ		$b = \mu_n \mu_p$	
	300 К	77 К	300 К	77 К	Электроны	Дырки	300 К	77 К
Ge	3 900	37 100	1900	43 700	1,66	2,33	2	0,8
Si	1 500	45 500	600	11 600	2,5	2,7	2,5	3,9
InSb	78 000	1 200 000	750	10 000	1,6	2,1	100	120
InAs	33 000	820 000	460	690	1,2	2,3	70	120
InP	4 600	24 000	150	1 200	2,0	2,4	30	20
GaSb	4 000	6 000	1400	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	1,2
AlSb	200		420			1,8	0,2	

Но, как следует из данных табл. 6-1, у большинства полупроводников температурная зависимость подвижности при рассеянии на тепловых колебаниях решетки не соответствует теоретическому значению показателя степени $3/2$. Это объясняется тем, что в реальных полупроводниках имеют место и другие механизмы рассеяния, такие как рассеяние на оптических колебаниях, двухфонное рассеяние, рассеяние на носителях заряда.

Е. Температурная зависимость удельной проводимости

Учитывая зависимость подвижности и концентрации носителей заряда от температуры, удельную проводимость собственного полупроводника можно записать в виде

$$\sigma = en_i \mu_n + ep_i \mu_p = AT^{3/2+p} e^{-E_g/2kT}. \quad (6-67)$$

Множитель $T^{3/2+p}$ медленно меняется с температурой, тогда как множитель $e^{-E_g/2kT}$ сильно зависит от температуры, если $E_g \gg kT$. Поэтому для не слишком высоких температур можно считать, что

$$AT^{3/2+p} = \sigma_0 = \text{const} \quad (6-68)$$

и выражение (6-67) можно заменить более простым:

$$\sigma = \sigma_0 e^{-E_g/2kT}. \quad (6-69)$$

В примесном полупроводнике, как следует из рис. 6-3, б, при достаточно высоких температурах электропроводность полупроводника является собственной. При низких температурах в области

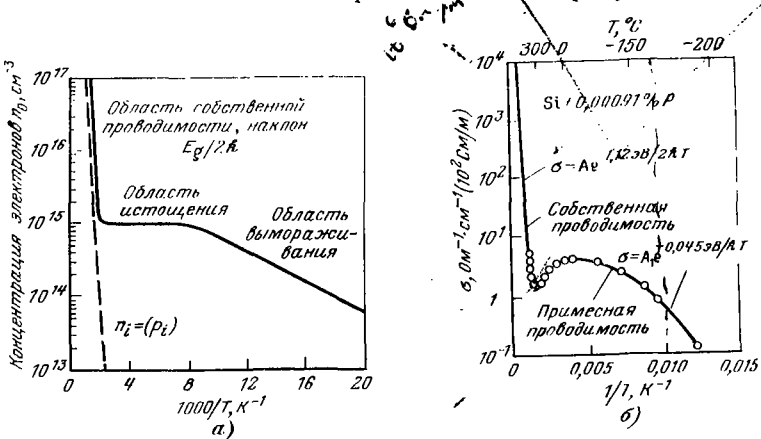


Рис. 6-3. Зависимость концентрации носителей заряда (а) и удельной проводимости (б) от обратной температуры для образца кремния с концентрацией донорных примесей около 10^{15} см^{-3} .

истощения примеси концентрация основных носителей заряда остается постоянной (рис. 6-3, а) и электропроводность меняется вследствие изменения подвижности носителей заряда с температурой.

6-4. ЭФФЕКТ ХОЛЛА

Кинетические эффекты, имеющие место при одновременном воздействии на проводник электрического и магнитного полей, называют гальваномагнитными эффектами. Эффект Холла является одним из гальваномагнитных эффектов. Рассмотрение эффекта Холла 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:

$$\tau \ll T_c. \quad (6-70)$$

Но частота обращения носителя заряда с эффективной массой m^* по круговой орбите в магнитном поле с индукцией B , т. е. частота ω_c циклотронного резонанса согласно выражению (2-150) равна:

$$\omega_c = \frac{2\pi}{T} = \frac{e}{m^*} B, \quad (6-71)$$

поэтому в слабых полях

$$\frac{\tau\omega_c}{2\pi} = \frac{e\tau}{2\pi m^*} B = \frac{\mu B}{2\pi} \ll 1, \quad (6-72)$$

а в сильных полях

$$\frac{\tau\omega_c}{2\pi} = \frac{\mu B}{2\pi} \gg 1. \quad (6-73)$$

Таким образом, условие слабого и сильного магнитного поля определяется не только значением индукции магнитного поля B , но и значением подвижности носителей заряда μ .

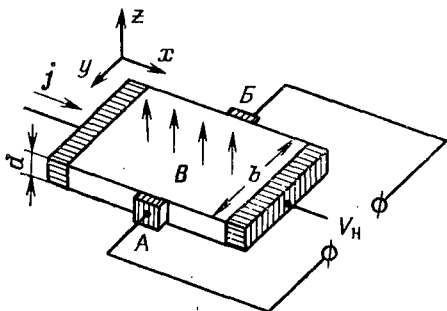


Рис. 6-4. Образец для измерения эффекта Холла.

Если воспользоваться длиной свободного пробега $l = vt$ и радиусом циклотронной орбиты $r = v/\omega_c$, по которой движется носитель заряда, то

$$\frac{l}{r} = \tau\omega_c. \quad (6-74)$$

Как видно, в слабом магнитном поле, поскольку $r \gg l$, носитель заряда, движущийся по круговому пути в плоскости, перпендикулярной \mathbf{B} , успевает пройти до столкновения малое расстояние по круговой орбите, а в сильном магнитном поле, для которого $r \ll l$, траектория носителя заряда искривляется очень сильно.

Пусть по полупроводнику, имеющему форму прямоугольной пластинки (рис. 6-4), под действием электрического поля \mathcal{E} протекает ток

$$\mathbf{j} = -en\mathbf{v} = \sigma\mathcal{E}. \quad (6-75)$$

Если полупроводник однородный, то эквипотенциальные поверхности расположены перпендикулярно к направлению электрического поля \mathcal{E} , а следовательно, и к вектору плотности тока \mathbf{j} . Поэтому разность потенциалов между точками A и B на рис. 6-4, лежащими в плоскости, перпендикулярной \mathbf{j} , будет равна нулю.

Теперь поместим полупроводник в магнитное поле, перпендикулярное вектору тока, как это показано на рис. 6-4. В этом случае на носитель заряда, движущийся с дрейфовой скоростью \mathbf{v} , будет действовать сила Лоренца

$$\mathbf{F} = \pm e[\mathbf{v}\mathbf{B}], \quad (6-76)$$

направленная перпендикулярно \mathbf{v} и \mathbf{B} . Здесь плюс соответствует дырке, а минус — электрону. Но

$$\mathbf{v} = \pm \mu\mathcal{E} = \pm \frac{e\langle\tau\rangle}{m^*} \mathcal{E}, \quad (6-77)$$

поэтому

$$\mathbf{F} = \frac{e^2 \langle \tau \rangle}{m^*} [\mathcal{E} \mathbf{B}]. \quad (6-78)$$

Из (6-78) следует, что направление силы Лоренца не зависит от знака носителя заряда, а определяется направлением векторов \mathcal{E} и \mathbf{B} или \mathbf{j} и \mathbf{B} . Следовательно, если скорость носителей заряда определяется электрическим полем, то электроны и дырки под действием силы Лоренца отклоняются в одну и ту же сторону.

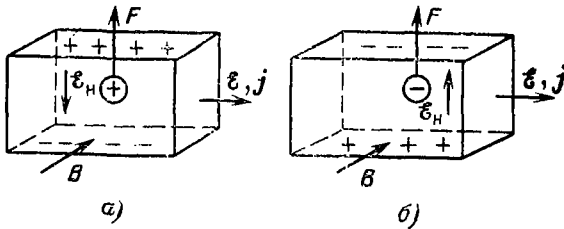


Рис. 6-5. Отклонение носителей заряда под воздействием магнитного поля в образцах с дырочной (а) и электронной (б) электропроводностью.

Для выбранных нами направлений \mathbf{B} и \mathcal{E} , представленных на рис. 6-5, сила Лоренца \mathbf{F} направлена вверх. Под действием этой силы дырки в акцепторном полупроводнике (рис. 6-5, а) и электроны в донорном полупроводнике (рис. 6-5, б) будут отеснены к верхней поверхности образца, вследствие чего на нижней поверхности возникнет их дефицит, что обусловит противоположный по знаку заряд по отношению к заряду на верхней поверхности. В результате разделения зарядов появится электрическое поле \mathcal{E}_H , перпендикулярное к направлению магнитного поля. Направление этого поля, которое называется полем Холла, зависит от знака носителей заряда. В нашем случае (рис. 6-5) поле Холла \mathcal{E}_H направлено вниз в p -образце и вверх в n -образце. Явление возникновения в полупроводнике с текущим по нему током поперечного электрического поля под действием магнитного поля называется эффектом Холла. Поле Холла \mathcal{E}_H будет расти до тех пор, пока сила, обусловленная этим полем, не скомпенсирует силу Лоренца:

$$-e\mathcal{E}_H = evB. \quad (6-79)$$

При этом условии носители заряда движутся вдоль образца под действием только продольного электрического поля \mathcal{E} , а следовательно, ток \mathbf{j} по направлению совпадает с полем \mathcal{E} . Вектор суммарного электрического поля $\mathcal{E}' = \mathcal{E} + \mathcal{E}_H$, как видно из рис. 6-6, повернут на некоторый угол φ относительно направления тока \mathbf{j} . Угол, заключенный между \mathbf{j} и \mathcal{E}' , носит название угла Холла. Теперь эквипотенциальные поверхности будут повернуты на угол φ от-

носителю первоначального положения, поэтому на образце между точками A и B (рис. 6-4) появится разность потенциалов, называемая э. д. с. Холла. Если ширина образца b , то холловская разность потенциалов

$$V_H = \mathcal{E}_H b = -v B b. \quad (6-80)$$

Если воспользоваться соотношением (6-75), будем иметь:

$$V_H = -\frac{1}{en} j B b = R j B b. \quad (6-81)$$

Величина R в уравнении (6-81) называется коэффициентом Холла (или постоянной) Холла, который в случае электронов равен:

$$R = -\frac{1}{en}. \quad (6-82)$$

Если носителями заряда являются дырки, концентрация которых равна p , то, как видно из рис. 6-5, a , $e\mathcal{E}_H = evB$, поэтому

$$R = \frac{1}{ep}. \quad (6-83)$$

Как следует из равенств (6-82) и (6-83), коэффициент Холла обратно пропорционален концентрации носителей заряда, а знак его совпадает со знаком носителей заряда. При таком рассмотрении эффекта Холла не принималось во внимание статистическое распре-

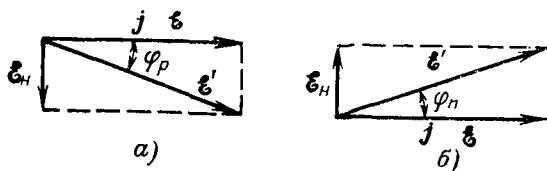


Рис. 6-6. Угол Холла при дырочной (а) и электронной (б) электропроводности.

деление носителей заряда по энергиям и не учитывалась зависимость времени релаксации от энергии. Это будет сделано в § 6-5.

Угол Холла φ (см. рис. 6-6) можно определить из соотношения

$$\operatorname{tg} \varphi = \frac{\mathcal{E}_H}{\mathcal{E}}, \quad (6-84)$$

причем знак $\operatorname{tg} \varphi$ определяется направлением поля Холла: положителен для дырочного полупроводника и отрицателен для электронного. Если магнитное поле слабое, то можно считать, что

$$\varphi = \frac{\mathcal{E}_H}{\mathcal{E}} = \frac{R j B}{\mathcal{E}} = R \sigma B, \quad (6-85)$$

откуда следует, что

$$R = \frac{\varphi}{\sigma B}. \quad (6-86)$$

Учитывая, что для электронного полупроводника

$$\sigma_n = en\mu_n, \quad (6-87)$$

то угол Холла:

$$\varphi_n = R_n \sigma_n B = \mu_n B. \quad (6-88)$$

Аналогично для дырочного полупроводника

$$\varphi_p = R_p \sigma_p B = \mu_p B. \quad (6-89)$$

6-5. ЭФФЕКТ ХОЛЛА В ПОЛУПРОВОДНИКАХ С ДВУМЯ ТИПАМИ НОСИТЕЛЕЙ ЗАРЯДА

Пусть в полупроводнике ток переносится электронами и дырками, концентрации и подвижности которых соответственно равны n , μ_n и p , μ_p .

Определение коэффициента Холла проведем для случая слабого магнитного поля. Пусть вектор напряженности электрического поля \mathcal{E} , вектор плотности тока \mathbf{j} и вектор магнитной индукции \mathbf{B} ориентированы относительно осей координат, как это показано на рис. 6-7.

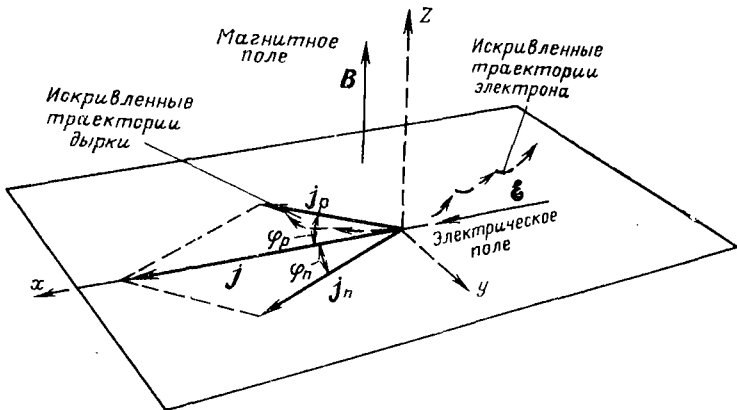


Рис. 6-7. Эффект Холла при наличии двух типов носителей заряда.

Под действием электрического поля носители заряда будут ускоряться. По мере увеличения дрейфовой скорости на носитель заряда будет действовать отклоняющая сила Лоренца. Траектория носителя заряда под действием магнитного поля искривляется до тех пор, пока он не испытывает столкновения с дефектами кристал-

лической решетки (магнитное поле слабое и $l \ll r$). При столкновении носитель заряда передаст решетке энергию, накопленную в электрическом поле, а затем он опять будет двигаться под действием поля \mathcal{E} . В результате ток дырок \mathbf{j}_p отклоняется от направления электрического поля на угол φ_p . Электроны под действием магнитного поля отклоняются в ту же сторону, что и дырки, а электронный ток \mathbf{j}_n отклоняется от направления электрического поля в противоположную сторону на угол φ_n (рис. 6-7).

При этом суммарная плотность тока, если использовать соотношение (6-32), будет равна:

$$\mathbf{j} = \frac{e\hbar}{4\pi^3 m_n^*} \int_{(V_B)} \frac{\partial f_0}{\partial E} (\chi_n \mathbf{k}) \mathbf{k} d\tau_{\mathbf{k}} - \frac{e\hbar}{4\pi^3 m_p^*} \int_{(V_B)} \frac{\partial f_{0p}}{\partial E'} (\chi_p \mathbf{k}') \mathbf{k}' d\tau_{\mathbf{k}'}. \quad (6-90)$$

Вычисление плотности электронного \mathbf{j}_n и дырочного \mathbf{j}_p токов проведем для однородного полупроводника со сферическими изоэнергетическими поверхностями при отсутствии градиента температуры, находящегося в слабом магнитном поле. Так как $\nabla_r f = \frac{E-F}{T} \nabla T + \nabla F = 0$, а $\mathcal{E} = -\nabla\varphi$, то для слабого магнитного поля, когда $\frac{e\tau}{m^*} B = \mu B \ll 1$, уравнения (6-25) и (6-26) примут вид:

$$\chi_n = -\frac{e\hbar\tau_e}{m_n^*} \left\{ \mathcal{E} + \frac{e\tau_e}{m_n^*} [\mathbf{B}\mathcal{E}] \right\}; \quad (6-91)$$

$$\chi_p = \frac{e\hbar\tau_h}{m_p^*} \left\{ \mathcal{E} - \frac{e\tau_h}{m_p^*} [\mathbf{B}\mathcal{E}] \right\}. \quad (6-92)$$

Находим теперь электронную составляющую тока, используя соотношение (6-91):

$$\mathbf{j}_n = -\frac{e^2 \hbar^2}{4\pi^3 m_n^{*2}} \int_{(V_B)} \frac{\partial f_0}{\partial E} \left\{ \tau_e (\mathbf{k}\mathcal{E}) + \frac{e\tau_e^2}{m_n^*} \mathbf{k} [\mathbf{B}\mathcal{E}] \right\} \mathbf{k} d\tau_{\mathbf{k}}. \quad (6-93)$$

Если учесть равенства (6-40), (6-41), (6-42), то уравнение (6-93) запишется в виде

$$\mathbf{j}_n = \frac{e^2 n}{m_n^*} \mathcal{E} \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-\alpha} \alpha^{3/2} d\alpha + \frac{e^3 n}{m_n^{*2}} [\mathbf{B}\mathcal{E}] \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e^2 e^{-\alpha} \alpha^{3/2} d\alpha. \quad (6-94)$$

Вводя среднее значение времени релаксации для электронов

$$\langle \tau_e \rangle = \frac{4}{3\sqrt{\pi}} \int_0^{\infty} \tau_e e^{-\alpha} \alpha^{3/2} d\alpha \quad (6-95)$$

и среднее значение квадрата времени релаксации

$$\langle \tau_e^2 \rangle = \frac{4}{3\sqrt{\pi}} \int_0^{\infty} \tau_e^2 e^{-\alpha} \alpha^{3/2} d\alpha, \quad (6-96)$$

а также учитывая, что $\mu_n = e \langle \tau_e \rangle / m_n^*$, получаем для плотности электронной составляющей тока:

$$\mathbf{j}_n = en\mu_n \mathcal{E} + \frac{\langle \tau_e^2 \rangle}{|\langle \tau_e \rangle|^2} en\mu_n^2 [\mathbf{B}\mathcal{E}] = en\mu_n \mathcal{E} + r_n en\mu_n^2 [\mathbf{B}\mathcal{E}], \quad (6-97)$$

где

$$r_n = \frac{\langle \tau_e^2 \rangle}{|\langle \tau_e \rangle|^2}. \quad (6-98)$$

По аналогии плотность дырочной составляющей тока запишется в виде

$$\mathbf{j}_p = ep\mu_p \mathcal{E} - r_p ep\mu_p^2 [\mathbf{B}\mathcal{E}], \quad (6-99)$$

где

$$r_p = \frac{\langle \tau_h^2 \rangle}{|\langle \tau_h \rangle|^2}. \quad (6-100)$$

Полная плотность тока

$$\mathbf{j} = e(n\mu_n + p\mu_p) \mathcal{E} - r_n e \left(\frac{r_p}{r_n} p\mu_p^2 - n\mu_n^2 \right) [\mathbf{B}\mathcal{E}]. \quad (6-101)$$

В уравнении (6-101) коэффициенты r_n и r_p зависят от механизма рассеяния носителей заряда, который определяет время релаксации $\tau(\mathbf{k})$. В случае упругого рассеяния, когда $k' = k$ и $E' = E$, для любого механизма рассеяния отношение

$$\frac{r_p}{r_n} = \frac{\langle \tau_h^2 \rangle}{|\langle \tau_h \rangle|^2} \frac{|\langle \tau_e \rangle|^2}{\langle \tau_e^2 \rangle} = 1. \quad (6-102)$$

Обозначая $r_n = r$ и учитывая (6-102), будем иметь:

$$\mathbf{j} = e(n\mu_n + p\mu_p) \mathcal{E} - re(p\mu_p^2 - n\mu_n^2) [\mathbf{B}\mathcal{E}]. \quad (6-103)$$

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 $B_y = B_x = 0$ (рис. 6-7), при этом равенство (6-103) распадется на два уравнения:

$$j_x = e(n\mu_n + p\mu_p) \mathcal{E}_x + re(p\mu_p^2 - n\mu_n^2) B \mathcal{E}_y = j; \quad (6-104)$$

$$j_y = e(n\mu_n + p\mu_p) \mathcal{E}_y - re(p\mu_p^2 - n\mu_n^2) B \mathcal{E}_x = 0. \quad (6-105)$$

Решая эти уравнения относительно \mathcal{E}_y , в случае слабого магнитного поля получаем для напряженности поля Холла:

$$\mathcal{E}_y = \frac{r}{e} \frac{p\mu_p^2 - n\mu_n^2}{(n\mu_n + p\mu_p)^2} jB = RjB, \quad (6-106)$$

где коэффициент Холла для полупроводника с двумя типами носителей заряда

$$R = \frac{r}{e} \frac{p\mu_p^2 - n\mu_n^2}{(n\mu_n + p\mu_p)^2}. \quad (6-107)$$

Для собственного полупроводника, у которого $n = p = n_i$,

$$R_i = \frac{r}{e} \frac{\mu_p - \mu_n}{\mu_p + \mu_n} = \frac{r}{e} \frac{1-b}{1+b}. \quad (6-108)$$

Так как обычно $b = \mu_n/\mu_p > 1$, то $R_i < 0$.

Как видим, коэффициент Холла зависит от механизма рассеяния, что учитывается сомножителем r , который носит название Холл-фактора. Определим значение r для разных механизмов рассеяния. При рассеянии на акустических колебаниях решетки время релаксации (5-96) с учетом $E = kT\alpha$ равно:

$$\tau_l = \frac{\tau_{0l}}{m_n^{*1/2}} T^{-1} E^{-1/2} = \frac{\tau_{0l}}{m_n^{*3/2} k^{1/2} T^{1/2}} \alpha^{-1/2}; \quad (6-109)$$

на основании (6-95), (6-96) и (6-98) найдем:

$$r = \frac{3\pi}{8}. \quad (6-110)$$

При рассеянии на ионах примеси, используя соотношение (5-70), получаем:

$$r = \frac{315\pi}{512} = 1,93. \quad (6-111)$$

Поскольку рассеяние носителей заряда зависит от температуры, то при определении коэффициента Холла при низких температурах, когда основную роль в рассеянии играют ионы примеси, необходимо полагать $r = 1,93$. Для температур, при которых имеет место рассеяние на акустических фононах, $r = \frac{3\pi}{8}$. В том случае, когда

в процессе рассеяния одновременно участвуют как колебания решетки, так и ионы примеси, r имеет более сложное выражение. На рис. 6-8 показана зависимость величины r от отношения удельного сопротивления полупроводника ρ_I при рассеянии только на ионах примеси к полному сопротивлению ρ при рассеянии на ионах и на фононах.

Для электронного полупроводника, у которого $n \gg p$, коэффициент Холла на основании (6-107) имеет вид:

$$R = -r \frac{1}{en}. \quad (6-112)$$

В то же время удельная проводимость такого полупроводника, определяемая дрейфовой подвижностью μ_d , равна $\sigma = e n \mu_d$. Следовательно,

$$\frac{1}{r} |R| \sigma = \mu_d. \quad (6-113)$$

Но с помощью равенства

$$|R| \sigma = \mu_H \quad (6-114)$$

можно определить μ_H , имеющую размерность подвижности и вследствие этого называемую холловской подвижностью. Из формул (6-113) и (6-114) следует, что

$$\mu_H = r \mu_d, \quad (6-115)$$

т. е. холловская подвижность, определяющая угол Холла, пропорциональна дрейфовой подвижности.

В случае, когда время релаксации $\tau = \text{const}$, т. е. не зависит от энергии,

$$r = 1, \quad (6-116)$$

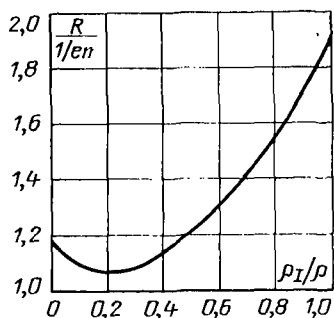


Рис. 6-8. Зависимость коэффициента $r = \frac{R}{1/en}$ от отношения ρ_I/ρ .

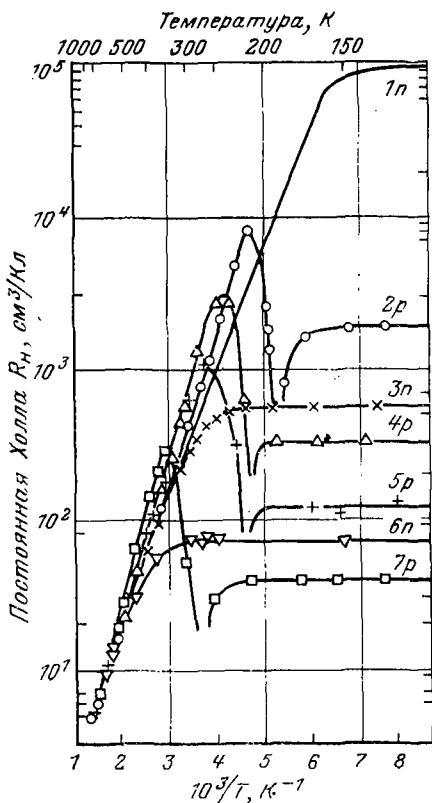


Рис. 6-9. Температурная зависимость постоянной Холла R в антимоноиде индия для образцов n - и p -типа.

Концентрация доноров и акцепторов, см^{-3} : $1n = 4,3 \cdot 10^{13}$; $2p = 4 \cdot 10^{13}$; $3n = 1,3 \cdot 10^{16}$; $4p = 2,2 \times 10^{16}$; $5p = 6 \cdot 10^{16}$; $6n = 1 \cdot 10^{17}$; $7p = 2 \cdot 10^{17}$.

поэтому

$$\mu_H = \mu_a \quad (6-117)$$

Это имеет место у металлов и вырожденных полупроводников. У них коэффициент Холла R не зависит от механизма рассеяния.

Поскольку n и p являются функциями температуры, то R также зависит от температуры. В качестве примера на рис. 6-9 представлена температурная зависимость постоянной Холла для образцов антимионид индия n - и p -типа. Как видно, в области низких температур концентрация электронов и дырок остается постоянной, что соответствует области истощения примеси. В области собственной электропроводности $\ln |R|$ линейно зависит от обратной температуры (кривые $1n$, $3n$ и $6n$). Так как у акцепторного полупроводника $R > 0$, то с ростом температуры R уменьшается, проходит через нуль и становится отрицательным. Температура, при которой $R = 0$, называется температурой инверсии. Из кривых рис. 6-9 для образцов $2p$, $4p$, $5p$ и $7p$ следует, что температура инверсии, так же как и температура перехода к собственной электропроводности тем выше, чем больше концентрация акцепторной примеси.

6-6. МАГНИТОРЕЗИСТИВНЫЙ ЭФФЕКТ

Искривление траектории носителей заряда в магнитном поле приводит к изменению составляющей тока, направленной вдоль вектора электрического поля, что эквивалентно изменению удельного сопротивления полупроводника. Это явление получило название эффекта магнитосопротивления (эффект Гаусса) или магниторезистивного эффекта.

Удельное сопротивление в отсутствие магнитного поля, когда направление тока \mathbf{j} совпадает с направлением вектора напряженности электрического поля \mathcal{E} , определяется выражением вида

$$\rho = \frac{1}{\sigma} = \frac{\mathcal{E}}{j} \quad (6-118)$$

При наличии магнитного поля, когда направление \mathcal{E} не совпадает с направлением \mathbf{j} , в это уравнение вместо величины \mathcal{E} следует подставить проекцию вектора \mathcal{E} на вектор \mathbf{j} :

$$\rho_B = \frac{\mathcal{E}_j}{j} = \frac{(\mathcal{E} \cdot \mathbf{j})}{j^2} \quad (6-119)$$

Выражение (6-103) для тока в слабых магнитных полях, полученное в линейном приближении относительно величины магнитной индукции B , с учетом равенств (6-107), (6-47) и (6-51) можно переписать в виде

$$\mathbf{j} = \sigma \mathcal{E} - R \sigma^2 [\mathbf{B} \mathcal{E}] \quad (6-120)$$

Оставаясь в рамках линейного приближения по магнитному полю и пренебрегая членами, содержащими B^2 , с учетом $\mathcal{E} [\mathbf{B}\mathcal{E}] = 0$ получаем:

$$j^2 = (\sigma\mathcal{E})^2. \quad (6-121)$$

Подстановка выражений (6-120) и (6-121) в (6-119) дает:

$$\rho_B = \frac{\sigma\mathcal{E}^2}{(\sigma\mathcal{E})^2} = \frac{1}{\sigma} = \rho. \quad (6-122)$$

Таким образом, в слабом магнитном поле в линейном приближении сопротивление полупроводника не меняется.

Чтобы определить величину удельного сопротивления полупроводника в магнитном поле ρ_B , следует использовать квадратичное приближение, решая уравнения (6-25) для однородного полупроводника ($\nabla F = 0$), находящегося в слабом магнитном поле $\frac{e\tau_e}{m_n} B \ll 1$ в отсутствие градиента температуры ($\nabla T = 0$). С учетом этого будем иметь:

$$\chi_n = -\frac{e\hbar\tau_e}{m_n^*} \left\{ \frac{\mathcal{E} + \frac{e\tau_e}{m_n^*} [\mathbf{B}\mathcal{E}] + \left(\frac{e\tau_e}{m_n^*}\right)^2 (\mathbf{B}\mathcal{E}) \mathbf{B}}{1 + \left(\frac{e\tau_e}{m_n^*} B\right)^2} \right\}. \quad (6-123)$$

Поскольку $\frac{e\tau_e}{m_n^*} B \ll 1$, то $\frac{1}{1 + \left(\frac{e\tau_e}{m_n^*} B\right)^2} \approx 1 - \left(\frac{e\tau_e}{m_n^*} B\right)^2$. В результате для вектора χ_n получим:

$$\begin{aligned} \chi_n &= -\frac{e\hbar\tau_e}{m_n^*} \left[\mathcal{E} + \frac{e\tau_e}{m_n^*} [\mathbf{B}\mathcal{E}] + \left(\frac{e\tau_e}{m_n^*}\right)^2 \{(\mathbf{B}\mathcal{E}) \mathbf{B} - B^2\mathcal{E}\} \right] = \\ &= -\frac{e\hbar\tau_e}{m_n^*} \left\{ \mathcal{E} + \frac{e\tau_e}{m_n^*} [\mathbf{B}\mathcal{E}] + \left(\frac{e\tau_e}{m_n^*}\right)^2 [\mathbf{B} [\mathbf{B}\mathcal{E}]] \right\}. \end{aligned} \quad (6-124)$$

Плотность тока для полупроводника с простой зонной структурой, содержащего один тип носителей заряда, например электроны, на основании выражения (6-32) запишется в виде

$$\begin{aligned} \mathbf{j} &= -\frac{e\hbar^2}{4\pi^3 m_n^{*2}} \int \frac{\partial f_0}{\partial E} \left\{ \tau_e(\mathbf{k}\mathcal{E}) + \left(\frac{e\tau_e}{m_n^*}\right)^2 \mathbf{k} [\mathbf{B}\mathcal{E}] + \right. \\ &\quad \left. + \frac{e^2\tau_e^3}{m_n^{*2}} \mathbf{k} [\mathbf{B} [\mathbf{B}\mathcal{E}]] \right\} \mathbf{k} d\tau_{\mathbf{k}}. \end{aligned} \quad (6-125)$$

Проводя интегрирование, находим с точностью до членов порядка B^2 :

$$\mathbf{j} = en\mu_n \left\{ \mathcal{E} + r\mu_n [\mathbf{B}\mathcal{E}] + \frac{3}{2} r\mu_n^2 [\mathbf{B} [\mathbf{B}\mathcal{E}]] \right\}. \quad (6-126)$$

Для решения полученного уравнения разложим вектор \mathcal{E} по некопланарным направлениям \mathbf{j} , \mathbf{B} и $[\mathbf{Bj}]$:

$$\mathcal{E} = \alpha \mathbf{j} + \beta \mathbf{B} + \gamma [\mathbf{Bj}]. \quad (6-127)$$

Коэффициенты разложения нетрудно найти, подставив выражение (6-126) в (6-127) и сравнив коэффициенты при векторах. В результате с точностью до B^2 получим:

$$\begin{aligned} \mathcal{E} = \frac{1}{en\mu_n} \left\{ \left[1 + \frac{3}{2} r \mu_n^2 \left(1 - \frac{\pi}{4} \right) B^2 \right] \mathbf{j} - \right. \\ \left. - \frac{3}{2} r \left(1 - \frac{\pi}{4} \right) \mu_n^2 (\mathbf{jB}) \mathbf{B} - r \mu_n [\mathbf{Bj}] \right\}, \end{aligned} \quad (6-128)$$

что при использовании формулы (6-119) приводит к

$$\frac{\rho_B - \rho}{\rho} = \frac{\Delta\rho}{\rho} = \frac{3}{2} r \left(1 - \frac{\pi}{4} \right) \mu_n^2 \left[B^2 - \frac{(\mathbf{jB})^2}{j^2} \right] = A \mu_n^2 B^2 \sin^2 \alpha, \quad (6-129)$$

где коэффициент $A = \frac{3}{2} r \left(1 - \frac{\pi}{4} \right)$ зависит от механизма рассеяния носителей заряда; α — угол между направлениями тока \mathbf{j} и магнитного поля \mathbf{B} :

$$(\mathbf{jB})^2 = j^2 B^2 \cos^2 \alpha. \quad (6-130)$$

Из выражения (6-129) следует, что если \mathbf{B} параллельно \mathbf{j} ($\alpha = 0$), то сопротивление полупроводника со сферически симметричной зонной структурой не изменяется, т. е. продольное магнитосопротивление равно нулю. Поперечное магнитосопротивление ($\mathbf{j} \perp \mathbf{B}$) в слабых полях пропорционально квадрату магнитной индукции и квадрату подвижности носителей заряда.

Чтобы найти магнитосопротивление при наличии носителей заряда обоих знаков, следует учесть дырочную составляющую тока, которая по аналогии с уравнением (6-126) равна:

$$\mathbf{j}_p = er\mu_p \left\{ \mathcal{E} - r\mu_p [\mathbf{B}\mathcal{E}] + \frac{3}{2} r\mu_p^2 [\mathbf{B}[\mathbf{B}\mathcal{E}]] \right\}. \quad (6-131)$$

Уравнение для полного тока тогда имеет вид:

$$\begin{aligned} \mathbf{j} = \mathbf{j}_n + \mathbf{j}_p = \sigma \left\{ \mathcal{E} + r \frac{n\mu_n^2 - p\mu_p^2}{n\mu_n + p\mu_p} [\mathbf{B}\mathcal{E}] + \right. \\ \left. + \frac{3}{2} r \frac{n\mu_n^3 + p\mu_p^3}{n\mu_n + p\mu_p} [\mathbf{B}[\mathbf{B}\mathcal{E}]] \right\} \end{aligned} \quad (6-132)$$

и решается подобно уравнению (6-126). В результате для полупроводника с простой зонной структурой, имеющего два типа носителей заряда, получим выражение

$$\frac{\Delta\rho}{\rho} = \frac{3}{2} r \frac{e}{\sigma} \left[(n\mu_n^3 + p\mu_p^3) - \frac{\pi}{4} \frac{(n\mu_n^2 - p\mu_p^2)^2}{n\mu_n + p\mu_p} \right] B^2 \sin^2 \alpha, \quad (6-133)$$

которое в случае $n = 0$ (или $p = 0$) сводится к формуле (6-129) для полупроводника с одним типом носителей заряда.

В материалах, не обладающих сферической симметрией зон, изменение сопротивления в магнитном поле зависит от ориентации вектора \mathbf{B} относительно кристаллографических осей. В общем случае наблюдается также эффект продольного магнитосопротивления.

Поворачивая образец в магнитном поле и измеряя $\Delta\rho/\rho$, можно получить важные сведения о параметрах зонной структуры данного материала.

Изменение сопротивления полупроводниковой пластины в магнитном поле существенно зависит от ее конфигурации (отношения длины к ширине). Возникающее в результате эффекта Холла поперечное электрическое поле препятствует изменению траектории носителей заряда под действием магнитного поля и снижает таким образом эффект магнитосопротивления. В частности, из формулы (6-133) следует, что отношение $\Delta\rho/\rho$ максимально в том случае, когда второе слагаемое обращается в нуль, что в соответствии с (6-107) имеет место при $R = 0$, т. е. в отсутствие поля Холла. Учитывая это свойство, для достижения наибольшего изменения сопротивления в магнитном поле следует использовать короткие, широкие образцы с омическими контактами на длинных сторонах, шунтирующими э. д. с. Холла. Оптимальной конструкцией является диск Корбино, у которого один контакт расположен в центре круговой пластины, а другой — по ее окружности. В таком образце ток протекает в радиальных направлениях, а э. д. с. Холла полностью отсутствует. На рис. 6-10 проиллюстрировано изменение сопротивления в магнитном поле для прямоугольных пластин антимионида индия различной конфигурации, а также для диска Корбино.

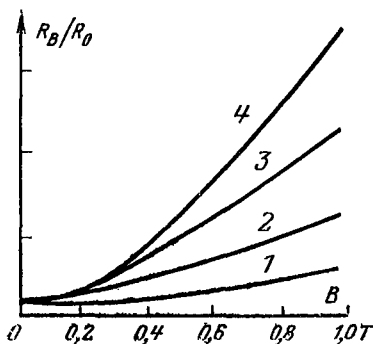


Рис. 6-10. Зависимость магнитосопротивления от величины отношения длины образца антимионида индия к ширине.
1 — 2; 2 — 1; 3 — 0,5; 4 — 0 (диск Корбино).

Рассмотрим теперь случай сильного магнитного поля, когда $\frac{e\tau_e}{m_n^*} B \gg 1$. Разлагая векторную функцию χ_n (6-25) в ряд по параметру $m_n^*/e\tau_e B$, для однородного полупроводника и в отсутствие градиента температуры $\nabla_r f = \frac{E-F}{T} \nabla T - \nabla F = 0$, будем иметь:

$$\begin{aligned} \chi_n &= -\frac{eh\tau_e}{m_n^*} \left\{ \frac{(\mathbf{B}\mathcal{E}) \mathbf{B}}{B^2} + \frac{m_n^*}{e\tau_e B^2} [B\mathcal{E}] + \right. \\ &\quad \left. + \left(\mathcal{E} - \frac{(\mathbf{B}\mathcal{E}) \mathbf{B}}{B^2} \right) \left(\frac{m_n^*}{e\tau_e B} \right)^2 \right\} = \\ &= -\frac{eh\tau_e}{m_n^*} \left\{ (\mathbf{e}_B \mathcal{E}) \mathbf{e}_B + \frac{m_n^*}{e\tau_e B} [\mathbf{e}_B \mathcal{E}] + \left(\mathcal{E} - (\mathbf{e}_B \mathcal{E}) \mathbf{e}_B \right) \left(\frac{m_n^*}{e\tau_e B} \right)^2 \right\} = \\ &= -\frac{eh\tau_e}{m_n^*} \mathcal{E}_B - \frac{h}{B} [\mathbf{e}_B \mathcal{E}] - \frac{hm_n^*}{e\tau_e B^2} \mathcal{E}_\perp, \end{aligned} \quad (6-134)$$

где $\mathbf{e}_B = \mathbf{B}/B$ — единичный вектор в направлении магнитного поля; $\mathcal{E}_B = (\mathbf{e}_B \mathcal{E}) \mathbf{e}_B$ — составляющая напряженности электрического поля в направлении магнитного поля; $\mathcal{E}_\perp = (\mathcal{E} - (\mathbf{e}_B \mathcal{E}) \mathbf{e}_B) = \mathcal{E} - \mathcal{E}_B$ — составляющая электрического поля, перпендикулярная к магнитному полю. Решая уравнение для плотности тока, получаем:

$$\mathbf{j} = en\mu_n \mathcal{E}_B + \frac{en}{B} [\mathbf{e}_B \mathcal{E}] + \frac{4}{3r} \frac{en}{\mu_n B^2} \mathcal{E}_\perp, \quad (6-135)$$

Пусть ток \mathbf{j} направлен по оси x , а магнитное поле — по оси z . Тогда

$$\left. \begin{aligned} i_x &= -\frac{en}{B} \mathcal{E}_y + \frac{4}{3r} \frac{en}{\mu_n B^2} \mathcal{E}_x = j; \\ j_y &= \frac{en}{B} \mathcal{E}_x + \frac{4}{3r} \frac{en}{\mu_n B^2} \mathcal{E}_y = 0. \end{aligned} \right\} \quad (6-136)$$

Решая эту систему уравнений с точностью до $1/B^2$ относительно \mathcal{E}_x и \mathcal{E}_y , будем иметь:

$$\left. \begin{aligned} \mathcal{E}_x &= \frac{4}{3r} \frac{1}{en\mu_n} j; \\ \mathcal{E}_y &= -\frac{1}{en} jB. \end{aligned} \right\} \quad (6-137)$$

Отсюда найдем выражения для постоянной Холла R и магнитосопротивления:

$$R = -\frac{1}{en}; \quad (6-138)$$

$$\rho_B = \frac{(\mathcal{E})}{j^2} \frac{4}{3r} \frac{1}{en\mu_n} = \frac{4}{3r} \rho. \quad (6-139)$$

В случае рассеяния на тепловых колебаниях решетки $r = 3\pi/8$, поэтому

$$\frac{\Delta\rho}{\rho} = \frac{4}{3r} - 1 = \frac{32}{9\pi} - 1 = 0,13. \quad (6-140)$$

Таким образом, в сильном магнитном поле постоянная Холла не зависит от механизма рассеяния, а для магнитосопротивления наступает насыщение. При этом величина магнитосопротивления при насыщении зависит от механизма рассеяния.

В случае полупроводника со смешанной проводимостью плотность тока будет равна:

$$\begin{aligned} \mathbf{j} = & e(n\mu_n + p\mu_p) \mathcal{E}_B + \frac{e}{B} (n - p) [\mathbf{e}_B \mathcal{E}] + \\ & + \frac{4}{3r} \frac{e}{B^2} \left(\frac{n}{\mu_n} + \frac{p}{\mu_p} \right). \end{aligned} \quad (6-141)$$

На основании этого выражения, находя $j_x = j$ и $j_y = 0$, а затем определяя \mathcal{E}_x и \mathcal{E}_y , получаем:

$$R = \frac{1}{e(p - n)}; \quad (6-142)$$

$$\rho_B = \frac{4}{3r} \frac{\frac{n}{\mu_n} + \frac{p}{\mu_p}}{e(p - n)}. \quad (6-143)$$

6-7. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ

К термоэлектрическим явлениям относятся три эффекта: 1) эффект Зеебека, 2) эффект Пельтье и 3) эффект Томсона. Рассмотрим эти эффекты.

1. Эффект Зеебека

Эффект Зеебека, или термоэлектрический эффект, заключается в том, что в замкнутой цепи, состоящей из двух различных полупроводников, места соединений которых находятся при разных температурах (рис. 6-11), возникает электрический ток, носящий на-

звание термоэлектрического. На концах такой разомкнутой цепи появляется разность потенциалов \mathcal{E} , которая носит название термоэ. д. с. (термо-э. д. с.). Величина этой разности потенциалов, зависящая от разности температур и вида материала, характеризуется коэффициентом α :

$$\alpha = \frac{d\mathcal{E}}{dT}. \quad (6-144)$$

Коэффициент α называется удельной термо-э. д. с. и представляет собой термо-э. д. с., отнесенную к единичной разности температур.

Если полупроводник нагрет неравномерно, то средняя энергия носителей заряда и их концентрация будут больше в той его области, где выше температура. Следовательно, градиент температуры в однородном полупроводнике приводит к градиенту средней энергии носителей заряда и градиенту их концентрации, вследствие чего возникает диффузионный поток носителей заряда и появляется электрический ток. В разомкнутой цепи в стационарном состоянии плотность тока в любой точке образца равна нулю. Это означает, что электрический ток, обусловленный градиентом температуры, компенсируется током, возникающим в электрическом поле при разделении зарядов. На образце возникает термоэлектродвижущая сила.

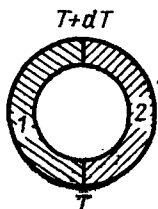


Рис. 6-11. Замкнутая цепь двух проводников.

Вычисление термо-э. д. с. проведем с помощью кинетического уравнения для однородного невырожденного полупроводника, имеющего сферические изоэнергетические поверхности для электронов $\left(E = \frac{\hbar k^2}{2m_n^*}\right)$ и дырок $\left(E' = -\frac{\hbar k'^2}{2m_p^*}\right)$.

Если химический потенциал для электронов F , то для дырок он будет $F' = -E_g - F$, где E_g — ширина запрещенной зоны. Поскольку полупроводник невырожденный, то равновесная функция распределения для электронов $f_0 = e^{-(E-F)/kT}$ и дырок $f_{0p} = e^{(E'-F')/kT} = e^{(E'+E_g+F)/kT}$, а их концентрации соответственно равны: $n = N_c e^{-(E_c-F)/kT} = N_c e^{F/kT}$ и $p = N_v e^{-(F'+E_v)/kT} = N_v e^{-F'/kT}$, так как $E_v = 0$. В этом случае, с учетом того, что $\mathbf{V} = 0$ (нет магнитного поля), выражения для χ_n (6-25) и χ_p (6-26) будут иметь вид:

$$\chi_n = \frac{\hbar \tau_e}{m_n^*} \left[\frac{F-E}{T} \nabla T - \nabla (F - e\varphi) \right]; \quad (6-145)$$

$$\chi_p = -\frac{\hbar \tau_h}{m_p^*} \left[\frac{E'+E_g+F}{T} \nabla T - \nabla (F - e\varphi) \right]. \quad (6-146)$$

Вычисление плотности тока проведем для полупроводника в предположении, что рассеяние носителей заряда осуществляется

на акустических колебаниях кристаллической решетки. В этом случае длина свободного пробега $l = vt$ не зависит от энергии носителей заряда (см. § 5-7) и время релаксации можно выразить через l , т. е.

$$\tau = \frac{l}{v} = \frac{m_n^* l}{\hbar} k^{-1}. \quad (6-147)$$

Электронную составляющую тока (6-32) с учетом (6-145), (6-147) и (6-40) можно представить в виде

$$\mathbf{j}_n = -\frac{e\hbar l}{3\pi^2 m_n^*} \left[\left\{ \frac{F}{T} \nabla T - \nabla(F - e\varphi) \right\} \int_0^\infty \frac{\partial f_0}{\partial E} k^3 dk - \Delta T \int_0^\infty \frac{\partial f_0}{\partial E} E k^3 dk \right]. \quad (6-148)$$

Если ввести $E = \alpha kT$, то $k^3 dk = \frac{(2m_n^* kT)^2}{2\hbar^4} \alpha d\alpha$. При этом, учитывая (6-6) и (6-145), получаем:

$$\mathbf{j}_n = -\frac{4enl}{3(2\pi m_n^* kT)^{1/2}} \left\{ \frac{F}{T} \nabla T - \nabla(F - e\varphi) - 2k\Delta T \right\}. \quad (6-149)$$

Используя подвижность носителей заряда (6-56), выраженную через длину свободного пробега, формулу (6-149) можно представить в виде

$$\mathbf{j}_n = n\mu_n \left\{ \nabla(F - e\varphi) + \left(2k - \frac{F}{T} \right) \nabla T \right\}. \quad (6-150)$$

Проведя расчет для дырочной составляющей тока, найдем:

$$\mathbf{j}_p = p\mu_p \left\{ \nabla(F - e\varphi) - \left(2k - \frac{F + E_g}{T} \right) \nabla T \right\}. \quad (6-151)$$

Следовательно, полная плотность тока

$$\mathbf{j} = n\mu_n \left\{ \nabla(F - e\varphi) + \left(2k - \frac{F}{T} \right) \nabla T \right\} + p\mu_p \left\{ \nabla(F - e\varphi) - \left(2k - \frac{F + E_g}{T} \right) \nabla T \right\}. \quad (6-152)$$

Из соотношения (6-152) следует, что в полупроводнике при наличии градиента температуры возникают электронный и дырочный токи, которые обусловлены действием градиента электрохимического потенциала $\nabla(F - e\varphi)$ и градиента температуры ∇T . Для нахождения термо-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 (6-152), приравняв ток нулю. Но при измерении термо-э. д. с. на границах полупроводника и измерительных металлических электродов существуют контактные разности потенциалов (см. § 9-4), которые не равны друг другу вследствие существующего температурного градиента. Измерительный прибор отметит 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 $\nabla(F - e\varphi)$. Абсолютная величина $|\nabla F/e|$ будет равна разности контактных потенциалов в граничных точках, если считать, что полупроводник и металл на контакте находятся в термодинамическом равновесии. Поэтому удельная термо-э. д. с. α определяется как

$$\alpha = \frac{|\nabla(\varphi - F/e)|}{|\nabla T|}. \quad (6-153)$$

Полагая ток в соотношении (6-152) равным нулю, получаем для удельной термо-э. д. с. полупроводника:

$$\alpha = -\frac{k}{e} \frac{1}{n\mu_n + p\mu_p} \left\{ n\mu_n \left(2 - \frac{F}{kT} \right) - p\mu_p \left(2 - \frac{F + E_g}{kT} \right) \right\}. \quad (6-154)$$

Если воспользоваться формулами для концентрации электронов и дырок невырожденного полупроводника, то выражение (6-154) можно записать в виде

$$\alpha = -\frac{k}{e} \left\{ \frac{\left[2 + \ln \frac{N_c}{n} \right] n\mu_n - \left[2 + \ln \frac{N_v}{p} \right] p\mu_p}{n\mu_n + p\mu_p} \right\}, \quad (6-155)$$

где N_c , N_v — эффективная плотность состояний в зоне проводимости и в валентной зоне.

Таким образом, термо-э. д. с. полупроводника определяется двумя слагаемыми, каждое из которых соответствует вкладу, вносимому электронами и дырками, причем эти слагаемые имеют противоположные знаки.

Из выражения (6-155) следует, что термо-э. д. с. отрицательна для электронного полупроводника. В этом случае на горячем торце образца возникает положительный объемный заряд, поскольку электроны диффундируют от горячего торца к холодному. В дырочном полупроводнике термо-э. д. с. положительна, в нем, наоборот, на горячем торце возникает отрицательный объемный заряд. Таким образом, если полупроводник примесный, то направление внутреннего электрического поля и полярность термо-э. д. с. определяются знаком основных носителей заряда и, следовательно, по знаку термо-э. д. с. можно определить тип примесной проводимости.

мости исследуемого образца. В случае смешанной проводимости знак термо-э. д. с. определяется не только соотношением концентраций носителей заряда, но и их подвижностью. В свою очередь, величина термо-э. д. с. примесных полупроводников уменьшается с ростом содержания примеси. При нагревании примесного образца его термо-э. д. с. также уменьшается.

Для собственного полупроводника $n = p = n_i$, $F = -\frac{E_g}{2}$ и формула (6-154) примет вид:

$$\alpha = -\frac{k}{e} \frac{b-1}{b+1} (2 + E_g/2kT), \quad (6-156)$$

где $b = \mu_n/\mu_p$.

Таким образом, величина термо-э. д. с. собственного полупроводника определяется лишь шириной запрещенной зоны и отношением подвижности электронов и дырок.

Изменение механизма рассеяния в полупроводнике, которое может иметь место при различных температурах, влияет на термо-э. д. с. через подвижность и сказывается на величине первого слагаемого (в квадратных скобках) выражения (6-155). Например, его 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 (6-155) нельзя объяснить большое возрастание термо-э. д. с., которое обнаруживается иногда в чистом веществе при очень низких температурах. Это явление определяется особым видом электрон-фононного взаимодействия, которое получило название эффекта фононного увлечения.

При наличии градиента температуры в полупроводнике имеет место анизотропия в распространении фононов, так как градиент концентрации фононов вызывает их движение в направлении от горячего к холодному краю образца. При каждом акте электрон-фононного рассеяния, которое сопровождается поглощением фонона, электрон приобретает энергию и квазиимпульс фонона. Поскольку фононов с квазиимпульсом, направленным от горячего к холодному краю образца, больше, чем с противоположно направленным квазиимпульсом, то электроны, поглощая фононный квазиимпульс, будут увлекаться фононным потоком. Вследствие этого электроны станут дрейфовать к холодной части образца. Дрейф электронов будет продолжаться до тех пор, пока в связи с их перераспределением не возникнет электрическое поле, которое уравновесит силу, действующую на электроны со стороны фононного потока. Эта разность потенциалов и представляет собой дополнительную термо-э. д. с., которая добавляется к термо-э. д. с., рассчитанной для случая отсутствия электрон-фононного взаимодействия. Эффект фононного увлечения приводит к значительному росту термо-э. д. с. только при низких температурах, так как из-за вза-

имных столкновений распределение фононов медленно возвращается к равновесному. При более высоких температурах рассеяния 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. 6-12).

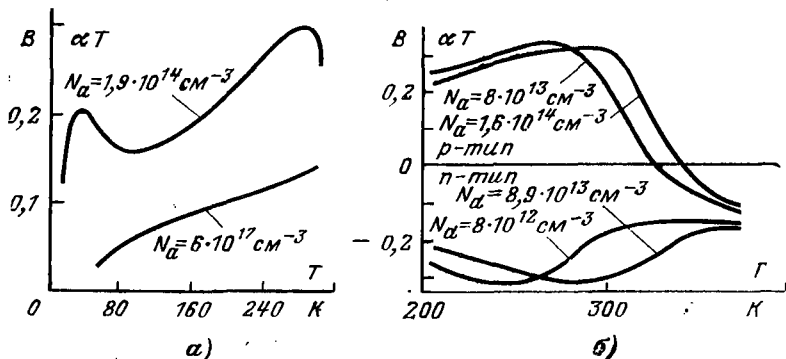


Рис. 6-12. Термо-э. д. с. германия при низких (а) и высоких (б) температурах.

У металлов концентрация электронов не зависит от температуры, а их средняя энергия*:

$$\langle E \rangle = \frac{5}{3} F_0 \left[1 + \frac{5\pi^2}{12} \left(\frac{kT}{F_0} \right)^2 \right], \quad (6-157)$$

где F_0 — уровень Ферми при температуре абсолютного нуля.

Так как для металлов $F_0 \approx 5$ эВ, то второй член в выражении (6-157) 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/К, а у полупроводников она достигает значения 100—1000 мкВ/К. 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, для значений термо-э. д. с. полупроводников не указывают, относительно какого металла она определена.

2. Эффект Пельтье

Эффект Пельтье заключается в том, что при прохождении электрического тока через контакт двух веществ в нем в зависимости от направления тока выделяется или поглощается тепло. Количе-

* Зейтц Ф. Современная теория твердого тела. М., Гостехиздат, 1949, гл. 4, § 26.

ство тепла при эффекте Пельтье пропорционально плотности тока и времени:

$$Q_{\Pi} = \Pi j t, \quad (6-158)$$

где Π — коэффициент Пельтье.

3. Эффект Томсона

Если вдоль однородного проводника существует градиент температуры, то при прохождении тока в объеме материала кроме тепла Джоуля выделяется (или поглощается) тепло, количество которого пропорционально количеству протекающего электричества и перепаду температур:

$$Q_T = I t \int_T^{T'} \tau_T dT. \quad (6-159)$$

Коэффициент τ_T называется коэффициентом Томсона.

Коэффициенты α , Π и τ_T связаны между собой соотношением Томсона. Для их вывода проведем элементарное рассмотрение термоэлектрических эффектов для замкнутой цепи из двух проводников, как это изображено на рис. 6-11. Пусть между контактами полупроводников существует разность температур dT . Наличие градиента температуры вызывает термоэлектрическую силу, равную $d\mathcal{E} = \alpha dT$, которая обеспечит в данной цепи прохождение тока dI . При этом на одном из контактов в единицу времени выделяется тепло Пельтье в количестве $\Pi' dI$, а на другом поглощается ΠdI . В результате в цепи выделяется тепло Пельтье, и если Π' мало отличается от Π , то количество этого тепла

$$dQ_{\Pi} = (\Pi' - \Pi) dI = \frac{d\Pi}{dT} dI dT. \quad (6-160)$$

Естественно, наличие электрического тока и градиента температуры вызовет появление тепла за счет эффекта Томсона, равного за 1 с

$$dQ_T = (\tau_{T_2} - \tau_{T_1}) dI dT. \quad (6-161)$$

Эти термодинамические явления можно считать обратимыми, поскольку они зависят от направления тока и градиента температуры. Но в данной замкнутой цепи имеют место и необратимые процессы — теплопроводность и выделение тепла Джоуля. В стационарных условиях теплопроводность не изменяет общего количества тепла в системе, имеет место только перенос тепла в неизменном количестве 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 имеем:

$$\alpha dI dT = \frac{d\Pi}{dT} dI dT + (\tau_{T_2} - \tau_{T_1}) dI dT, \quad (6-162)$$

откуда находим:

$$\alpha = \frac{d\Pi}{dT} + (\tau_{T_2} - \tau_{T_1}). \quad (6-163)$$

Если положить для горячего конца T' , Π' , τ_{T_2} , а для холодного T , Π , τ_{T_1} , то на основании второго закона термодинамики, по которому полное изменение энтропии dS замкнутой системы при обратимых процессах равно нулю, имеем:

$$\frac{\Pi' dI}{T'} - \frac{\Pi dI}{T} + dI \int_T^{T'} \frac{\tau_{T_2} - \tau_{T_1}}{T} dT = 0. \quad (6-164)$$

Если за начало отсчета энтропии взять ее значение на горячем конце ($\frac{\Pi' dI}{T'} = 0$), то, продифференцировав выражение (6-164) по T , получим:

$$\frac{d}{dT} \left(\frac{\Pi}{T} \right) + \frac{\tau_{T_2} - \tau_{T_1}}{T} = 0, \quad (6-165)$$

откуда найдем:

$$\frac{d\Pi}{dT} + (\tau_{T_2} - \tau_{T_1}) = \frac{\Pi}{T}. \quad (6-166)$$

Сравнивая (6-163) и (6-166), видим, что

$$\Pi = \alpha T. \quad (6-167)$$

Если продифференцируем равенство (6-167) по T , найдем:

$$\frac{d\alpha}{dT} = \frac{1}{T} \frac{d\Pi}{dT} - \frac{\Pi}{T^2} = \frac{1}{T} \left(\frac{d\Pi}{dT} - \frac{\Pi}{T} \right). \quad (6-168)$$

Учитывая выражение (6-166), получаем:

$$\tau_{T_1} - \tau_{T_2} = \tau_T = T \frac{d\alpha}{dT}. \quad (6-169)$$

Таким образом, коэффициенты α , Π , τ_T , характеризующие термоэлектрические свойства данного вещества, связаны между собой термодинамическими соотношениями. Обычно проводится экспериментальное измерение α , а трудно измеряемые коэффициенты Π и τ_T определяются на основе этих соотношений,

6-8. ТЕПЛОПРОВОДНОСТЬ ПОЛУПРОВОДНИКОВ

Как известно, все тела проводят тепло. Это явление носит название *теплопроводности*. В том случае, если в веществе нет тепловых процессов, приводящих к появлению тока, т. е. $\mathbf{j} = 0$, то имеющийся в веществе градиент температуры ∇T вызовет поток энергии \mathbf{W} в направлении, противоположном ∇T :

$$\mathbf{W} = -\kappa \nabla T. \quad (6-170)$$

Здесь $\kappa = \frac{|\mathbf{W}|}{|\nabla T|}$ есть коэффициент теплопроводности, который численно равен количеству энергии, проходящей в единицу времени через единичное сечение образца, на концах которого создана разность температуры в один градус.

Перенос тепловой энергии в твердом теле осуществляется свободными носителями заряда и фононами. Теплопроводность, обусловленная движением носителей заряда, называется *электронной* (или *дырочной*). Она характеризуется коэффициентом теплопроводности κ_e , величина которой в сильной степени зависит от свойств полупроводника и температуры. Теплопроводность, обусловленная колебаниями решетки, называется *решеточной*, или *фононной*, теплопроводностью. Она обозначается κ_L . Таким образом, полная теплопроводность твердого тела складывается из решеточной κ_L и электронной (или дырочной) κ_e теплопроводности, т. е.

$$\kappa = \kappa_L + \kappa_e. \quad (6-171)$$

Теплопроводность диэлектриков в основном является фононной теплопроводностью; у них $\kappa_L \gg \kappa_e$. В металлах высокой чистоты вклад решеточной теплопроводности обычно мал, поэтому $\kappa_e \gg \kappa_L$. В полупроводниках нередко решеточная теплопроводность такого же порядка, что и электронная (или дырочная).

Решеточная теплопроводность определяется потоком фононов, движущихся от горячего конца образца к холодному. Если для фононов ввести длину свободного пробега l_L , то коэффициент фононной теплопроводности может быть определен на основе кинетической теории. При этом выражение для κ_L имеет вид:

$$\kappa_L = \frac{1}{3} C v_{зв} l_L, \quad (6-172)$$

где C — теплоемкость решетки; $v_{зв}$ — фононная скорость (скорость звука).

В линейном образце длиной L , перепад температуры между концами которого равен $T_1 - T_2$, в единицу времени через единичную площадку проходит тепло в количестве

$$Q = \frac{\kappa L}{L} (T_1 - T_2). \quad (6-173)$$

Воспользовавшись для κ_L соотношением (6-172), получим:

$$Q = \frac{1}{3} C (T_1 - T_2) \frac{l_L v_{зв}}{L}. \quad (6-174)$$

Определение теплопроводности κ_e , обусловленной движением носителей заряда, проведем для невырожденного донорного полупроводника, в котором рассеяние носителей заряда осуществляется на акустических колебаниях. Для этого вычислим плотность потока энергии, переносимого электронами,

$$\begin{aligned} \mathbf{W} &= \frac{1}{4\pi^3} \int_{(V_B)} f_1^{(n)}(\mathbf{k}) (E - e\varphi) \mathbf{v}_n d\tau_{\mathbf{k}} = \\ &= - \frac{e\hbar}{4\pi^3 m_n^*} \int_{(V_B)} \frac{\partial f_0}{\partial E} (\chi_n \mathbf{k}) (E - e\varphi) \mathbf{k} d\tau_{\mathbf{k}} \end{aligned} \quad (6-175)$$

в предположении, что

$$\mathbf{j}_n = 0. \quad (6-176)$$

Решая уравнение (6-175) с учетом значения χ_n (6-25) для случая $\mathbf{V} = 0$ и $\mathbf{j}_n = 0$, получаем:

$$\mathbf{W} = -n\mu_n \frac{2k^2 T}{e} \nabla T. \quad (6-177)$$

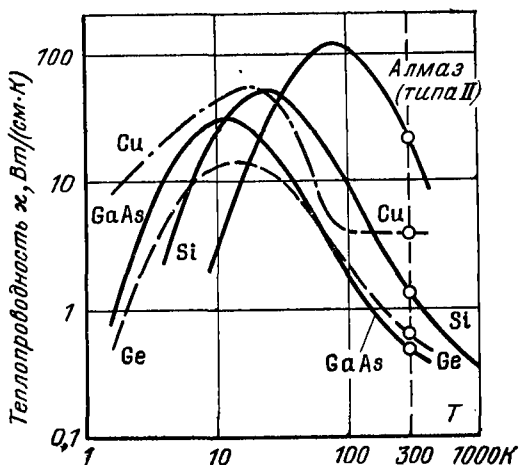


Рис. 6-13. Температурная зависимость теплопроводности для чистых германия, кремния, арсенида галлия, меди и алмаза (типа II).

В этом выражении множитель, стоящий перед ∇T , и есть коэффициент теплопроводности κ_e , равный:

$$\kappa_e = \frac{2k^2T}{e} n\mu_n = 2 \left(\frac{k}{e} \right)^2 \sigma T. \quad (6-178)$$

Следовательно, *теплопроводность вещества, обусловленная движением свободных носителей заряда, определяется температурой и его удельной проводимостью.*

Из равенства (6-178) найдем отношение

$$\frac{\kappa_e}{\sigma} = 2 \left(\frac{k}{e} \right)^2 T, \quad (6-179)$$

которое выражает закон Видемана — Франца.

Проведя подобные вычисления теплопроводности для невырожденного полупроводника с двумя типами носителя заряда, получим формулу

$$\kappa_e = \frac{2k^2T}{e} n\mu_n + \frac{2k^2T}{e} p\mu_p + \frac{n\mu_n p\mu_p (E_g + 4kT)^2}{eT (n\mu_n + p\mu_p)}. \quad (6-180)$$

Установлено, что теплопроводность твердых тел, как видно из рис. 6-13, сначала растет с ростом температуры, а затем убывает при высоких температурах.

6.9. ЭЛЕКТРОПРОВОДНОСТЬ ПОЛУПРОВОДНИКОВ В СИЛЬНОМ ЭЛЕКТРИЧЕСКОМ ПОЛЕ

При выводе кинетического уравнения мы предполагали, что время релаксации τ не зависит от электрического поля. Затем при помощи кинетического уравнения для невырожденного полупроводника со сферическими изоэнергетическими поверхностями было показано, что плотность тока подчиняется закону Ома, т. е. пропорциональна напряженности электрического поля:

$$\mathbf{j} = en\mu_0 \mathcal{E} = \sigma_0 \mathcal{E}, \quad (6-181)$$

где μ_0 , σ_0 — подвижность и проводимость при слабом поле.

В электрическом поле напряженности \mathcal{E} электрон на длине свободного пробега приобретает энергию, равную $e\mathcal{E}l$. При столкновении с акустическим фононом электрон передает эту энергию решетке, приходя в тепловое равновесие с ней. Но при столкновении с фононом электрон может либо передать энергию фонону, либо получить энергию от фонона. Пусть при этом изменение энергии электрона будет ΔE .

Если из общего числа столкновений подсчитать число столкновений, идущих с передачей энергии фонону, то согласно (5-99) на одно такое столкновение приходится $(2N_q + 1)$ столкновений.

Поэтому относительная потеря энергии электрона при столкновении составляет величину

$$\delta = \frac{\Delta E}{kT} = \frac{\hbar\omega_q}{kT} \frac{1}{2N_q + 1} = \frac{\hbar\omega_q}{kT} \operatorname{th} \frac{\hbar\omega_q}{2kT}. \quad (6-182)$$

Следовательно, изменение энергии электрона при столкновении с фононом равно $\Delta E = kT\delta$. Соударение будет упругим в том случае, если энергия, набираемая электроном в электрическом поле на длине свободного пробега $e\mathcal{E}l$, будет значительно меньше энергии ΔE , теряемой в среднем при одном столкновении с решеткой:

$$\frac{e\mathcal{E}l}{\Delta E} \ll 1.$$

Это означает, что закон Ома выполняется в том случае, когда

$$e\mathcal{E}l \ll kT\delta, \quad (6-183)$$

или

$$\mathcal{E} \ll \delta \frac{kT}{el}. \quad (6-184)$$

При этом условии распределение скоростей носителей заряда остается почти максвелловским с температурой, равной температуре решетки. В этом случае на хаотическое движение носителей заряда накладывается медленный их дрейф, обусловленный действием слабого внешнего электрического поля. При этом скорость дрейфа носителей заряда v значительно меньше их тепловой скорости v_0 , т. е.

$$v \ll v_0. \quad (6-185)$$

В сильных электрических полях имеет место нарушение закона Ома. Это может происходить по двум причинам. Во-первых, в сильных электрических полях скорость дрейфа носителей заряда становится соизмеримой с их тепловой скоростью. Это скажется на процессе рассеяния, что проявляется в изменении времени релаксации τ , а следовательно, и подвижности μ . Сюда необходимо отнести явление «разогрева» электронно-дырочного газа и эффект Ганна. Во-вторых, сильные электрические поля могут вызвать изменение концентрации носителей заряда. При этом рост концентрации носителей заряда может обуславливаться рядом причин. Такими эффектами являются ударная ионизация, эффект Зинера и электростатическая ионизация.

Рассмотрим электропроводность полупроводника в сильном электрическом поле. В сильных электрических полях, для которых не выполняется условие (6-185), носители заряда при столкновении с акустическими фононами не успевают отдавать набираемую ими в поле энергию. В результате этого средняя энергия повышается

и электронно-дырочный газ «разогревается». Но с ростом энергии носителей заряда увеличивается их скорость, а значит уменьшается время свободного пробега и, следовательно, увеличивается вероятность рассеяния носителей заряда на тепловых колебаниях решетки. В силу этого рост энергии носителей заряда с возрастанием поля будет продолжаться до тех пор, пока электрон-фононное взаимодействие не усилится настолько, что установится стационарное состояние, при котором энергия, приобретаемая носителем заряда в поле за единицу времени $e\mathcal{E}v$, сравнивается с энергией, теряемой этим носителем заряда за одну секунду в результате столкновения $\Delta E/\tau = kT\delta/\tau$, т. е.

$$e\mathcal{E}v = \delta \frac{kT}{\tau}. \quad (6-186)$$

При рассеянии на тепловых колебаниях решетки согласно (5-93) длина свободного пробега l не зависит от энергии, поэтому время свободного пробега $\tau = l/v_0$. Но дрейфовая скорость носителей заряда на основании (6-186) равна:

$$v = \frac{\delta kT}{e\mathcal{E}\tau} = \frac{\delta kT v_0}{e\mathcal{E}l}. \quad (6-187)$$

В то же время дрейфовая скорость

$$v = \frac{e\tau}{m^*} \mathcal{E} = \frac{el}{m^* v_0} \mathcal{E}. \quad (6-188)$$

Из сравнения формул (6-187) и (6-188) следует, что

$$(e\mathcal{E}l)^2 = 2\delta kT \frac{m^* v_0^2}{2}. \quad (6-189)$$

Но так как $\frac{1}{2} m^* v_0^2 = kT$, то из соотношения (6-189) найдем критическую напряженность поля

$$\mathcal{E}_{\text{кр}} = \frac{kT}{el} (2\delta)^{1/2}, \quad (6-190)$$

при которой начинается разогрев электронно-дырочного газа. При этом, как следует из (6-190), величина критического поля уменьшается при понижении температуры. Это согласуется с экспериментальными данными, которые представлены на рис. 6-14.

Если воспользоваться равенствами (6-187) и (6-190), получим:

$$v = \left(\frac{\delta}{2}\right)^{1/2} v_0. \quad (6-191)$$

Следовательно, разогрев электронно-дырочного газа имеет место, когда дополнительная скорость, приобретенная носителем заряда в электрическом поле, становится сравнимой с его тепловой скоростью.

Из (6-189) можно определить энергию носителей заряда, выраженную через тепловую скорость

$$\left(\frac{m^* v_0^2}{2} \right)^2 = \frac{(e\mathcal{E}l)^2}{2\delta}, \quad (6-192)$$

откуда

$$v_0 = \left[\frac{2e\mathcal{E}l}{m^* (2\delta)^{1/2}} \right]^{1/2}. \quad (6-193)$$

Подставляя (6-193) в выражение (6-191), получаем:

$$v = \left(\frac{\delta}{2} \right)^{1/2} \left[\frac{2e\mathcal{E}l}{m^* (2\delta)^{1/2}} \right]^{1/2} = \left[\frac{el}{m^*} \left(\frac{\delta}{2} \right)^{1/2} \right]^{1/2} \mathcal{E}^{1/2}. \quad (6-194)$$

Таким образом, в сильных электрических полях, когда напряженность электрического поля больше критического, т. е.

$$\mathcal{E} > \mathcal{E}_{кр} \quad \text{или} \quad \mathcal{E} > \frac{kT}{el} (2\delta)^{1/2}, \quad (6-195)$$

дрейфовая скорость движения носителей заряда будет увеличиваться как

$$v \sim \mathcal{E}^{1/2}, \quad (6-196)$$

а время релаксации соответственно будет уменьшаться:

$$\tau = \frac{l}{v_0 + v} \sim \mathcal{E}^{-1/2}. \quad (6-197)$$

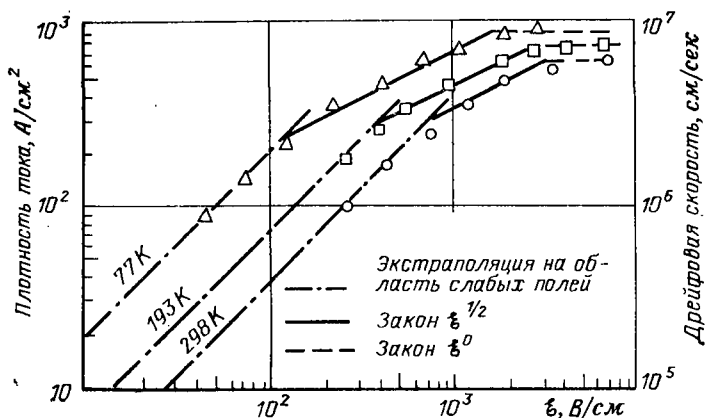


Рис. 6-14. Зависимость плотности тока от напряженности электрического поля в электронном германии.

Отсюда следует, что

$$\mu = \frac{e\tau}{m^*} = \mu_0 \mathcal{E}^{-1/2}, \quad (6-198)$$

а

$$\sigma = en\mu_n = \sigma_0 \mathcal{E}^{-1/2} \text{ и } j = \sigma \mathcal{E} = \delta_0 / \mathcal{E}^{1/2}. \quad (6-199)$$

Следовательно, при электрических полях $\mathcal{E} > \mathcal{E}_{кр}$ закон Ома не выполняется. Например, в германии отклонения от омической проводимости при комнатной температуре наступают при напряженности поля около 700 В/см (рис. 6-14).

В области еще более сильных электрических полей энергии носителей заряда увеличивается настолько, что возможно рассеяние их на оптических колебаниях. В этом случае при столкновении с оптическим фооном энергия носителя изменяется на величину, равную $\hbar\omega_0$, и скорость потерь энергии сильно возрастает. В результате устанавливается новое стационарное состояние, при котором энергия $e\mathcal{E}v$, набираемая носителем заряда в поле за единицу времени, равна энергии $\Delta E/\tau$, теряемой этим носителем заряда за то же время, т. е.

$$e\mathcal{E}v = \frac{\Delta E}{\tau}. \quad (6-200)$$

Но энергия, которую носитель заряда теряет при одном столкновении, согласно (6-182) равна:

$$\Delta E = \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2kT}, \quad (6-201)$$

поэтому

$$e\mathcal{E}v = \frac{\hbar\omega_0}{\tau} \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2kT}. \quad (6-202)$$

Если из соотношения (6-188) найти напряженность электрического поля \mathcal{E} и подставить ее в выражение (6-202), получим дрейфовую скорость носителей заряда

$$v = \left(\frac{\hbar\omega_0}{m^*} \operatorname{th} \frac{\hbar\omega_0}{2kT} \right)^{1/2}. \quad (6-203)$$

Предельное значение скорости дрейфа

$$v_s = \left(\frac{\hbar\omega_0}{m^*} \right)^{1/2}. \quad (6-204)$$

Таким образом, в пределе плотность тока, текущего через образец, стремится к насыщению, величина которого

$$j_s = env_s = en \left(\frac{\hbar\omega_0}{m^*} \right)^{1/2}. \quad (6-205)$$

Это значит, что

$$\sigma = \frac{j_s}{\mathcal{E}} = \sigma_0 \mathcal{E}^{-1}. \quad (6-206)$$

Из соотношений (6-202) и (6-203) следует, что при увеличении напряженности электрического поля до величины

$$\mathcal{E}'_{кр} = \frac{v_0}{el} \left(m^* \hbar \omega_0 \operatorname{th} \frac{\hbar \omega_0}{2kT} \right)^{1/2} \quad (6-207)$$

в валентных кристаллах наступает насыщение тока, а следовательно, и насыщение дрейфовой скорости. При полях $\mathcal{E} > \mathcal{E}'_{кр}$ ток и дрейфовая скорость не зависят от поля. Оценки показывают,

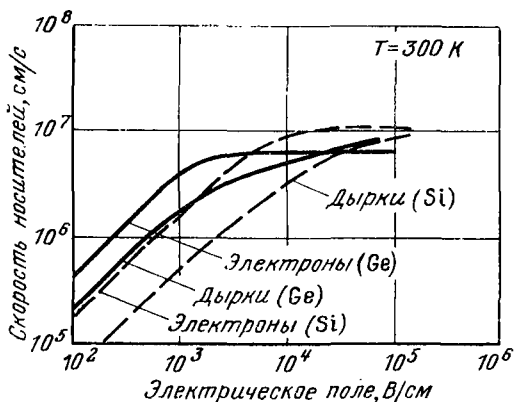


Рис. 6-15. Зависимость скорости носителей заряда от электрического поля для чистых германия и кремния,

что насыщение тока и дрейфовой скорости наступает в полях порядка 10^4 В/см, а величина дрейфовой скорости $v \approx 10^7$ см/с. Это согласуется с экспериментальными данными рис. 6-15, на котором представлена зависимость скорости электронов и дырок от электрического поля для чистых германия и кремния. Из данных рис. 6-14 видно, что при слабых полях выполняется закон Ома. Начиная с критического значения напряженности поля, зависящего от температуры, ток пропорционален $\mathcal{E}^{1,2}$. При очень высоких полях ток не зависит от поля.

В полярных кристаллах рассеяние на оптических колебаниях становится преобладающим уже при низких значениях энергии носителей заряда. В электрических полях, при которых энергия электрона $E \ll \hbar \omega_0$, 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.

В полупроводниках, зона проводимости которых имеет более одного минимума энергии, электрон с волновым вектором k , соответствующим одному из минимумов, при рассеянии может оказаться в состоянии с волновым вектором k' , принадлежащим другому минимуму. В результате такого рассеяния будет иметь место переброс электронов из одного минимума в другой минимум зоны проводимости. Такой вид рассеяния получил название «междолинного».

Междолинное рассеяние носителей заряда в определенных условиях может приводить к возникновению колебаний тока с частотой порядка 10^{10} Гц при приложении к однородному полупроводнику сильного постоянного электрического поля. Это явление, названное эффектом Ганна, впервые наблюдалось на GaAs.

На рис. 6-16 изображена энергетическая структура арсенида галлия в направлении оси $[100]$.

Существенным здесь является наличие двух минимумов A и B , разделенных зазором $\Delta E = 0,36$ эВ, в которых эффективные массы электронов различны. В области нижней долины A электроны легкие, с эффективной массой $m_1^* = 0,068 m_0$, они имеют высокую подвижность [$\mu_1 \approx \approx 4000 \div 8000$ см²/(В·с)]. В области высокой долины B электроны тяжелые с $m_2^* = 1,2 m_0$ и обладают низкой подвижностью [$\mu_2 \approx 100$ см²/(В·с)]. Плотность состояний в верхней долине примерно в 70 раз выше, чем в нижней долине.

При малых внешних электрических полях электроны будут находиться в термодинамическом равновесии с решеткой, имеющей температуру T . Поскольку энергия электронов значительно меньше энергетического зазора $kT \ll \Delta E$, то электроны в основном будут занимать энергетические уровни в нижней долине зоны проводимости (рис. 6-17, а), где они имеют высокую подвижность, малую эффективную массу и малую плотность состояний. В этом случае плотность тока определяется концентрацией электронов в нижней долине $n_1 = n_0$ (в верхней долине $n_2 = 0$), их подвижностью μ_1 и будет линейно возрастать с ростом напряженности поля до некоторого критического значения \mathcal{E}_a (рис. 6-18):

$$j = en_0\mu_1\mathcal{E}. \quad (6-208)$$

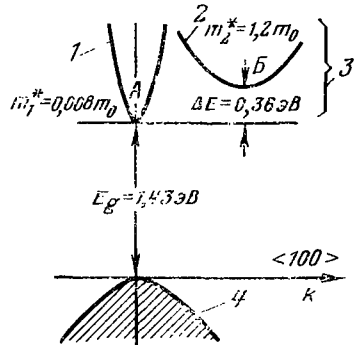


Рис. 6-16. Схематическое изображение структуры зон арсенида галлия n -типа в направлении оси $[100]$.

1 — нижний минимум (A); 2 — верхний минимум (B); 3 — зона проводимости; 4 — валентная зона.

По мере возрастания напряженности электрического поля средняя энергия электронов повышается и в результате рассеяния на оптических фононах становится возможным переход электронов в верхнюю долину B (рис. 6-17, б), где они будут иметь низкую подвижность, большую эффективную массу и высокую плотность состояний (широкий максимум), а суммарная концентрация электронов будет $n_0 = n_1 + n_2$. Таким образом, с ростом напряженности поля (от \mathcal{E}_a до \mathcal{E}_b — рис. 6-18) будет иметь место уменьшение подвижности электронов, а следовательно, уменьшение плотности тока, и на вольт-амперной характеристике появится падающий участок.

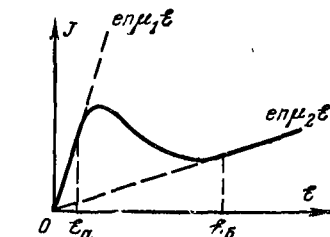
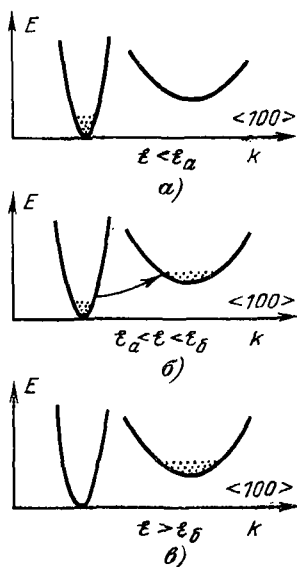


Рис. 6-18. Зависимость плотности тока от напряженности электрического поля для двух-долинного полупроводника.

←
Рис. 6-17. Распределение электронов при различных значениях напряженности поля.

При напряженности поля $\mathcal{E} > \mathcal{E}_b$ все электроны будут находиться в верхней долине $n_2 = n_0$, а в нижней долине $n_1 = 0$ (рис. 6-17, в), и плотность тока:

$$j = en_0\mu_2\mathcal{E}. \quad (6-209)$$

Рассмотрим подробнее механизм такой неустойчивости. Пусть к образцу длиной \mathcal{L} приложено внешнее напряжение. В однородном полупроводнике электрическое поле примерно одинаково по всей длине образца. Но если в образце имеется локальная неоднородность с повышенным сопротивлением, то напряженность электрического поля в этом месте образца будет несколько выше. Следовательно, критическое значение величины $\mathcal{E}_{кр} = \mathcal{E}_a$ при повышении напряженности поля возникнет в первую очередь в этом сечении образца. Как только напряженность поля в области локальной неоднородности достигнет критического значения \mathcal{E}_a , имеет место переход электронов в верхнюю долину B и в этой области

образца с повышенной напряженностью поля появятся тяжелые электроны. Подвижность электронов в этой части образца уменьшается и сопротивление ее возрастает. Это приводит к возрастанию напряженности поля в этом месте образца, что в свою очередь вызывает более интенсивный переход электронов в верхнюю долину. Но так как напряженне, приложенное к образцу, не меняется, то напряженность поля справа и слева от этой области образца будет падать. В результате распределение электрического поля станет резко неоднородным и образуется область сильного электрического поля, которая носит название электрического домена (рис. 6-19, а).

Область тяжелых электронов под действием электрического поля будет перемещаться вдоль образца с относительно низкой скоростью, обусловленной низкой подвижностью тяжелых электронов. 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, поэтому образуется область, обедненная электронами, — область положительного объемного заряда. Следовательно, в пределах области сильного электрического поля на кривой распределения концентрации электронов имеется обедненный участок с $n < n_0$, соответствующий переднему фронту домена, и обогащенный участок с $n > n_0$, соответствующий заднему фронту домена (рис. 6-19, б).

Поскольку внутри домена напряженность поля сильно возросла, возрастает в нем и скорость движения электронов. Вне домена напряженность поля резко уменьшилась, поэтому скорость движения электронов падает. Через некоторый промежуток времени установится стационарное состояние, при котором скорость движения домена v_d будет равна дрейфовой скорости электронов вне домена v_b , т. е.

$$v_d = v_b \quad (6-210)$$

или

$$\mu_2 \mathcal{E}_d = \mu_1 \mathcal{E}_b, \quad (6-211)$$

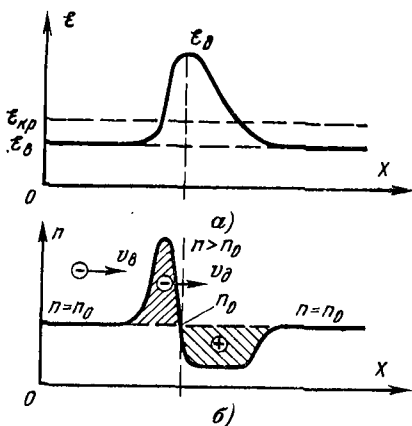


Рис. 6-19. Распределение электрического поля (а) и плотности электронов (б) вдоль образца в случае стабильного домена сильного поля, движущегося со скоростью v_d .

т. е. стационарному состоянию будет соответствовать напряженность поля \mathcal{E}_δ в домене и \mathcal{E}_n вне домена. При этом установившаяся скорость движения домена v_δ будет меньше максимальной скорости движения электронов, которую они имеют при \mathcal{E}_a . Поэтому в момент подключения к образцу напряжения (через $t = t_1$) ток будет иметь максимальное значение $I_{\text{макс}}$, определяемое $v_{\text{макс}}$. Сразу же начнется процесс образования домена, и поскольку этот процесс кратковременный, так как постоянная времени, связанная с междолинным переходом, порядка 10^{-12} с, ток очень быстро спадает

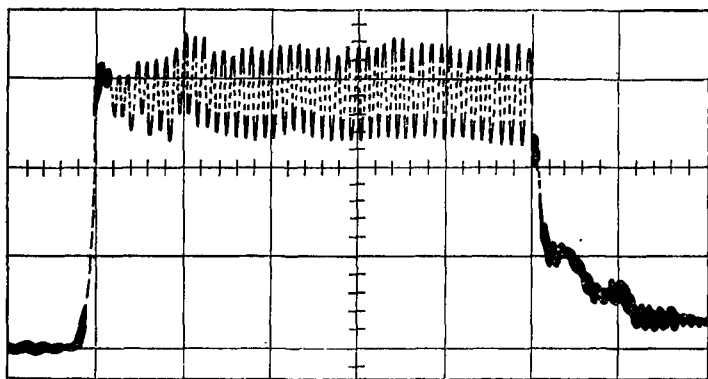


Рис. 6-20. Форма импульса тока при подаче на образец арсенида галлия длиной $2,5 \cdot 10^{-3}$ см импульса напряжения амплитудой 16 В и длительностью 16 нс. Частота переменной составляющей 4,5 ГГц.

до значения $I_{\text{мин}}$, определяемого скоростью установившегося движения домена v_δ :

$$I_{\text{мин}} = s n_0 v_\delta, \quad (6-212)$$

где s — площадь сечения образца.

Минимальное значение тока сохраняется в течение всего времени движения домена вдоль образца. Как установлено, это время определяется длиной образца и скоростью движения домена:

$$T = \frac{L}{v_\delta}. \quad (6-213)$$

По достижении анода область сильного поля выходит из образца и ток в образце начинает возрастать. Как только ток в образце достигнет значения $I_{\text{макс}}$, происходит образование нового домена и ток спадает до $I_{\text{мин}}$. В результате движения домена по кристаллу во внешней цепи появляется импульс тока. Скважность импульсов тока определяется временем T прохождения домена. При длине образца 50 мкм частота колебаний тока должна составлять около 2 ГГц. Примерный вид этих колебаний показан на рис. 6-20.

Поскольку эффект образования домена представляет собой объемную неустойчивость, он не зависит от внешних элементов схемы. Это значит, что скорость движения домена не зависит от величины внешнего напряжения и домен передвигается со скоростью дрейфа 10^7 см/с.

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. В пластиночных образцах домены зарождаются в области повышенного поля вблизи катода, обусловленной неоднородностью уровня легирования, которое возникает в процессе изготовления электродов. Важно также, чтобы расстояние между минимумами *A* и *B* зоны проводимости не было весьма велико, ибо для перехода электрона во вторую долину потребуются поле большой напряженности, при котором возможно увеличение концентрации электронов за счет ударной ионизации или тушельного эффекта.

6-11. УДАРНАЯ ИОНИЗАЦИЯ

В сильном электрическом поле, если оно возрастает примерно до 10^5 В/см, электроны (или дырки) приобретают энергию, достаточную для ионизации атомов вещества. В результате этого возникают электронно-дырочные пары (рис. 6-21), которые в свою очередь ускоряются полем и генерируют дополнительные носители заряда. Этот процесс создания носителей заряда носит название ударной ионизации. Наблюдать, однако, это явление возможно лишь в *p-n* переходах, так как создать поля 10^5 — 10^6 В/см в однородном полупроводнике чрезвычайно трудно при разумных толщинах последнего. Исключением является ударная ионизация примесных атомов, которая требует небольших полей и наблюдается при низких температурах, когда атомы примеси не ионизированы. Ударная ионизация доноров и акцепторов — элементов пятой и третьей групп — в германии происходит при полях, приблизительно равных 5—10 В/см. Например, из рис. 6-22, на котором представлена зависимость плотности тока от электрического поля в электронном германии при 4,2 К, видно, что при $\mathcal{E} \approx 5$ В/см происходит резкое скачкообразное возрастание плотности тока, вызванное ударной ионизацией атомов примеси.

Скорость генерации электронно-дырочных пар *G* вследствие ударной ионизации дается в виде

$$G = \alpha_n n \mu_n + \alpha_p p \mu_p, \quad (6-214)$$

где α_n и α_p — коэффициент ударной ионизации электронов и дырок соответственно. Он определяется как число электронно-дырочных пар, генерируемых электроном (дыркой) на единице пути перемещения. Коэффициент ударной ионизации как электронов α_n , так и дырок α_p сильно зависит от значения напряженности электрического поля. На рис. 6-23 приведены экспериментально измеренные значения коэффициентов ударной ионизации для германия, кремния, арсенида галлия и фосфида галлия.

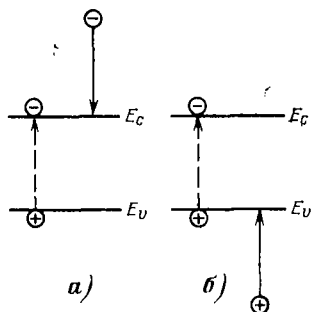


Рис. 6-21. Два возможных варианта (а, б) ударной ионизации.

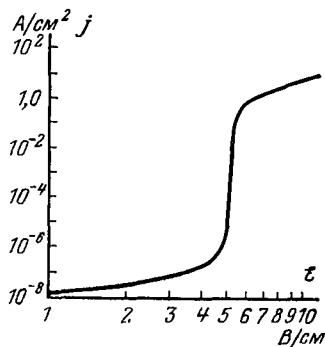


Рис. 6-22. Ударная ионизация в электронном германии при $T = 4,2$ К.

Коэффициент ударной ионизации $\alpha(E)$ определяется соотношением

$$\alpha(E) = \frac{1}{v} \int_0^{\infty} W_n(E) f(E) N(E) dE, \quad (6-215)$$

где $W_n(E)$ — вероятность ионизации, рассчитываемая квантовомеханическим путем; $f(E)$ — неравновесная функция распределения; $N(E)$ — плотность квантовых состояний; v — скорость дрейфа. Для полупроводников с большой диэлектрической проницаемостью вероятность ионизации (для Ge и Si):

$$W_n(E) \approx A(E - E_n)^2, \quad (6-216)$$

где A — коэффициент пропорциональности, а E_n — энергия ионизации.

При ударной ионизации участвуют три носителя заряда, которые примерно поровну распределяют между собой энергию первоначального носителя заряда; пороговая энергия ионизации $E_n = \frac{3}{2} E_g$, где E_g — ширина запрещенной зоны. Зависимость

коэффициента ударной ионизации $\alpha(E)$ от поля в основном определяется видом функции распределения $f(E)$. Как показывает детальный анализ с помощью кинетического уравнения, полное аналитическое выражение для функции распределения в области энер-

гии $0 < E < \infty$ получить нельзя; $f(E)$ можно получить в явном виде для двух участков: в области $E < E_n$ и $E > E_n$.

Ускорение электронов до порога ионизации E_n зависит от соотношения двух факторов — ускорения во внешнем электрическом поле и рассеяния энергии при столкновении 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

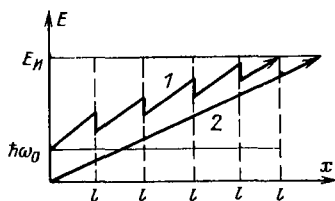
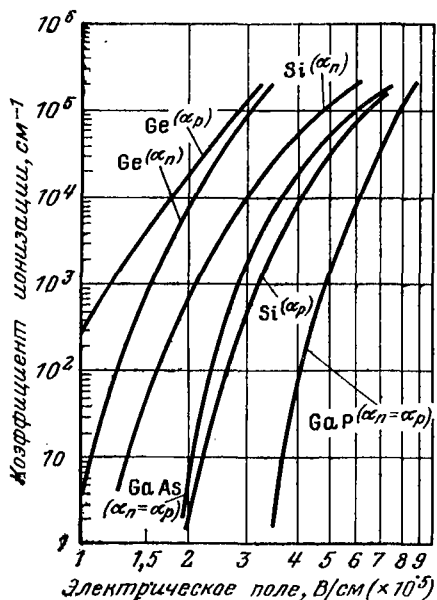


Рис. 6-24. Условная диаграмма ударной ионизации.

Рис. 6-23. Зависимость измеренных коэффициентов ударной ионизации от электрического поля для германия, кремния, арсенида галлия и фосфида галлия.

не зависит от энергии, т. е. $l = \text{const}$. Ускорение до порога ионизации E_n можно проиллюстрировать следующей условной диаграммой, изображенной на рис. 6-24. Здесь по вертикали отложена энергия, приобретаемая электроном в электрическом поле, а по горизонтали — расстояние, проходимое электроном в среднем между двумя ионизирующими столкновениями. Как видно из рис. 6-24, существуют две возможности ускорения электронов до порога ионизации.

В первом случае в очень сильных электрических полях, когда выполняется условие $eEl \gg \hbar\omega_0$, электроны могут набрать энергию в результате многократных столкновений с фононами, так как относительная потеря энергии в каждом таком столкновении мала (линия 1). В этом случае, как показывает теория,

$$\alpha(E) \sim e^{-E_n/E^{5/2}}. \quad (6-217)$$

Здесь можно провести аналогию с диффузионным движением частицы, поэтому указанное приближение для $\alpha(E)$ называют диффузионным.

Во втором случае, когда поля относительно слабые ($e\mathcal{E}l \ll \hbar\omega_0$), носители заряда набирают энергию на расстоянии, равном многим длинам свободного пробега (линия 2). В такой ситуации существует только одна возможность набрать энергию E_n , когда носители заряда не сталкиваются с фононами до порога ионизации. Поскольку путь до порога ионизации $\mathcal{L}_n = E_n/e\mathcal{E}$, а средний свободный пробег между столкновениями с фононами равен l , то вероятность пробега до порога ионизации без столкновений есть $e^{-\mathcal{L}_n/l} \approx e^{-E_n/e\mathcal{E}l}$. В связи с этим функция распределения, а стало быть, и коэффициент ударной ионизации пропорциональны величине

$$\alpha(E) \sim e^{-E_n/e\mathcal{E}l}. \quad (6-218)$$

Функция распределения будет сильно вытянута вдоль направления электрического поля, т. е. в направлении дрейфа носителей, поэтому это приближение для $\alpha(E)$ называют дрейфовым.

Как показали эксперименты, в Ge и Si в основном $\alpha(E)$ определяется законом вида (6-218). Однако в узких p - n переходах (особенно четко это проявляется в Si) наблюдается зависимость вида (6-217). Это имеет место при напряженности поля $\mathcal{E} \geq 5 \cdot 10^5$ В/см.

6-12. ТУННЕЛЬНЫЙ ЭФФЕКТ И ЭЛЕКТРОСТАТИЧЕСКАЯ ИОНИЗАЦИЯ

В очень сильных электрических полях становится возможным еще один механизм образования свободных носителей заряда — туннельный эффект, или эффект Зинера. Как известно, у полупроводника, находящегося во внешнем электрическом поле, имеет место наклон энергетических зон (см. § 2-10). Наклон зон тем больше, чем больше величина напряженности электрического поля \mathcal{E} . На рис. 6-25 представлена зонная структура собственного полупроводника при наличии сильного электрического поля.

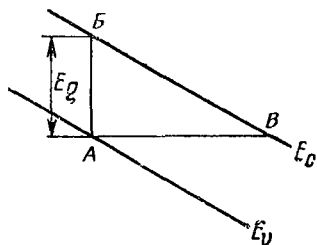


Рис. 6-25. Зонная структура собственного полупроводника при наличии сильного электрического поля.

В этом случае, как следует из рис. 6-25, возможен переход электрона из валентной зоны в зону проводимости благодаря туннельному эффекту. Вероятность туннельного перехода зависит от высоты и ширины потенциального барьера. В данном случае высота потенциального барьера AB представляет собой ширину запрещенной зоны E_g . Эффективная ширина барьера $AB = \Delta x$ может быть определена из разности потенциальной энергии электрона в зоне проводимости в точке B и в валентной зоне — в точке A . Так

В данном случае высота потенциального барьера AB представляет собой ширину запрещенной зоны E_g . Эффективная ширина барьера $AB = \Delta x$ может быть определена из разности потенциальной энергии электрона в зоне проводимости в точке B и в валентной зоне — в точке A . Так

как потенциальная энергия с точностью до аддитивной постоянной есть

$$U(B) - U(A) = -e\mathcal{E}\Delta x = -E_g, \quad (6-219)$$

то эффективная ширина барьера

$$\Delta x = \frac{E_g}{e\mathcal{E}}. \quad (6-220)$$

Следовательно, ширина потенциального барьера зависит от напряженности электрического поля.

Переход электрона из точки A в точку B связан с переходом сквозь треугольный потенциальный барьер ABB . Но, как известно из квантовой механики*, вероятность перехода для барьера треугольной формы имеет вид:

$$D = D_0 e^{-2\sqrt{2m^*E_g} \frac{3:2/e\mathcal{E}s}}. \quad (6-221)$$

Из этого выражения следует, что вероятность туннельного перехода из зоны в зону растет экспоненциально с ростом напряженности электрического поля \mathcal{E} . Кроме того, вероятность туннельного перехода будет больше у полупроводников с малыми значениями ширины запрещенной зоны E_g и эффективной массы m^* (рис. 6-26).

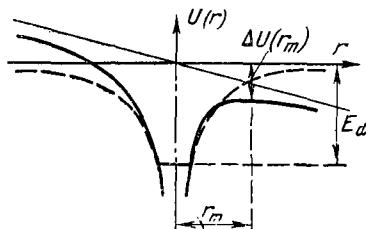
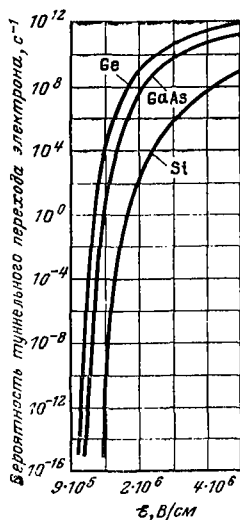


Рис. 6-27. Понижение потенциального барьера под действием внешнего электрического поля.

Рис. 6-26. Вероятность туннельного перехода электронов в зависимости от напряженности поля в p - n переходе.

Вероятность туннельного перехода одна и та же как для переходов из валентной зоны в зону проводимости, так и из зоны проводимости в валентную зону. Но поскольку в валентной зоне электронов намного больше, чем в зоне проводимости, то будет иметь

* Шпольский Э. В. Атомная физика. М., «Наука», 1974, т. 1, § 154.

место преимущественный переход электронов из валентной зоны в зону проводимости. Поэтому туннельный эффект может приводить к значительному увеличению концентрации свободных носителей заряда. Он наблюдается в полях 10^6 В/см.

Увеличение концентрации свободных носителей заряда может быть также обусловлено уменьшением энергии ионизации атомов донорной примеси в сильном электрическом поле. На рис. 6-27 представлена потенциальная энергия электрона в атоме донорной примеси без поля (пунктирная кривая) и при наличии внешнего электрического поля (сплошная кривая). При отсутствии электрического поля электрон локализован в потенциальной яме, образованной положительно заряженным донором, а энергия, необходимая для его ионизации, равна E_d .

Внешнее электрическое поле понижает потенциальный барьер в направлении, противоположном направлению электрического поля.

Потенциальная энергия электрона с учетом действия внешнего электрического поля, направленного вдоль вектора \mathbf{r} , может быть записана в виде:

$$U(\mathbf{r}) = -e\mathcal{E}r - \frac{e^2}{4\pi\epsilon\epsilon_0 r}. \quad (6-222)$$

Первое слагаемое в этом равенстве отражает изменение потенциальной энергии во внешнем электрическом поле напряженности \mathcal{E} , а второе — в кулоновском поле донора. Как следует из (6-222), энергия $U(\mathbf{r})$ в некоторой точке \mathbf{r}_m принимает экстремальное значение, а потенциальный барьер в этой точке понижается на величину

$$\Delta U(\mathbf{r}_m) = -e \sqrt{\frac{e\mathcal{E}}{\pi\epsilon\epsilon_0}}, \quad (6-223)$$

а следовательно, на такую же величину уменьшается и энергия ионизации донора.

Уменьшение энергии ионизации донора увеличивает вероятность термического возбуждения согласно статистике Больцмана на величину

$$e^{-\Delta U/kT} = e^{\frac{(e^2/\pi\epsilon\epsilon_0)^{1/2}}{kT} \sqrt{\mathcal{E}}}. \quad (6-224)$$

Напряженность электрического поля, при которой имеет место заметное возрастание концентрации электронов, обычно составляет 10^4 — 10^5 В/см. Этот механизм образования свободных носителей заряда проявляется только в том случае, когда концентрация неионизированной донорной примеси значительно превышает концентрацию свободных электронов.

ГЕНЕРАЦИЯ И РЕКОМБИНАЦИЯ
НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА

7-1. РАВНОВЕСНЫЕ И НЕРАВНОВЕСНЫЕ НОСИТЕЛИ ЗАРЯДА

Рассмотрим донорный полупроводник, находящийся в состоянии термодинамического равновесия. В результате тепловой генерации электроны с донорной примеси переходят в зону проводимости. При достаточно высоких температурах будет существенным переброс электронов непосредственно из валентной зоны в зону проводимости. На рис. 7-1 стрелками показаны переходы электронов при тепловом возбуждении. Свободные носители заряда, возникающие в результате термической генерации и находящиеся в тепловом равновесии с кристаллической решеткой, называются равновесными.

На рис. 7-1 изображены плотность квантовых состояний $N(E)$ в зоне проводимости и валентной зоне, функция распределения Ферми—Дирака $f_0(E)$ и концентрация равновесных электронов n_0 и дырок p_0 , которые занимают состояния вблизи краев соответствующих зон (заштрихованные области).

Одновременно с генерацией свободных носителей заряда идет процесс рекомбинации: электроны возвращаются в свободные состояния в валентной зоне, в результате чего исчезают свободный электрон и свободная дырка. В условиях термодинамического равновесия 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 концентраций свободных носителей заряда, поэтому

$$R_0 = \gamma_r n_0 p_0. \quad (7-1)$$

Здесь γ_r — коэффициент пропорциональности, называемый коэффициентом рекомбинации. Для равновесного состояния полупроводника справедливо равенство

$$G_0 = R_0, \quad (7-2)$$

которое является выражением принципа детального равновесия.

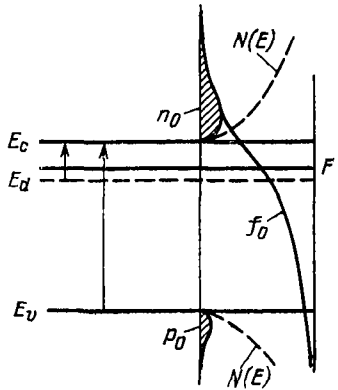


Рис. 7-1. Тепловая генерация свободных носителей заряда и их концентрация в состоянии термодинамического равновесия.

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 возникает некоторая концентрация электронов n и дырок p , которая отличается от термодинамически равновесной. *Подвижные носители заряда, не находящиеся в термодинамическом равновесии как по концентрации, так и по энергетическому распределению, называются неравновесными носителями заряда, а их концентрация n , p называется неравновесной концентрацией. Избыток неравновесной концентрации носителей заряда Δn , Δp в полупроводнике по сравнению с равновесной n_0 , p_0 называется избыточной концентрацией носителей заряда.*

При генерации, например при поглощении света, носители заряда могут иметь кинетическую энергию, значительно превышающую среднюю тепловую энергию равновесных частиц. В результате рассеяния на дефектах кристаллической решетки носители заряда довольно быстро передают ей свою избыточную энергию. Пусть энергия генерированных светом электронов была в зоне проводимости примерно 1 эВ. Рассеиваясь на длинноволновых фононах, они за каждое столкновение будут передавать решетке энергию порядка $6 \cdot 10^{-4}$ эВ (см. с. 165). Для рассеивания своей избыточной энергии избыточные электроны должны совершить около 1600 столкновений с акустическими фононами. Так как средняя длина свободного пробега электронов порядка 10^{-6} см, а их тепловая скорость при комнатной температуре составляет приблизительно 10^7 см/с, при этом среднее время между двумя столкновениями равно $\tau \approx 10^{-13}$ с, то уже через $1,6 \cdot 10^{-10}$ с избыточные электроны приобретут температуру кристаллической решетки и не будут отличаться от равновесных носителей заряда. Поэтому распределение по энергиям неравновесных и равновесных носителей заряда будет одинаковым. Этот процесс сводится к тому, что неравновесные электроны, рассеивая свою избыточную энергию, как бы «опускаются» к нижнему краю зоны проводимости, а неравновесные дырки, рассеивая избыточную энергию, «поднимаются» к верхнему краю валентной зоны. Если при этом концентрация неравновесных носителей заряда мало отличается от равновесной, можно считать, что энергия кристалла практически не изменится, а значит, не меняется температура кристалла и, следовательно, не меняется концентрация равновесных носителей заряда. В этом случае общее количество электронов и дырок будет равно соответственно:

$$\left. \begin{aligned} n &= n_0 + \Delta n; \\ p &= p_0 + \Delta p. \end{aligned} \right\} \quad (7-3)$$

Стационарную концентрацию неравновесных электронов n можно выразить с помощью формулы, аналогичной (4-40),

$$n = n_0 + \Delta n = \int_{E_c}^{\infty} P_e(E) N(E) dE, \quad (7-4)$$

где $P_e(E)$ — функция распределения для неравновесных электронов, отличная от равновесной функции распределения $f_0(E)$ и стремящаяся к ней при приближении неравновесной системы к состоянию термодинамического равновесия. Это выражение можно представить в виде, аналогичном соотношению (4-42), введя к в а з и у р о в е н ь Ф е р м и для электронов в неравновесном состоянии F_n так, чтобы выполнялось условие

$$n = N_c F_{1,2}(\eta_n), \quad (7-5)$$

где

$$\eta_n = \frac{F_n - E_c}{kT} \quad (7-6)$$

есть приведенный квазиуровень Ферми для электронов.

Концентрацию неравновесных электронов в невырожденном полупроводнике с учетом равенств (4-65), (4-66) и (4-69) можно представить в виде

$$n = N_c e^{\eta_n} = n_0 e^{\eta_n - \eta_i} = n_i e^{\eta_n - \eta_i}. \quad (7-7)$$

Аналогично концентрацию неравновесных дырок можно выразить следующим образом:

$$p = p_0 + \Delta p = \int_{-\infty}^{E_v} P_h(E) N(E) dE = N_v F_{1,2}(-\eta_p - \varepsilon_i), \quad (7-8)$$

где

$$\eta_p = \frac{F_p - E_v}{kT} \quad (7-9)$$

есть приведенный квазиуровень Ферми для дырок. При отсутствии вырождения, учитывая формулы (4-67) и (4-69), можно написать:

$$p = N_v e^{-\eta_p - \varepsilon_i} = p_0 e^{\eta_i - \eta_p} = n_i e^{\eta_i - \eta_p}. \quad (7-10)$$

Таким образом, в неравновесном состоянии уровень Ферми как бы расщепляется на два квазиуровня — для электронов F_n и для дырок F_p (рис. 7-2).

Произведение концентраций электронов и дырок для неравновесного состояния отличается теперь от его значения для равновесного состояния:

$$np = n_0 p_0 e^{\eta_n - \eta_p} = n_i^2 e^{\eta_n - \eta_p}. \quad (7-11)$$

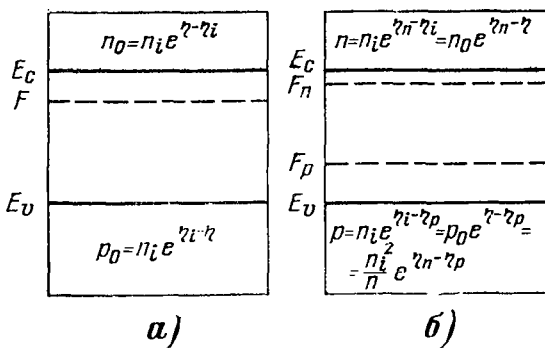


Рис. 7-2. Расщепление уровня Ферми на два квазиуровня — для электронов F_n и дырок F_p .

а — равновесное состояние; б — неравновесное состояние.

Расстояние между приведенными квазиуровнями Ферми ($\eta_n - \eta_p$) характеризует отклонение системы от состояния термодинамического равновесия:

$$\frac{np}{n_0 p_0} = \frac{np}{n_i^2} = e^{\eta_n - \eta_p}. \quad (7-12)$$

Это соотношение выражает связь между концентрациями электронов и дырок в неравновесном состоянии. При этом чем сильнее различаются квазиуровни Ферми электронов и дырок, тем сильнее отличается произведение концентраций неравновесных носителей заряда от произведения равновесных концентраций.

7-2. БИПОЛЯРНАЯ СВЕТОВАЯ ГЕНЕРАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА

Рассмотрим полупроводник, в котором в результате поглощения света создается некоторая избыточная концентрация электронов и дырок. *Генерацию носителей заряда, при которой в результате возбуждения возникают пары — электрон и дырка, будем называть биполярной генерацией.* В этом случае поглощение кванта света сопровождается разрывом валентной связи и количество генерируемых избыточных электронов и дырок одинаково:

$$\Delta n = \Delta p. \quad (7-13)$$

Параллельно с процессом генерации идет процесс рекомбинации, и в стационарном состоянии число электронов и дырок, возбуждаемых за единицу времени в результате разрыва валентных связей, равно числу электронов и дырок, рекомбинирующих за это же время. Поскольку неравновесные носители заряда через очень малое время становятся физически неотличимыми по энергии от равновесных, можно считать, что они имеют тот же коэффициент рекомбинации γ_r , что и равновесные носители заряда.

После выключения возбуждающего света концентрации электронов и дырок уменьшаются в результате рекомбинации. При этом скорость убывания числа свободных электронов и дырок определяется разностью скоростей рекомбинации и тепловой генерации

$$-\left(\frac{dn}{dt}\right)_r = -\left(\frac{dp}{dt}\right)_r = \gamma_r np - G_0. \quad (7-14)$$

Принимая во внимание соотношения (7-1), (7-2), (7-3) и (7-13), равенство (7-14) можно переписать следующим образом:

$$-\left(\frac{dn}{dt}\right)_r = \gamma_r (np - n_0 p_0) = \gamma_r (n + p_0 + \Delta n) \Delta n. \quad (7-15)$$

Учитывая, что в случае малого уровня возбуждения $\Delta n \ll n_0 + p_0$, получаем:

$$\left(\frac{dn}{dt}\right)_r = -\gamma_r (n_0 + p_0) \Delta n = -\gamma_r (n_0 + p_0) (n - n_0). \quad (7-16)$$

Введем обозначение

$$\tau = \frac{1}{\gamma_r (n_0 + p_0)}. \quad (7-17)$$

Тогда уравнение (7-16) примет вид:

$$\left(\frac{dn}{dt}\right)_r = -\frac{n - n_0}{\tau} = -\frac{\Delta n}{\tau}, \quad (7-18)$$

откуда

$$\Delta n = \Delta n(0) e^{-t/\tau}, \quad (7-19)$$

где $\Delta n(0)$ — избыточная концентрация электронов в момент выключения возбуждающего света.

Таким образом, в случае малого уровня биполярной генерации избыточная концентрация электронов и дырок после прекращения возбуждения уменьшается по экспоненциальному закону и за время τ число их в результате рекомбинации убывает в e раз. Следовательно, τ представляет собой среднее время существования избыточной концентрации электронов и дырок и называется временем жизни неравновесных 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 неравновесных носителей заряда в зависимости от типа полупроводникового материала и от степени чистоты может изменяться в очень широких пределах — от 10^{-2} до 10^{-8} с. Поскольку величина τ достаточно велика, свободные электроны и дырки, диффундируя из приповерхностной области, где их концентрация повышена, в объем полупроводника, за время жизни успеют пройти большое расстояние. Следовательно, в случае биполярной генерации область генерации носителей заряда и область их рекомбинации пространственно не совпадают.

В случае большого уровня биполярного возбуждения, когда $\Delta n \gg n_0 + p_0$, из (7-15) найдем, что

$$-\left(\frac{dn}{dt}\right)_r = \gamma_r (\Delta n)^2, \quad (7-20)$$

т. е. скорость рекомбинации носителей заряда зависит от Δn по квадратичному закону. Разделяя переменные, запишем:

$$\frac{dn}{(\Delta n)^2} = -\gamma_r dt, \quad (7-21)$$

откуда

$$\Delta n = \frac{\Delta n(0)}{1 + \gamma_r \Delta n(0) t}. \quad (7-22)$$

Из равенства (7-22) следует, что при квадратичной рекомбинации избыточная концентрация носителей заряда уменьшается по гиперболическому закону.

Если ввести мгновенное время жизни $\tau_{\text{мгн}}$, то в случае квадратичной рекомбинации согласно уравнению (7-20)

$$\left(\frac{dn}{dt}\right)_r = -\frac{\Delta n}{\tau_{\text{мгн}}} \quad (7-23)$$

и

$$\tau_{\text{мгн}} = \frac{1}{\gamma_r \Delta n}, \quad (7-24)$$

т. е. $\tau_{\text{мгн}}$ зависит от концентрации избыточных носителей заряда и в связи с этим является переменной величиной. Однако в каждый момент времени $\tau_{\text{мгн}}$ имеет определенный смысл, являясь мгновенным временем жизни носителей заряда. В силу этого при большом уровне инжекции имеем дело с мгновенным временем жизни, которое согласно уравнению (7-23) равно:

$$\tau_{\text{мгн}} = -\frac{\Delta n}{\frac{d\Delta n}{dt}}. \quad (7-25)$$

7.3. 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. Пусть облучение полупроводника светом приводит к перебросу электронов с донорных уровней в зону проводимости. Такой процесс возбуждения носителей заряда будем называть монополярной

световой генерацией. Она характеризуется образованием избыточной концентрации основных носителей заряда (в данном случае электронов). При этом электронейтральность полупроводника не нарушается, так как избыточный заряд свободных электронов скомпенсирован зарядом образовавшихся положительных ионов донорной примеси.

Если неравновесные электроны генерируются в некоторой области образца, то концентрация электронов в этой области будет повышенной $n = n_0 + \Delta n$. Наличие избыточной концентрации электронов вызовет перемещение их в область, где генерация отсутствует, что приведет к возникновению объемного заряда. Вне области генерации возникнет отрицательный объемный заряд избыточных электронов, а в области генерации будет сосредоточен положительный заряд, обусловленный ионами донорной примеси. Если в момент времени $t = 0$ генерация прекратится, то электрическое поле напряженности \mathcal{E} , порожденное объемным зарядом, вызовет ток проводимости, который в течение некоторого времени полностью уничтожит объемный заряд.

Изменение плотности пространственного заряда ρ в результате протекания тока, плотность которого равна \mathbf{j} , как известно, подчиняется уравнению непрерывности электрического заряда

$$\operatorname{div} \mathbf{j} = \operatorname{div} (\sigma \mathcal{E}) = -\frac{\partial \rho}{\partial t}, \quad (7-26)$$

а электрическое поле и объемный заряд связаны уравнением Пуассона

$$\operatorname{div} \mathcal{E} = \frac{\rho}{\epsilon \epsilon_0}. \quad (7-27)$$

Здесь ϵ — диэлектрическая проницаемость полупроводника; ϵ_0 — диэлектрическая проницаемость вакуума.

Из уравнений (7-26) и (7-27) следует, что

$$\frac{\partial \rho}{\partial t} = -\frac{\sigma}{\epsilon \epsilon_0} \rho. \quad (7-28)$$

Отсюда найдем закон изменения объемного заряда во времени

$$\rho = \rho_0 e^{-t/\tau_\mu}, \quad (7-29)$$

где ρ_0 — плотность объемного заряда в момент $t = 0$;

$$\tau_\mu = \frac{\epsilon \epsilon_0}{\sigma} \quad (7-30)$$

есть диэлектрическое, или максвелловское, время релаксации. Оценим численное значение максвелловского времени релаксации. Для германия, имеющего $\sigma = 1 \text{ Ом}^{-1} \cdot \text{см}^{-1} = 10^2 \text{ См/м}$ и $\epsilon = 16$, $\tau_\mu = 10^{-12} \text{ с}$.

Как следует из равенства (7-29), в случае монополярной генерации носителей заряда возникает объемный заряд, который со временем уменьшается по экспоненциальному закону с постоянной времени τ_{μ} . Другими словами, объемный заряд, созданный избыточной концентрацией основных носителей заряда, в результате тока проводимости исчезнет в среднем через τ_{μ} . Поскольку максвелловское время релаксации τ_{μ} для полупроводников достаточно малая величина, электронное облако избыточных носителей заряда за это время не может сместиться на значительное расстояние по отношению к ионам примеси, а поэтому повышенная концентрация носителей заряда будет в той же области полупроводника, где происходит их генерация. Следовательно, монополярная генерация и рекомбинация носителей заряда имеют место в одной и той же области примесного полупроводника.

7-4. 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 квазиимпульса. Если энергия и квазиимпульс электрона в зоне проводимости (до рекомбинации) были E' и p' , а после рекомбинации они равны соответственно E и p , то законы сохранения энергии и квазиимпульса запишутся в виде

$$E' = E + \Delta E; \quad p' = p + Q. \quad (7-31)$$

Здесь Q — квазиимпульс, передаваемый электроном кристаллической решетке, а ΔE — энергия, которая выделяется при процессе рекомбинации.

В зависимости от того, каким образом расходуется энергия ΔE , межзонная рекомбинация разделяется на три типа:

если энергия, освобождающаяся в процессе рекомбинации, излучается в виде кванта света, рекомбинация называется излучательной или фотонной;

если энергия, освобождающаяся в процессе рекомбинации, передается решетке, т. е. расходуется на образование фононов, рекомбинация называется безызлучательной или фононной;

если в акте рекомбинации выделяющаяся энергия передается третьему свободному носителю заряда, рекомбинация называется ударной или рекомбинацией Оже.

В полупроводниковом кристалле всегда имеются дефекты, энергетические уровни которых находятся в запрещенной зоне. По-

этому наряду с межзонной рекомбинацией может идти процесс рекомбинации через локальные центры. Дефект решетки, способный захватить электрон из зоны проводимости и дырку из валентной зоны, осуществляя их рекомбинацию, называется рекомбинационной ловушкой. В этом случае исчезновение пары — электрона проводимости и дырки проводимости — осуществляется следующим образом. Нейтральная рекомбинационная ловушка захватывает электрон из зоны проводимости, который затем через некоторое время перейдет в валентную зону. В зависимости от способа расходования энергии ΔE при переходе электрона на более-низкий энергетический уровень рекомбинация через рекомбинационные ловушки также может быть фотонной или фононной.

Рекомбинация носителей заряда, идущая на поверхности полупроводника, называется поверхностной рекомбинацией. Поверхностная рекомбинация происходит через поверхностные рекомбинационные ловушки и может протекать с излучением и без излучения света.

7.5. МЕЖЗОННАЯ ИЗЛУЧАТЕЛЬНАЯ РЕКОМБИНАЦИЯ

При межзонной излучательной рекомбинации электрон из зоны проводимости переходит в валентную зону, при этом излучается квант света, энергия которого равна ширине запрещенной зоны:

$$h\nu = E_g. \quad (7-32)$$

При термодинамическом равновесии количество носителей заряда, рекомбинирующих в единице объема за единицу времени с испусканием фотонов, в соответствии с принципом детального равновесия равно числу носителей заряда, возбуждаемых при поглощении фотонов. Иными словами, плотность равновесного излучения R_n^0 равна числу поглощаемых квантов R_p^0 в единице объема за одну секунду. Но интенсивность процесса излучательной рекомбинации пропорциональна произведению концентраций равновесных электронов и дырок:

$$R_n^0 = R_p^0 = \gamma_r n_0 p_0 = \gamma_r n_i^2. \quad (7-33)$$

Здесь γ_r — коэффициент межзонной излучательной рекомбинации.

При отклонении от равновесного состояния концентрации неравновесных носителей заряда равны $n = n_0 + \Delta n$ и $p = p_0 + \Delta p$. Они, как отмечалось в § 7-2, характеризуются тем же коэффициентом рекомбинации, что и равновесные. При этом, если не наступает вырождения, сохраняется пропорциональность скорости рекомбинации R произведению концентраций неравновесных электронов и дырок. Поэтому, определяя γ_r из соотношения (7-33), можно написать:

$$R = \gamma_r n p = R_n^0 \frac{n p}{n_i^2}. \quad (7-34)$$

Определим время жизни неравновесных носителей заряда при межзонной излучательной рекомбинации. Для электронов в соответствии с равенством (7-18) имеем:

$$\tau_{nr} = -\frac{\Delta n}{dn/dt}. \quad (7-35)$$

Когда внешнее возбуждение снято, скорость изменения концентрации свободных электронов определяется разностью интенсивностей рекомбинации R и равновесной генерации R_n^0 . Так как $R_n^0 = R_n^0$, то

$$-\frac{dn}{dt} = R - R_n^0. \quad (7-36)$$

С учетом этого равенство (7-35) запишется в виде

$$\tau_{nr} = \frac{\Delta n}{R - R_n^0} = \frac{\Delta n}{\Delta R}, \quad (7-37)$$

где ΔR — изменение скорости рекомбинации при отклонении системы от равновесного состояния.

Аналогичное равенство можно записать для дырок:

$$\tau_{pr} = \frac{\Delta p}{R - R_n^0} = \frac{\Delta p}{\Delta R}. \quad (7-38)$$

Величину ΔR определим с помощью формул (7-34) и (7-33):

$$\Delta R = R - R_n^0 = R_n^0 \frac{np - n_i^2}{n_i^2}. \quad (7-39)$$

Подставив в соотношение (7-39) значения n и p и заменив n_i^2 на $n_0 p_0$, получим:

$$\Delta R = R_n^0 \frac{n_0 \Delta p + p_0 \Delta n + \Delta n \Delta p}{n_0 p_0}. \quad (7-40)$$

Теперь выражение (7-37) принимает вид:

$$\tau_{nr} = \Delta n \frac{n_0 p_0}{R_n^0 (n_0 \Delta p + p_0 \Delta n + \Delta n \Delta p)}. \quad (7-41)$$

Если $\Delta n = \Delta p$, то формула (7-41) упрощается:

$$\tau_{nr} = \tau_{pr} = \frac{n_0 p_0}{R_n^0 (n_0 + p_0 + \Delta n)}. \quad (7-42)$$

Проведем анализ полученного выражения. В случае малого уровня возбуждения $\Delta n \ll n_0 + p_0$ с учетом (7-33) имеем:

$$\tau_{nr} = \tau_{pr} = \frac{n_i^2}{R_n^0 (n_0 + p_0)} = \frac{1}{\gamma_r (n_0 + p_0)}. \quad (7-43)$$

Для собственного полупроводника, у которого $n_0 = p_0 = n_i$,

$$\tau_{ir} = \frac{n_i}{2R_{II}^0} = \frac{1}{2\gamma_r n_i} = \frac{e^{E_g/2kT}}{2\gamma_r (N_c N_v)^{1/2}}. \quad (7-44)$$

Для материала n -типа ($n_0 \gg p_0$)

$$\tau_r^n = \frac{p_0}{R_{II}^0} = 2 \frac{n_i}{n_0} \tau_{ir} = \frac{1}{\gamma_r n_0}, \quad (7-45)$$

а для дырочного полупроводника ($p_0 \gg n_0$)

$$\tau_r^p = \frac{n_0}{R_{II}^0} = 2 \frac{n_i}{p_0} \tau_{ir} = \frac{1}{\gamma_r p_0}. \quad (7-46)$$

Анализ формул (7-44) — (7-46) показывает, что время жизни неравновесных носителей заряда в собственном полупроводнике при межзонной излучательной рекомбинации тем меньше, чем выше температура и чем меньше ширина запрещенной зоны (см. табл. 7-1), а в примесном полупроводнике время жизни неравновесных носителей заряда меньше, чем в собственном, и с ростом степени легирования и температуры оно уменьшается.

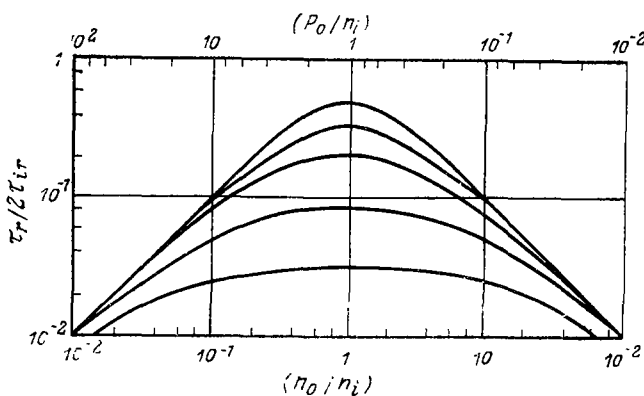


Рис. 7-3. Зависимость времени жизни от концентрации носителей заряда при заданной температуре в случае низкого уровня возбуждения. В максимумах отношение $\Delta n/n_i = 0$ для верхней кривой и соответственно равно 1, 3, 10, 30 для последующих кривых.

Зависимость времени жизни для излучательной межзонной рекомбинации от степени легирования, т. е. концентрации носителей заряда при постоянной температуре, в случае малого уровня возбуждения представлена на рис. 7-3. На этом рисунке логарифмическую шкалу концентрации носителей заряда можно рассматривать так же, как линейную шкалу положения уровня Ферми, так как $n = N_c e^{-(E_c - F)/kT}$, причем ее средняя точка соответствует значению уровня Ферми для собственного полупроводника. Из приведенных зависимостей видно, что с увеличением уровня возбужде-

ния время жизни в собственном полупроводнике резко снижается, а в примесном изменяется сравнительно слабо.

Рассчитаем величину $R_{\text{н}}^0$, используя для этого теорию Эйнштейна для равновесного излучения. Число квантовых состояний для фотонов в интервале импульса между P и $P + dP$ составляет:

$$\frac{P^2 dP}{h^3}. \quad (7-47)$$

Чтобы определить количество ячеек, занятых фотонами, эту величину следует умножить на функцию распределения Бозе—Эйнштейна:

$$f_{\Phi} = \frac{1}{e^{h\nu/kT} - 1}. \quad (7-48)$$

Общее число фотонов с импульсом P , излучаемых в телесный угол $d\Omega$, равно:

$$dn_{\Phi} = \frac{2}{h^3} f_{\Phi} P^2 dP d\Omega, \quad (7-49)$$

где коэффициент 2 учитывает различную поляризацию света.

Импульс фотона связан с энергией ($E = h\nu$) следующим образом:

$$P = \frac{E}{v_{\Phi}} = \frac{h\nu}{v_{\Phi}} = \frac{\bar{n}h\nu}{c}. \quad (7-50)$$

Здесь $v_{\Phi} = c/\bar{n}$ — скорость фотона в данной среде; c — скорость света в вакууме; \bar{n} — показатель преломления среды. Из формулы (7-50) получим:

$$dP = \frac{\bar{n}h}{c} dv, \quad (7-51)$$

что при подстановке в уравнение (7-49) после интегрирования по всем углам от 0 до 4π дает:

$$dN = \int_0^{4\pi} \frac{2}{h^3} f_{\Phi} P^2 dP d\Omega = \frac{8\pi}{h^3} f_{\Phi} P^2 dP = \frac{8\pi}{c^3} \frac{\bar{n}^3 v^2 dv}{e^{h\nu/kT} - 1}. \quad (7-52)$$

Если обозначить вероятность поглощения фотона с частотой ν через $g(\nu)$, то для интенсивности поглощения, которая в условиях термодинамического равновесия равна интенсивности излучения, получим следующее выражение:

$$R_{\text{н}}^0 = R_{\text{н}}^0 = \int_0^{\infty} g(\nu) dN = \frac{8\pi}{c^3} \int_0^{\infty} \frac{\bar{n}^3 g(\nu) v^2 dv}{e^{h\nu/kT} - 1}. \quad (7-53)$$

Время жизни фотона τ_{ϕ} есть 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 $\alpha(v)$, т. е. $l_{\phi} = 1/\alpha(v)$. В результате будем иметь:

$$\tau_{\phi} = \frac{1}{g(v)} = \frac{l_{\phi}}{v_{\phi}} = \frac{\bar{n}}{c\alpha(v)}. \quad (7-54)$$

Подставляя значение $g(v)$ из формулы (7-54) в (7-53), находим:

$$R_n^0 = \gamma_r n_i^2 = \frac{8\pi}{c^2} \int_0^{\infty} \frac{\bar{n}^2 \alpha(v) v^2 dv}{e^{h\nu/kT} - 1}. \quad (7-55)$$

Значения n_i , R_n^0 и значения времени жизни τ_{ir} для излучательных переходов зона проводимости — валентная зона для некоторых полупроводников приведены в табл. 7-1.

Таблица 7-1

Излучательная рекомбинация при 300 К некоторых полупроводников

Полупроводник	Ширина запрещенной зоны, эВ	n_i , см ⁻³	R_n^0 , см ⁻³ ·с ⁻¹	τ_{ir}	$\tau_{\text{экс}}$, мкс
Si	1,08	$1,5 \cdot 10^{10}$	$4,5 \cdot 10^5$	4,6 ч	2500
Ge	0,66	$2,4 \cdot 10^{13}$	$2,0 \cdot 10^{13}$	0,61 с	150
GaAs	0,71	$4,3 \cdot 10^{12}$	$2,4 \cdot 10^{14}$	0,009 с	0,37
IaAs	0,31	$1,6 \cdot 10^{15}$	$5,4 \cdot 10^{19}$	15 мкс	0,24
InSb	0,18	$2,0 \cdot 10^{15}$	$1,6 \cdot 10^{22}$	0,62 мкс	0,12
PbS	0,41	$7,1 \cdot 10^{14}$	$2,4 \cdot 10^{19}$	15 мкс	0,21
PbTe	0,32	$4,0 \cdot 10^{15}$	$8,3 \cdot 10^{20}$	2,4 мкс	0,19
PbSe	0,29	$6,2 \cdot 10^{15}$	$1,5 \cdot 10^{21}$	2 мкс	0,25

Из данных табл. 7-1 следует, что время жизни неравновесных носителей заряда в случае межзонной излучательной рекомбинации снижается с уменьшением ширины запрещенной зоны. Рассчитанные значения τ_{ir} относительно близки к экспериментальным только для полупроводников с малой шириной запрещенной зоны. Для материалов с большой шириной запрещенной зоны наблюдается существенное расхождение расчетных данных с экспериментально измеренными. Это обусловлено тем, что у этих полупроводников велика вероятность безызлучательных переходов.

Согласно формулам (7-45) и (7-46) при увеличении степени легирования время жизни неравновесных носителей заряда снижается. Например, для сильно легированного германия ($n_0 \approx 10^{19}$ см⁻³) оно составляет:

$$\tau_r^n = 2 \frac{n_i}{n_0} \tau_{ir} \approx 2 \frac{2,4 \cdot 10^{13}}{10^{19}} 0,61 \approx 3 \cdot 10^{-6} \text{ с} = 3 \text{ мкс.}$$

Казалось бы, что с ростом степени легирования материала должна увеличиваться вероятность межзонной излучательной рекомбинации. Однако с повышением концентрации примеси появляется большое количество дефектов и начинает преобладать рекомбинация через рекомбинационные ловушки. Поэтому для исследования процессов межзонной излучательной рекомбинации следует использовать наиболее чистые полупроводниковые материалы.

7-6. МЕЖЗОННАЯ УДАРНАЯ РЕКОМБИНАЦИЯ

Если происходит столкновение одновременно двух свободных электронов и одной дырки или двух дырок и одного свободного электрона, может иметь место рекомбинация двух из них с передачей энергии третьему носителю заряда, который переходит на более высокий энергетический уровень в соответствующей зоне, становясь «горячим».

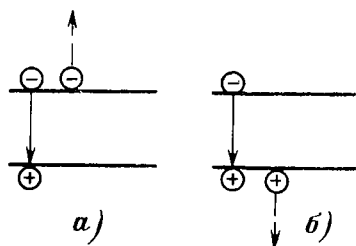


Рис. 7-4. Схема процессов ударной рекомбинации.

«Горячий» носитель в результате столкновений передаст свою энергию фононам и придет в равновесное состояние с решеткой. Этот процесс рекомбинации является процессом обратным ударной ионизации, при котором за счет перехода свободного носителя заряда в зону на более низкий энергетический уровень за счет выделившейся энергии происходит возникновение пары электрон—дырка. На рис. 7-4 схематично изображены процессы ударной рекомбинации с участием в качестве третьего носителя заряда электрона (а) или дырки (б).

Вероятность столкновения пары электрон—дырка со свободным электроном пропорциональна n^2p , а с дыркой — p^2n . Таким образом, уменьшение концентрации свободных электронов или дырок в результате межзонной ударной рекомбинации определится выражением

$$-\frac{dn}{dt} = -\frac{dp}{dt} = \eta^{(n)} n^2 p + \eta^{(p)} p^2 n - \eta^{(n)} n_0^2 p_0 - \eta^{(p)} p_0^2 n_0, \quad (7-56)$$

где $\eta^{(n)}$ и $\eta^{(p)}$ — коэффициенты ударной рекомбинации с участием в качестве третьего носителя заряда электрона или дырки. Два последних члена в выражении (7-56) учитывают интенсивность генерации при термодинамическом равновесии. Следует отметить, что коэффициенты ударной рекомбинации отличаются по размерности от соответствующих коэффициентов в случае рекомбинации другого типа.

Подставив в выражение (7-56) значения n и p , приравняв $\Delta n = \Delta p$ и расположив члены уравнения по степеням Δn , получим:

$$-\frac{d}{dt} \Delta n = -\frac{d}{dt} \Delta p = \Delta n (\eta^{(n)} n_0^2 + 2\eta^{(n)} n_0 p_0 + 2\eta^{(p)} n_0 p_0 + \eta^{(p)} p_0^2) + \Delta n^2 (\eta^{(n)} p_0 + 2\eta^{(n)} n_0 + 2\eta^{(p)} p_0 + \eta^{(p)} n_0) + \Delta n^3 (\eta^{(n)} + \eta^{(p)}). \quad (7-57)$$

Найдем выражение для времени жизни неравновесных носителей заряда в случае межзонной ударной рекомбинации согласно (7-18), считая, что уровень инжекции мал ($\Delta n \ll n_0 + p_0$). Для этого отбросим в (7-57) члены, содержащие Δn во второй и третьей степени, и будем иметь:

$$\tau_{nA} = \tau_{pA} = \frac{\Delta n}{\frac{d}{dt} \Delta n} = \frac{1}{\eta^{(n)} n_0^2 + 2\eta^{(n)} n_0^2 + 2\eta^{(p)} n_0^2 + \eta^{(p)} p_0^2}. \quad (7-58)$$

Рассмотрим частные случаи.

1. Собственный полупроводник ($n_0 = p_0 = n_i$)

$$\tau_{iA} = \frac{1}{3(\eta^{(n)} + \eta^{(p)}) n_i^2} = \frac{1}{3(\eta^{(n)} + \eta^{(p)}) N_c N_v} e^{\frac{E_g}{kT}}. \quad (7-59)$$

Как видно, время жизни неравновесных носителей заряда в собственном полупроводнике при межзонной ударной рекомбинации снижается с ростом температуры и с уменьшением ширины запрещенной зоны.

2. Донорный полупроводник ($n_0 \gg p_0$; $n_0 \gg n_i$)

$$\tau_A^n = \frac{1}{\eta^{(n)} n_0^2}. \quad (7-60)$$

3. Акцепторный полупроводник ($p_0 \gg n_0$; $p_0 \gg n_i$)

$$\tau_A^p = \frac{1}{\eta^{(p)} p_0^2}. \quad (7-61)$$

Из полученных выражений следует, что время жизни неравновесных носителей заряда для собственного полупроводника при межзонной ударной рекомбинации максимально.

В примесном полупроводнике время жизни при этом типе рекомбинации сильно зависит от концентрации основных носителей заряда. При ее увеличении возрастает число столкновений неравновесных электронно-дырочных пар с основными носителями заряда и время жизни снижается. Зависимость τ_A/τ_i от положения уровня Ферми приведена на рис. 7-5.

Экспериментальные исследования показывают, что ударная межзонная рекомбинация наблюдается очень редко. Она становится

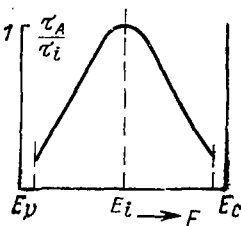


Рис. 7-5. Время жизни при ударной межзонной рекомбинации в зависимости от положения уровня Ферми.

заметной лишь при довольно высоких температурах в полупроводниках с малой шириной запрещенной зоны (например, в InSb).

При высоком уровне возбуждения можно считать, что $\Delta n \gg n_0 + p_0$ и $\Delta p \gg n_0 + p_0$. Для этого случая в уравнении (7-57) можно отбросить два первых слагаемых, т. е. члены, содержащие Δn и Δn^2 . В результате будем иметь:

$$-\frac{d}{dt} \Delta n = -\frac{d}{dt} \Delta p = \Delta n^3 (\eta^{(n)} + \eta^{(p)}). \quad (7-62)$$

Мгновенное время жизни при этом:

$$\tau_{\text{мгн}} = \frac{\Delta n}{\frac{d}{dt} \Delta n} = \frac{1}{(\eta^{(n)} + \eta^{(p)}) \Delta n^2}. \quad (7-63)$$

Из (7-63) следует, что мгновенное время жизни для ударной межзонной рекомбинации при высоком уровне возбуждения очень сильно зависит от концентрации избыточных носителей заряда.

7-7. РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА ЧЕРЕЗ РЕКОМБИНАЦИОННЫЕ ЛОВУШКИ

Реальные полупроводниковые материалы содержат обычно примеси нескольких типов, каждая из которых может создавать один или несколько уровней в запрещенной зоне полупроводника. Дефекты решетки, обычно нейтральные в состоянии термодинамического равновесия и способные захватывать подвижные носители заряда одного знака и освобождать их, называются ловушками захвата.

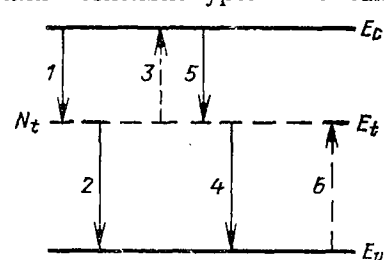


Рис. 7-6. Переходы электронов, связанные с рекомбинацией через локальные центры.

Мы ограничимся рассмотрением простейшего случая, когда в полупроводнике имеется один тип ловушек, создающих единичный энергетический уровень.

Пусть концентрация ловушек равна N_t , а их энергетический уровень E_t лежит вблизи центра запрещенной зоны. Процесс рекомбинации через эти ловушки проиллюстрирован на рис. 7-6.

В полупроводнике донорного типа, имеющем большое число свободных электронов, рекомбинация происходит следующим образом. Нейтральная ловушка захватывает электрон из зоны проводимости, приобретая отрицательный заряд (переход 1), а затем электрон с уровня ловушки переходит в валентную зону, что эквивалентно захвату отрицательно заряженной ловушкой дырки (переход 2). При снижении концентрации свободных дырок вероят-

ность второго процесса уменьшается и возможна эмиссия электрона, т. е. обратный переход электрона с ловушки в зону проводимости (переход 3).

Если материал содержит большое количество свободных дырок, они интенсивно захватываются ловушками, т. е. электроны переходят с ловушки в валентную зону (переход 4), а затем происходит захват электрона из зоны проводимости (переход 5). При малой концентрации электронов проводимости возможен процесс освобождения дырки, т. е. возвращение дырки в валентную зону (переход 6).

Перейдем к количественному описанию рассмотренных процессов. Если через $f_t = f(E_t)$ обозначим вероятность заполнения ловушки электроном, то $1 - f_t$ будет вероятность того, что на ловушке нет электрона, т. е. она свободна. Тогда интенсивность захвата ловушками электронов $R_{\text{зах}}^n$, которая пропорциональна концентрации электронов в зоне проводимости n и числу свободных мест на уровне ловушек $N_t(1 - f_t)$, равна:

$$R_{\text{зах}}^n = \gamma_n n N_t (1 - f_t), \quad (7-64)$$

где γ_n — коэффициент захвата неравновесных электронов ловушками.

Количество электронов $R_{\text{воз}}^n$, возвращающихся с ловушек в зону проводимости в результате эмиссии, пропорционально концентрации электронов на ловушках $N_t f_t$:

$$R_{\text{воз}}^n = \beta_n f_t N_t, \quad (7-65)$$

где β_n — коэффициент ионизации электронов с ловушек.

Изменение концентрации неравновесных электронов в зоне проводимости определяется разницей интенсивностей процессов захвата $R_{\text{зах}}^n$ и освобождения $R_{\text{воз}}^n$:

$$-\frac{dn}{dt} = R_{\text{зах}}^n - R_{\text{воз}}^n = \gamma_n n N_t (1 - f_t) - \beta_n N_t f_t. \quad (7-66)$$

Чтобы определить связь между коэффициентами γ_n и β_n , рассмотрим состояние термодинамического равновесия. В состоянии термодинамического равновесия вероятность заполнения уровня ловушки электроном есть

$$f_{0t} = f(E_t) = \frac{1}{e^{(E_t - F)/kT} + 1}.$$

Если в неравновесном состоянии не наступает вырождение, то неравновесные носители заряда большую часть своего времени жизни физически неотличимы от равновесных. Поэтому можно считать, что и равновесные носители заряда характеризуются также коэффициентами захвата γ_n и ионизации β_n . На основании

равенства (7-66) в случае термодинамического равновесия, для которого $\frac{dn_0}{dt} = 0$, найдем, что

$$\begin{aligned} \beta_n &= \gamma_n n_0 \left(\frac{1}{f_{0t}} - 1 \right) = \gamma_n N_c e^{-(E_c - F)/kT} e^{(E_t - F)/kT} = \\ &= \gamma_n N_c e^{-(E_c - E_t)/kT} = \gamma_n n_1. \end{aligned} \quad (7-67)$$

Здесь

$$n_1 = N_c e^{-(E_c - E_t)/kT} \quad (7-68)$$

есть величина равновесной концентрации электронов в зоне проводимости, когда уровень Ферми совпадает с уровнем ловушки.

Используя равенства (7-67), преобразуем уравнение (7-66) к виду

$$-\frac{dn}{dt} = \gamma_n N_t [n(1 - f_t) - n_1 f_t]. \quad (7-69)$$

Это выражение определяет скорость изменения концентрации неравновесных электронов в зоне проводимости в результате процессов захвата электронов ловушками и освобождения их из ловушек в зону проводимости.

Скорость изменения концентрации неравновесных дырок в валентной зоне запишется следующим образом:

$$-\frac{dp}{dt} = \gamma_p p N_t f_t - \beta_p N_t (1 - f_t). \quad (7-70)$$

Здесь первое слагаемое определяет количество дырок, захватываемых ловушкой, а второе -- число дырок, возвращающихся из ловушек в валентную зону. Величины γ_p и β_p представляют собой соответственно коэффициенты захвата и ионизации для дырок.

Из условия термодинамического равновесия $\left(\frac{dp_0}{dt} = 0 \right)$ найдем:

$$\begin{aligned} \beta_p &= \gamma_p p_0 \frac{f_{0t}}{1 - f_{0t}} = \gamma_p N_v e^{(E_v - F)/kT} e^{-(E_t - F)/kT} = \\ &= \gamma_p N_v e^{(E_v - E_t)/kT} = \gamma_p p_1, \end{aligned} \quad (7-71)$$

где

$$p_1 = N_v e^{(E_v - E_t)/kT} \quad (7-72)$$

есть равновесная концентрация дырок в валентной зоне, когда уровень Ферми совпадает с уровнем ловушек.

Подставляя выражение (7-71) в формулу (7-70), имеем:

$$-\frac{dp}{dt} = \gamma_p N_t [p f_t - p_1 (1 - f_t)]. \quad (7-73)$$

Это выражение определяет скорость изменения концентрации неравновесных дырок в валентной зоне в результате процессов захвата дырок ловушками и освобождения их из ловушек в валентную зону.

Используя уравнения (7-69) и (7-73), находим изменение количества электронов, находящихся на ловушках. Оно определяется изменением концентраций неравновесных электронов в зоне проводимости и дырок в валентной зоне:

$$N_t \frac{df_t}{dt} = \frac{dn}{dt} - \frac{dp}{dt} = \\ = \gamma_p N_t [pf_t - p_1(1-f_t)] - \gamma_n N_t [n(1-f_t) - n_1 f_t]. \quad (7-74)$$

Запишем условие локальной электронейтральности для однородного полупроводника, содержащего некоторое количество ионов донорной примеси N_d^+ . Для равновесного случая условие электронейтральности имеет вид:

$$n_0 + N_t f_{0t} = p_0 + N_d^+. \quad (7-75)$$

Будем считать, что число ионов донорной примеси не меняется и в неравновесном состоянии. Тогда

$$n + N_t f_t = p + N_d^+. \quad (7-76)$$

Вычитая (7-75) из (7-76), получаем соотношение для избыточной концентрации носителей заряда:

$$n - n_0 + N_t (f_t - f_{0t}) = p - p_0$$

или

$$\Delta n + N_t \Delta f_t = \Delta p. \quad (7-77)$$

Если концентрация ловушек в полупроводнике мала по сравнению с плотностью избыточных носителей заряда, т. е. $N_t \ll \Delta n$, то количеством электронов $N_t \Delta f_t$, оседающих на ловушках, можно пренебречь. При этом условии избыточная концентрация электронов равна избыточной концентрации дырок, т. е.

$$\Delta n = \Delta p. \quad (7-78)$$

В этом случае $\frac{dn}{dt} = \frac{dp}{dt}$, и из уравнения (7-74), приравняв его нулю, найдем функцию распределения для электронов, находящихся на ловушках:

$$f_t = \frac{\gamma_n n + \gamma_p p_1}{\gamma_n (n + n_1) + \gamma_p (p + p_1)}. \quad (7-79)$$

Подставляя найденное значение f_t в выражение (7-69), получаем:

$$-\frac{dn}{dt} = \frac{\gamma_n \gamma_p N_t (np - n_1 p_1)}{\gamma_n (n + n_1) + \gamma_p (p + p_1)}. \quad (7-80)$$

На основании этого соотношения можно определить время жизни неравновесных носителей заряда при рекомбинации через рекомбинационные ловушки. Учитывая, что

$$n_1 p_1 = n_i^2 = n_0 p_0, \quad (7-81)$$

в результате будем иметь:

$$\begin{aligned} \tau &= -\frac{\Delta n}{\frac{dn}{dt}} = \frac{\gamma_n (n_0 + n_1 + \Delta n) + \gamma_p (p_0 + p_1 + \Delta p)}{\gamma_n \gamma_p N_t (n_0 + p_0 + \Delta n)} = \\ &= \frac{1}{\gamma_p N_t} \frac{n_0 + n_1 + \Delta n}{n_0 + p_0 + \Delta n} + \frac{1}{\gamma_n N_t} \frac{p_0 + p_1 + \Delta p}{n_0 + p_0 + \Delta n}. \end{aligned} \quad (7-82)$$

Введем обозначения:

$$\tau_{p0} = \frac{1}{\gamma_p N_t}; \quad \tau_{n0} = \frac{1}{\gamma_n N_t}. \quad (7-83)$$

С учетом равенства (7-83) формула (7-81) может быть записана в виде

$$\tau = \tau_{p0} \frac{n_0 + n_1 + \Delta n}{n_0 + p_0 + \Delta n} + \tau_{n0} \frac{p_0 + p_1 + \Delta p}{n_0 + p_0 + \Delta p}. \quad (7-84)$$

В случае малого уровня возбуждения, когда $\Delta n \ll n_0 + p_0$, время жизни неравновесных носителей заряда в зависимости от положения уровня Ферми при постоянной температуре, т. е. от степени легирования материала, примет вид:

$$\tau = \tau_{p0} \frac{n_0 + n_1}{n_0 + p_0} + \tau_{n0} \frac{p_0 + p_1}{n_0 + p_0}. \quad (7-85)$$

Итак, при малом уровне возбуждения время жизни электронно-дырочной пары не зависит от концентрации избыточных носителей заряда, а определяется значениями равновесных концентраций электронов и дырок и положением энергетического уровня рекомбинационной ловушки, задающим величины n_1 и p_1 .

На основании формулы (7-85) для зависимости $\ln \tau$ от положения уровня Ферми можно выделить четыре основные области (рис. 7-7).

Первая область. Сильно легированный полупроводник n -типа. Уровень Ферми расположен ниже дна зоны проводимости, но выше энергетического уровня рекомбинационной ловушки, т. е. $E_t < F < E_c$. В этом случае справедливы следующие соотношения концентраций носителей заряда:

$$n_0 \gg p_0; \quad n_0 \gg n_1; \quad n_0 \gg p_1.$$

Учитывая эти неравенства, получаем:

$$\tau = \tau_{p0} = \frac{1}{\gamma_p N_t}. \quad (7-86)$$

Таким образом, в сильно легированном полупроводнике *n*-типа величина времени жизни пары неравновесных электронов и дырок постоянна и равна $\tau = \tau_{p0}$, т. е. определяется временем жизни неосновных равновесных носителей заряда — дырок. В этом случае скорость процесса рекомбинации определяется числом актов захвата дырок рекомбинационными ловушками, занятыми электронами. Действительно, как показывает эксперимент (рис. 7-8), время жизни дырок в сильно легированном электронном германии постоянно.

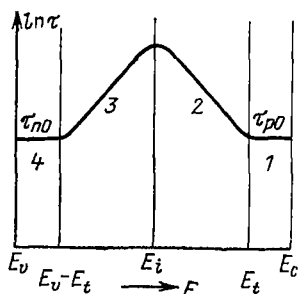


Рис. 7-7. Изменение времени жизни в зависимости от положения уровня Ферми.

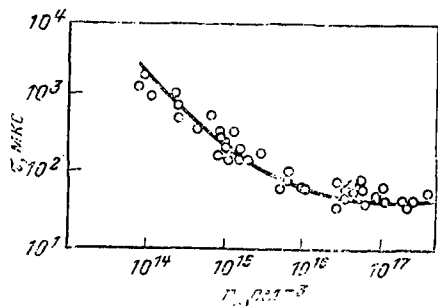


Рис. 7-8. Зависимость времени жизни дырок в электронном германии от равновесной концентрации электронов n_0 .

Проведем анализ формулы (7-86). Это равенство выражает собой тот факт, что каждая дырка в материале *n*-типа, перемещаясь в валентной зоне, обладает определенной вероятностью быть захваченной отрицательно заряженной рекомбинационной ловушкой. Для количественной оценки этого процесса используем эффективное сечение захвата дырки σ_p . Согласно выражению (5-44) вероятность захвата носителя заряда пропорциональна поперечному сечению захвата центра и тепловой скорости носителей v_0 , а поэтому

$$\gamma_p = \sigma_p v_0. \quad (7-87)$$

Следовательно, время жизни дырки в электронном полупроводнике равно

$$\tau = \tau_{p0} = \frac{1}{\gamma_p N_t} = \frac{1}{\sigma_p N_t v_0}. \quad (7-88)$$

Оно уменьшается при увеличении концентрации рекомбинационных ловушек и температуры, приводящей к росту v_0 .

Вторая область. Полупроводник *n*-типа легирован слабо, так что уровень Ферми лежит ниже энергетического уровня ловушек, но выше середины запрещенной зоны, т. е. $E_t < F < E_i$. В этом случае имеют место следующие соотношения для концентраций носителей заряда:

$$n_0 \gg p_0; \quad n_0 \gg p_1; \quad n_0 < n_1.$$

Соответственно выражение для τ приобретает вид:

$$\tau \approx \tau_{p0} \frac{n_1}{n_0} = \tau_{p0} e^{(E_t - F)/kT}. \quad (7-89)$$

Из соотношения (7-89) следует, что по мере понижения уровня Ферми время жизни τ растет по экспоненциальному закону. При этом чем ниже опустается уровень Ферми, тем меньше степень заполнения ловушек электронами, поэтому вероятность захвата ловушкой дырки снижается, что ведет к увеличению времени жизни дырки. Это подтверждается данными, представленными на рис. 7-7.

Третья область. В слабо легированном полупроводнике p -типа положение уровня Ферми определяется неравенством $E_v - E_t < F < E_t$, что дает следующие соотношения для концентраций носителей заряда:

$$p_0 \gg n_0; \quad p_0 \gg p_1; \quad n_1 > p_0.$$

На основании этого из (7-85) получаем формулу для времени жизни:

$$\tau \approx \tau_{p0} \frac{n_1}{p_0} = \tau_{p0} \frac{N_c}{N_v} e^{-\frac{E_c + E_v - E_t - F}{kT}}. \quad (7-90)$$

Видно, что для акцепторного полупроводника по мере понижения уровня Ферми от середины запрещенной зоны в результате легирования время жизни уменьшается по экспоненциальному закону. Это происходит потому, что в слабо легированном полупроводнике p -типа почти все ловушки свободные и охотно захватывают электроны из зоны проводимости. С увеличением количества дырок проводимости растет вероятность их рекомбинации с электронами, попавшими на ловушки. При этом интенсивность процесса освобождения электронов с ловушек снижается и время жизни уменьшается (рис. 7-7).

Четвертая область. Сильно легированный материал p -типа, для которого положение уровня Ферми определяется равенством вида $E_v < F < E_v - E_t$.

В этом случае

$$p_0 \gg n_0; \quad p_0 \gg p_1; \quad p_0 \gg n_1,$$

а из соотношения (7-85) следует:

$$\tau = \tau_{n0} = \frac{1}{\gamma_n N_t} = \frac{1}{\sigma_n N_t v_0}, \quad (7-91)$$

где σ_n — эффективное сечение захвата электрона.

Таким образом, в сильно легированном полупроводнике p -типа время жизни пары электрон — дырка постоянно равно $\tau = \tau_{n0}$ и определяется временем жизни равновесных электронов — неосновных носителей заряда и не зависит от положения уровня Ферми.

У такого полупроводника все ловушки свободны от электронов, и каждый электрон, захваченный ловушкой, немедленно рекомбинирует с дыркой, количество которых в валентной зоне очень велико. Процесс обратного переброса электронов с ловушек в зону проводимости практически полностью отсутствует и не влияет на время жизни.

Рассмотрим теперь случай большого уровня возбуждения избыточных носителей заряда, когда $\Delta n \gg n_0$; $\Delta p \gg p_0$, а значит $\Delta n \gg n_1$ и $\Delta p \gg p_1$. Из равенства (7-84) получаем:

$$\tau_{\infty} \approx \tau_{p0} + \tau_{n0} = \frac{\gamma_n + \gamma_p}{\gamma_n \gamma_p N_t}. \quad (7-92)$$

Как видно из этого соотношения, *при большом уровне возбуждения время жизни неравновесных электронов и дырок τ_{∞} при рекомбинации через рекомбинационные ловушки не зависит от концентраций электронов и дырок, а определяется лишь количеством и свойствами ловушек.*

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

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

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

По мере дальнейшего повышения температуры при переходе к области полной ионизации донорной примеси уровень Ферми по-

нижается и достигает уровня рекомбинационной ловушки. На этом участке концентрация электронов n_0 постоянна, а концентрация $n_1 = N_c e^{-(E_c - E_l)/kT}$ растет с температурой экспоненциально, но она по-прежнему значительно меньше равновесной концентрации n_0 . Кроме того, n_0 также больше концентраций дырок p_0 и p_1 . На этом участке согласно выражению (7-85) время жизни также постоянно и равно τ_{p0} (участок 1 на рис. 7-9, в).

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

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

т. е. время жизни увеличивается с ростом температуры. При этом оно увеличивается вплоть до температуры, при которой наступает собственная проводимость (участок 2 на рис. 7-9, в). Рост τ в этой области температур происходит потому, что с увеличением температуры

имеет место интенсивная эмиссия электронов с уровня рекомбинационных ловушек. В этом случае уменьшается заполнение ловушек электронами, а следовательно, увеличивается время жизни неосновных носителей — дырок, которое и определяет измеряемое время τ .

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

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

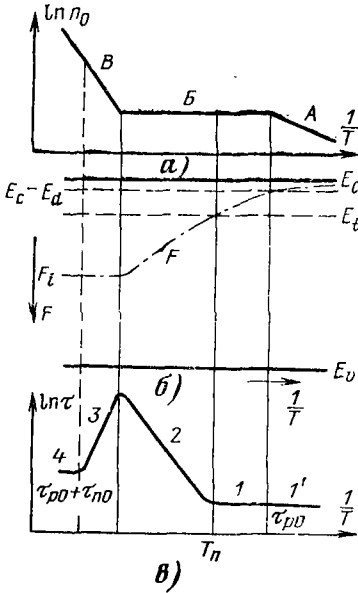


Рис. 7-9. Зависимость равновесной концентрации (а), уровня Ферми (б) и времени жизни (в) от температуры для донорного полупроводника.

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

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

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

$$\tau \approx \frac{\tau_{n0}}{2} \left(\frac{N_v}{N_c} \right)^{1/2} e^{\frac{E_c + E_v}{2} - E_t \over kT}. \quad (7-96)$$

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

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

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

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

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

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

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

$$n_0 = N_c e^{- (E_c - E_t) / k T_{II}}, \quad (7-99)$$

откуда

$$E_c - E_t = k T_{II} \ln \frac{N_c}{n_0}. \quad (7-100)$$

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

7.9. ЛОВУШКИ ЗАХВАТА И РЕКОМБИНАЦИОННЫЕ ЛОВУШКИ

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

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

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

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

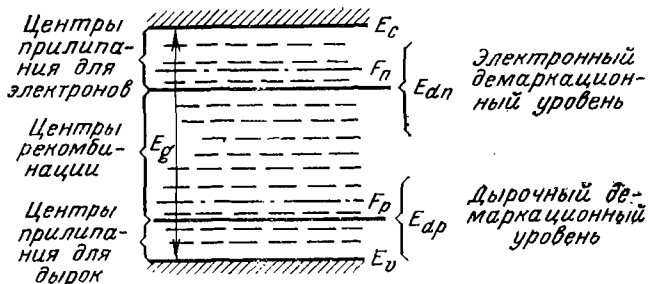


Рис. 7-10. Полупроводник, содержащий ловушки захвата и рекомбинационные ловушки.

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

Учитывая это, для коэффициента k_n получаем следующее выражение:

$$k_n = \frac{\gamma_p p}{\gamma_n n_1} = \frac{\gamma_p p}{\gamma_n N_{ce}^{- (E_c - E_t)/kT}} \quad (7-101)$$

Ловушки, для которых вероятность рекомбинации больше вероятности теплового освобождения, т. е. $k_n > 1$, будут являться рекомбинационными ловушками, а ловушки с $k_n < 1$ относятся к ловушкам захвата. Энергетический уровень ловушки, который характеризуется значением $k_n = 1$, т. е. для которого вероятность

рекомбинации и теплового освобождения одинакова, называется электронным демаркационным уровнем E_{dn} . Его положение можно найти из условия

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

Отсюда будем иметь:

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

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

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

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

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

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

или

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

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

Глава восьмая

ДИФFUЗИЯ И ДРЕЙФ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА

8-1. УРАВНЕНИЕ НЕПРЕРЫВНОСТИ

Рассмотрим полупроводник, в котором в результате воздействия каких-либо внешних возбуждающих факторов возникли неравновесные носители заряда. В этом случае концентрации неравновесных электронов и дырок являются функциями пространственных координат и времени, $n(x, y, z, t)$ и $p(x, y, z, t)$, и скорости их изменения могут быть выражены с помощью уравнения непрерывности.

Вывод уравнения непрерывности проведем для одномерного однородного образца, в котором концентрация электронов изменяется только в направлении оси x , как это представлено на рис. 8-1. Выделим в полупроводнике слой толщиной dx и сечением 1 см^2 . Объем этого слоя будет численно равен dx . 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:

$$n(x, t + dt) dx - n(x, t) dx = \frac{\partial n}{\partial t} dx dt. \quad (8-1)$$

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 заряда как в результате тепловых процессов, так и в результате внешних воздействий. К внешним воздействиям относятся действие света, ядерных частиц, неравновесного излучения, испускаемого самим полупроводником в его объеме, и т. п. Обозначим скорость этих двух типов генерации соответственно через G_0 и G . Тогда полная скорость генерации носителей заряда будет равна $G_0 + G$.

Будем считать, что в рассматриваемом случае возбуждение электронов, происходит за счет поглощения света полупроводни-

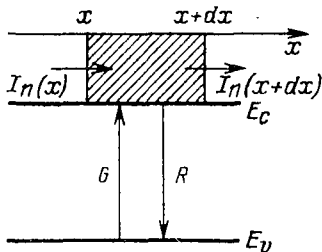


Рис. 8-1. К выводу уравнения непрерывности.

ком, и пусть число пар электрон—дырка, создаваемых светом за 1 с в 1 см³, будет равно G . Тогда световая генерация в объеме dx за время dt создает электроны в количестве

$$G dx dt. \quad (8-2)$$

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 рекомбинации при тепловом равновесии, R — скорость рекомбинации при наличии внешнего воздействия. В случае линейной рекомбинации, когда время жизни неравновесных электронов τ_n не зависит от концентрации электронов, изменение числа электронов вследствие рекомбинации R в объеме dx за время dt согласно выражению (7-18) составит:

$$R dx dt = - \left(\frac{dn}{dt} \right)_r dx dt = - \frac{n - n_0}{\tau_n} dx dt. \quad (8-3)$$

3. Изменение количества носителей заряда в объеме dx может быть обусловлено, кроме того, процессами диффузии и дрейфа. Если $I_n(x, t)$ — число электронов, проходящих через 1 см² поверхности за 1 с, то за время dt через границу слоя x в рассматриваемый объем войдут электроны в количестве $I_n(x, t) dx$, 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:

$$I_n(x, t) dt - I_n(x + dx, t) dt = - \frac{\partial I_n}{\partial x} dx dt. \quad (8-4)$$

Полное изменение числа электронов в объеме dx за время dt составит величину

$$\frac{\partial n}{\partial t} dx dt = G dx dt - \frac{\partial I_n}{\partial x} dx dt - \frac{n - n_0}{\tau_n} dx dt, \quad (8-5)$$

откуда имеем:

$$\frac{\partial n}{\partial t} = G - \frac{\partial I_n}{\partial x} - \frac{n - n_0}{\tau_n}. \quad (8-6)$$

Уравнение (8-6) называют уравнением непрерывности для электронов.

Аналогично уравнение непрерывности для дырок запишется в виде

$$\frac{\partial p}{\partial t} = G - \frac{\partial I_p}{\partial x} - \frac{p - p_0}{\tau_p}. \quad (8-7)$$

Потоки электронов I_n и дырок I_p можно выразить через плотности тока, которые с учетом знака носителя равны:

$$\left. \begin{aligned} j_n &= -eI_n; \\ j_p &= eI_p. \end{aligned} \right\} \quad (8-8)$$

Когда концентрации носителей заряда являются функциями координат (x, y, z) , уравнения непрерывности запишутся через соответствующие токи следующим образом:

$$\frac{\partial n}{\partial t} = G - \frac{1}{e} \operatorname{div} \mathbf{j}_n = \frac{n - n_0}{\tau_n}; \quad (8-9)$$

$$\frac{\partial p}{\partial t} = G - \frac{1}{e} \operatorname{div} \mathbf{j}_p = \frac{p - p_0}{\tau_p}, \quad (8-10)$$

где

$$\operatorname{div} \mathbf{j} = \frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z}.$$

Пусть в точке $x = 0$ имеется источник неравновесных носителей заряда. Рассмотрим стационарный случай. Концентрации электронов n и дырок p не изменяются во времени, т. е. $\frac{\partial n}{\partial t} = \frac{\partial p}{\partial t} = 0$.

Для стационарного условия уравнения непрерывности в одномерном случае принимают вид:

$$-\frac{1}{e} \frac{\partial j_n}{\partial x} = G - \frac{n - n_0}{\tau_n}; \quad (8-11)$$

$$\frac{1}{e} \frac{\partial j_p}{\partial x} = G - \frac{p - p_0}{\tau_p}. \quad (8-12)$$

Уравнения (8-11) и (8-12) выражают собой условия сохранения числа носителей заряда, т. е. в стационарном состоянии поток носителей заряда, вытекающих из объема, равен количеству носителей заряда, созданных внешним возбуждением, за вычетом носителей заряда, прорекомбинировавших в этом объеме.

8-2. ДИФFUЗИОННЫЙ И ДРЕЙФОВЫЙ ТОКИ

Рассмотрим теперь неоднородный полупроводник, в котором концентрации электронов $n(\mathbf{r})$ и дырок $p(\mathbf{r})$ изменяются от точки к точке. По этой причине в неоднородном полупроводнике должен возникнуть диффузионный ток электронов и дырок, который будет определяться диффузией носителей заряда из областей, где их концентрация больше, в области с меньшей концентрацией.

Пусть в полупроводнике концентрация носителей заряда возрастает в направлении оси x , как это изображено на рис. 8-2. Проведем через точку x плоскость, перпендикулярную к оси x , и рассмотрим движение носителей заряда в слоях 1 и 2 толщиной dx , расположенных справа и слева от этой плоскости. В результате хаотического движения носители заряда уйдут из слоя 1, 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 электронов будет

больше прямого. Если $n\left(x - \frac{dx}{2}\right)$ — средняя концентрация электронов в слое 1, а $n\left(x + \frac{dx}{2}\right)$ — в слое 2, то разность концентраций электронов в этих слоях будет равна:

$$n\left(x - \frac{dx}{2}\right) - n\left(x + \frac{dx}{2}\right) = -\frac{dn}{dx} dx. \quad (8-13)$$

Согласно выражению (8-13) разность концентраций электронов пропорциональна градиенту их концентрации, поэтому и поток электронов I_n , возникающий в результате их диффузии в направлении x , будет пропорционален градиенту концентрации электронов в этом направлении. Его можно записать уравнением

$$I_n = -D_n \frac{dn}{dx}, \quad (8-14)$$

где D_n — коэффициент диффузии электронов.

Аналогично диффузионный поток дырок

$$I_p = -D_p \frac{dp}{dx}, \quad (8-15)$$

где D_p — коэффициент диффузии дырок.

Потоки электронов и дырок, как следует из уравнений (8-14) и (8-15), текут в сторону меньших концентраций носителей заряда.

Диффузионным потокам носителей заряда соответствуют диффузионные токи электронов $j_{n \text{ диф}}$ и дырок $j_{p \text{ диф}}$:

$$j_{n \text{ диф}} = eD_n \frac{dn}{dx}; \quad (8-16)$$

$$j_{p \text{ диф}} = -eD_p \frac{dp}{dx}. \quad (8-17)$$

В том случае, если n является функцией координат (x, y, z) , диффузионный ток в векторной форме имеет вид для электронов

$$\mathbf{j}_{n \text{ диф}} = eD_n \text{grad } n(\mathbf{r}) \quad (8-18)$$

и для дырок

$$\mathbf{j}_{p \text{ диф}} = -eD_p \text{grad } p(\mathbf{r}). \quad (8-19)$$

Диффузионный ток, возникший из-за наличия градиента концентрации носителей заряда, приведет к пространственному разделению зарядов. Разделение зарядов породит статическое электрическое поле, которое создаст дрейфовые токи электронов и дырок. При термодинамическом равновесии в каждой точке полупроводника дрейфовый ток будет уравновешивать диффузионный ток, поэтому суммарный ток 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 слабое и не изменяет характера движения носителей заряда, то дрейфовые составляющие плотности тока запишутся на основании закона Ома в виде

$$\left. \begin{aligned} j_{n \text{ др}} &= en\mu_n \mathcal{E}; \\ j_{p \text{ др}} &= ep\mu_p \mathcal{E}. \end{aligned} \right\} \quad (8-20)$$

Полный ток будет складываться из диффузионного и дрейфового токов. Для электронов и дырок он будет равен:

$$j_n = j_{n \text{ др}} + j_{n \text{ диф}} = en\mu_n \mathcal{E} + eD_n \frac{dn}{dx}; \quad (8-21)$$

$$j_p = j_{p \text{ др}} + j_{p \text{ диф}} = ep\mu_p \mathcal{E} - eD_p \frac{dp}{dx}. \quad (8-22)$$

Таким образом, плотность общего тока j в любой точке неоднородного полупроводника в любой момент времени будет определяться уравнением

$$j = j_n + j_p = e(n\mu_n + p\mu_p) \mathcal{E} + e \left(D_n \frac{dn}{dx} - D_p \frac{dp}{dx} \right). \quad (8-23)$$

Необходимо отметить, что диффузионный ток существует только в полупроводниках. Это происходит потому, что в полупроводниках концентрации электронов и дырок могут изменяться в широких пределах. В металлах концентрация электронов практически постоянна.

8-3. СООТНОШЕНИЕ ЭЙНШТЕЙНА

В неоднородном полупроводнике при термодинамическом равновесии ток равен нулю, т. е. $j = j_n + j_p = 0$: В этом случае токи проводимости уравновешивают диффузионные токи и на основании (8-21) для электронов можно записать:

$$n\mu_n \mathcal{E}_{\text{ст}} = -D_n \frac{dn}{dx}. \quad (8-24)$$

Поскольку в полупроводнике имеется статическое электрическое поле $\mathcal{E}_{\text{ст}}$, то электроны, находящиеся в этом поле, будут обладать потенциальной энергией $U = -e\varphi$. Поэтому при отсутствии вырождения концентрация электронов в зоне проводимости будет удовлетворять соотношению Больцмана вида

$$n = N_c e^{-(E_c + U - F)/kT} = n_0 e^{e\varphi/kT}, \quad (8-25)$$

где $n_0 := N_c e^{-(E_c - F)/kT}$ — равновесная концентрация электронов; φ — электростатический потенциал.

Учитывая, что $\mathcal{E}_{ст} := -\frac{d\varphi}{dx}$, и подставляя значения n и $\frac{dn}{dx}$ в уравнение (8-24), получаем:

$$- \mu_n n_0 e^{e\varphi/kT} \frac{d\varphi}{dx} = - D_n \frac{en_0}{kT} e^{e\varphi/kT} \frac{d\varphi}{dx}, \quad (8-26)$$

откуда для электронов будем иметь:

$$\frac{\mu_n}{D_n} = \frac{e}{kT}. \quad (8-27)$$

Аналогично для дырок

$$\frac{\mu_p}{D_p} = \frac{e}{kT}. \quad (8-28)$$

Соотношение, связывающее коэффициент диффузии носителей заряда, подчиняющихся статистике Максвелла, 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 заряда за малое время, намного меньше их времени жизни, обмениваясь энергией с решеткой, приходят в тепловое равновесие с решеткой, и их распределение по энергиям при отсутствии вырождения не отличается от распределения равновесных носителей заряда.

8-4. ДИФФУЗИЯ И ДРЕЙФ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА В СЛУЧАЕ МОНОПОЛЯРНОЙ ПРОВОДИМОСТИ

Рассмотрим диффузию и дрейф неравновесных носителей заряда в случае монополярной проводимости, когда свободные носители заряда возникают только в результате возбуждения их с уровней примеси. Пусть часть достаточно длинного однородного полупроводника, например n -типа, освещается слабопоглощающим светом. В освещенной части полупроводника при $x < 0$ (рис. 8-3, а) имеет место однородная генерация электронов в результате перевода их светом с донорных уровней в зону проводимости. В этом случае концентрация неравновесных электронов n определяется концентрацией равновесных электронов n_0 и избыточных электронов Δn , возбужденных с уровней донорной примеси. И если нет ловушек захвата, то число избыточных электронов Δn равно количеству положительных ионов донорной примеси ΔN_d^+ , возбужденных светом, т. е.

$$\Delta n = \Delta N_d^+. \quad (8-29)$$

Это равенство выражает собой условие электронейтральности в случае монополярной проводимости.

Поскольку концентрация электронов в освещенной части образца больше, чем в неосвещенной, то неравновесные электроны из освещенной части образца будут диффундировать в неосвещенную часть (рис. 8-3, б). Вследствие этого нарушится электронейтральность в некоторой области полупроводника и возникнет объемный заряд, а следовательно, и электрическое поле.

В неосвещенной части образца, куда в результате диффузии пришли избыточные электроны, объемный заряд ρ будет отрицательный, а в освещенной области, откуда они ушли — положительный, обусловленный ионами донорной примеси (рис. 8-3, в). Эти заряды создадут статическое электрическое поле напряженностью $\mathcal{E}_{ст}$, направленное так, что оно будет препятствовать диффузии неравновесных электронов (рис. 8-3, г). Таким образом, возникновение диффузионного тока $j_{диф} = eD_n \frac{dn}{dx}$ приводит к появлению статического электрического поля напряженностью $\mathcal{E}_{ст}$, а следовательно, и тока проводимости $j_{др} = en\mu_n \mathcal{E}_{ст}$, направленного против тока диффузии. В стационарном состоянии полный ток равен нулю:

$$j = j_{др} + j_{диф} = en\mu_n \mathcal{E}_{ст} + eD_n \frac{dn}{dx} = 0. \quad (8-30)$$

Из равенства (8-30) 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:

$$\mathcal{E}_{ст} = - \frac{D_n}{\mu_n n} \frac{d \Delta n}{dx} = - \frac{kT}{e(n_0 + \Delta n)} \frac{d \Delta n}{dx} \approx - \frac{kT}{en_0} \frac{d \Delta n}{dx}. \quad (8-31)$$

Продифференцировав (8-31), получим:

$$\frac{d \mathcal{E}_{ст}}{dx} = - \frac{kT}{en_0} \frac{d^2 \Delta n}{dx^2}. \quad (8-32)$$

Величину $\frac{d \mathcal{E}_{ст}}{dx}$ можно найти, воспользовавшись уравнением Пуассона:

$$\frac{d \mathcal{E}_{ст}}{dx} = \frac{1}{\epsilon \epsilon_0} \rho = - \frac{e}{\epsilon \epsilon_0} \Delta n, \quad (8-33)$$

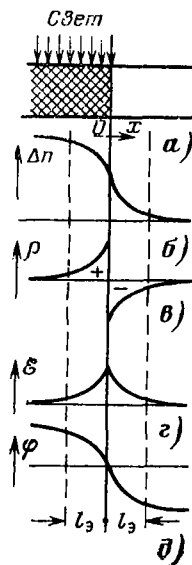


Рис. 8-3. Распределение концентрации (б), объемного заряда (в), электрического поля (г), потенциала (д) при частичном освещении донорного полупроводника (а).

где $\rho = -e\Delta n$ — объемный заряд в неосвещенной части образца.

Из равенств (8-32) и (8-33) следует, что

$$\frac{d^2 \Delta n}{dx^2} - \frac{e^2 n_0}{\epsilon \epsilon_0 k T} \Delta n = 0. \quad (8-34)$$

Если ввести обозначение

$$l_3 = \sqrt{\frac{\epsilon \epsilon_0 k T}{e^2 n_0}}, \quad (8-35)$$

то уравнение (8-34) запишется следующим образом:

$$\frac{d^2 \Delta n}{dx^2} - \frac{\Delta n}{l_3^2} = 0. \quad (8-36)$$

Общее решение этого уравнения имеет вид:

$$\Delta n = C_1 e^{\alpha_1 x} + C_2 e^{\alpha_2 x}, \quad (8-37)$$

где C_1 и C_2 — постоянные, определяемые из граничных условий; α_1 , α_2 — корни характеристического уравнения, равные:

$$\alpha_1 = \frac{1}{l_3}; \quad \alpha_2 = -\frac{1}{l_3}. \quad (8-38)$$

Для неосвещенной области полупроводника, в которой концентрация избыточных электронов уменьшается по мере удаления от освещенной части образца, имеет смысл только член решения с отрицательным показателем степени, поэтому

$$\Delta n = \Delta n(0) e^{-x/l_3}. \quad (8-39)$$

Таким образом, в случае монополярной проводимости концентрация избыточных носителей заряда в неосвещенной части образца по мере удаления от освещенной области уменьшается по экспоненциальному закону с постоянной спада l_3 , называемой радиусом (длиной) экранирования или дебаевским радиусом. Длина экранирования, как следует из (8-35), зависит от концентрации основных носителей заряда, поэтому ее значение может изменяться в широких пределах в зависимости от удельной проводимости полупроводника. Например, для таких полупроводников, как германий и кремний, радиус экранирования составляет 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 (7-29), из которой следует, что объемный заряд, введенный в полупроводник, в результате тока проводимости существует в среднем в течение времени τ_{μ} , т. е.

$$\rho = \rho_0 e^{-t/\tau_{\mu}}. \quad (8-40)$$

Если плотность объемного заряда поделить на величину единичного заряда, то получим изменение концентрации избыточных носителей заряда во времени:

$$\Delta n = \Delta n(0) e^{-t/\tau_{\mu}}. \quad (8-41)$$

Из сравнения формул (8-39) и (8-41) следует, что распространение носителей заряда в монополярном случае на расстояние длины экранирования l_{Σ} осуществляется в течение максвелловского времени релаксации τ_{μ} , которое в данном случае является эффективным временем установления диффузионно-дрейфового равновесия.

8-5. ДИФФУЗИЯ И ДРЕЙФ НЕОСНОВНЫХ ИЗБЫТОЧНЫХ НОСИТЕЛЕЙ ЗАРЯДА В ПРИМЕСНОМ ПОЛУПРОВОДНИКЕ

Рассмотрим диффузию и дрейф неосновных избыточных носителей заряда в примесном полупроводнике при наличии внешнего однородного электрического поля \mathcal{E} .

Допустим, что имеется полупроводник n -типа ($p_0 \ll n_0$), поперечные размеры которого значительно меньше его длины. Очень узкая область полупроводника ($-l \leq x \leq 0$) освещается светом, так что во всем объеме освещенной области образца происходит равномерная генерация электронов и дырок (рис. 8-4). Пусть концентрация основных носителей заряда n_0 велика по сравнению с избыточной концентрацией неосновных носителей заряда Δp , время жизни которых τ_p .

В освещенной области электронного полупроводника, где возбуждаются в равном количестве избыточные электроны и дырки $\Delta n = \Delta p$, концентрация неравновесных дырок $p = p_0 + \Delta p$ будет значительно больше, чем их концентрация p_0 в неосвещенной части образца. Благодаря наличию градиента концентрации дырок возникнет их диффузия. В неосвещенной части образца, куда продиффундируют дырки, появится объемный положительный заряд и электронейтральность образца нарушится. Этот объемный заряд создает статическое электрическое поле, которое вызовет перераспределение носителей заряда, так что в эту область из освещенной части образца в течение максвелловского времени релаксации будут подтянуты электроны и заряд избыточных дырок будет скомпенсирован. Полная компенсация объемного заряда будет

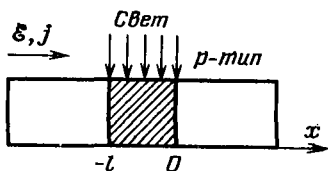


Рис. 8-4. Одномерная модель полупроводникового образца.

В том случае, если число избыточных электронов будет равно числу избыточных дырок в любой точке образца. Таким образом, *избыточные неосновные носители заряда—дырки, диффундируя в глубь электронного полупроводника, увлекают за собой равное количество основных носителей заряда — электронов, так что объемный заряд не создается, а диффузия неосновных носителей заряда—дырок происходит как диффузия незаряженных частиц. По мере продвижения в глубь полупроводника избыточные дырки и электроны будут рекомбинировать и их концентрации с расстоянием будут убывать.*

Если к такому полупроводнику приложить однородное электрическое поле \mathcal{E} , то избыточные носители заряда будут дрейфовать в этом поле, и их дрейф будет осуществляться со скоростью дрейфа неосновных носителей заряда — дырок таким образом, что будет сохраняться условие электронейтральности. При ускоряющем поле дырки будут затягивать за собой и электроны, а если поле будет замедлять движение дырок, то соответствующим образом произойдет и перераспределение основных носителей заряда. Следовательно, *и во внешнем электрическом поле избыточные неосновные носители заряда дрейфуют как незаряженные частицы, не создавая объемного заряда.*

Концентрация избыточных дырок Δp может быть определена в результате решения уравнения непрерывности

$$\frac{\partial p}{\partial t} = G - \frac{1}{e} \frac{djp}{dx} - \frac{p - p_0}{\tau_p} \quad (8-42)$$

Проведем решение этого уравнения для стационарного распределения избыточных носителей заряда ($\partial p / \partial t = 0$) в неосвещенной части полупроводника ($G = 0$). Для этого, продифференцировав выражение для плотности дырочного тока (8-22) и подставив его в уравнение (8-42), получим:

$$D_p \frac{d^2 \Delta p}{dx^2} - \mu_p \mathcal{E} \frac{d \Delta p}{dx} - \frac{\Delta p}{\tau_p} = 0. \quad (8-43)$$

Это равенство справедливо для такого внешнего электрического поля, когда D_p , μ_p и τ_p не зависят от поля.

Разделив обе части уравнения (8-43) на D_p и введя обозначения

$$L_p = \sqrt{D_p \tau_p} \quad (8-44)$$

и

$$L_{\mathcal{E}} = \tau_p \mu_p \mathcal{E}, \quad (8-45)$$

преобразуем уравнение (8-43) к виду

$$\frac{d^2 \Delta p}{dx^2} - \frac{L_{\mathcal{E}}}{L_p^2} \frac{d \Delta p}{dx} - \frac{\Delta p}{L_p^2} = 0. \quad (8-46)$$

Общим решением этого уравнения будет:

$$\Delta p = C_1 e^{\alpha_1 x} + C_2 e^{\alpha_2 x}, \quad (8-47)$$

где C_1 и C_2 — постоянные, определяемые из граничных условий; $\Delta p|_{x=0} = \Delta p|_{x=-l} = \Delta p(0)$, а при $x \rightarrow \pm \infty \Delta p \rightarrow 0$; α_1, α_2 — корни характеристического уравнения:

$$\alpha^2 - \frac{L_s}{L_p^2} \alpha - \frac{1}{L_p^2} = 0; \quad (8-48)$$

$$\alpha_{1,2} = \frac{L_s \pm \sqrt{L_s^2 + 4L_p^2}}{2L_p^2}. \quad (8-49)$$

Учитывая уменьшение концентрации неосновных неравновесных носителей заряда по мере удаления от освещенной области образца, окончательно получаем:

при $x \geq 0$

$$\Delta p = C_1 e^{-x/L_1}, \quad (8-50)$$

где

$$L_1 = \frac{2L_p^2}{\sqrt{L_s^2 + 4L_p^2} - L_s}, \quad (8-51)$$

а при $x \leq -l$

$$\Delta p = C_2 e^{x/L_2}, \quad (8-52)$$

где

$$L_2 = \frac{2L_p^2}{\sqrt{L_s^2 + 4L_p^2} + L_s}. \quad (8-53)$$

Таким образом, по обе стороны от освещенной области образца концентрация избыточных неосновных носителей заряда снижается по экспоненциальному закону с постоянными спада L_1 и L_2 , которые называются длиной затягивания.

Проведем анализ полученных выражений. Пусть внешнее электрическое поле отсутствует ($\mathcal{E} = 0$). В отсутствие внешнего электрического поля имеет место только диффузия избыточных дырок. Так как $L_s = \tau_p \mu_p \mathcal{E} = 0$, то избыточная концентрация неосновных носителей заряда — дырок в результате рекомбинации будет изменяться с расстоянием по закону

$$\Delta p = \Delta p(0) e^{-x \cdot L_p}. \quad (8-54)$$

Из этого выражения следует, что экспоненциальный спад концентрации неосновных носителей заряда, обусловленный рекомбинацией, будет симметричным по обе стороны от освещенной части

образца (рис. 8-5, а) и определяется величиной L_p , называемой диффузионной длиной неосновных носителей заряда — в данном случае дырок. L_p — это то среднее расстояние, на которое смещаются неравновесные дырки при диффузии за время жизни. Следовательно, в отсутствие электрического поля избыточная концентрация дырок в результате рекомбинации будет изменяться с расстоянием по экспоненциальному закону (8-54). Это распределение справедливо и для неравновесных электронов, так как при решении задачи предполагалось выполнение электронейтральности, по которому $\Delta n = \Delta p$.

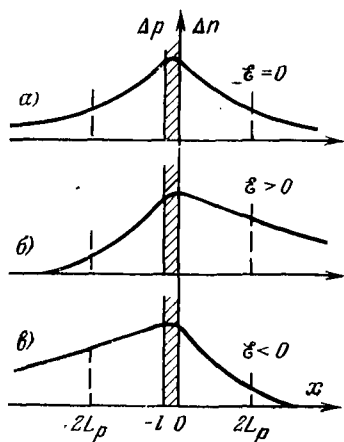


Рис. 8-5. Распределение концентрации избыточных носителей заряда вдоль образца.

называется диффузионной скоростью. Численно она равна скорости, с которой неравновесные носители заряда за время жизни проходят путь, равный диффузионной длине.

Рассмотрим теперь случай, когда $\mathcal{E} \neq 0$. Если $L_{\mathcal{E}} \ll 2L_p$, то $L_1 = L_2 = L_p$. Следовательно, в слабом электрическом поле распределение избыточной концентрации неосновных носителей заряда определяется диффузией и описывается уравнением (8-54).

При наличии сильного внешнего электрического поля, когда $|L_{\mathcal{E}}| > 2L_p$, постоянные спада L_1 и L_2 отличаются от диффузионной длины L_p . В зависимости от направления электрического поля ($\mathcal{E} > 0$ и $\mathcal{E} < 0$) они будут больше или меньше L_p . Например, при $\mathcal{E} > 0$ в неосвещенной части для областей $x < -l$ и $x > 0$

$$L_2 = \frac{2L_p^2}{\sqrt{L_{\mathcal{E}}^2 + 4L_p^2 + L_{\mathcal{E}}}} < L_p < \frac{2L_p^2}{\sqrt{L_{\mathcal{E}}^2 + 4L_p^2 - L_{\mathcal{E}}}} = L_1. \quad (8-57)$$

Отсюда следует, что внешнее электрическое поле искажает симметрию в распределении избыточной концентрации носителей заряда.

Рассмотрим теперь случай таких больших полей, для которых выполняется неравенство $L_\epsilon^2 \gg 4L_p^2$. Это условие согласно равенствам (8-45) и (8-56) можно записать следующим образом:

$$\frac{L_\epsilon^2}{4L_p^2} = \frac{(\tau_p \mu_p \mathcal{E})^2}{4L_p^2} = \frac{\tau_p^2 v^2}{4L_p^2} = \frac{v^2}{4v_D^2} \gg 1, \quad (8-58)$$

где $v = \mu_p \mathcal{E}$ — скорость дрейфа. Величина L_ϵ , численно равная пути, пройденному неравновесным носителем заряда за время жизни со скоростью дрейфа, называется длиной дрейфа. Следовательно, напряженность поля \mathcal{E} будет большой, если длина дрейфа намного превосходит диффузионную длину.

Определим длины затягивания L_1 и L_2 в случае, когда $\mathcal{E} > 0$. Для области $x > 0$ из выражения (8-51) можно записать:

$$L_1 = \frac{2L_p^2}{L_\epsilon} \frac{1}{\sqrt{1 + 4L_p^2/L_\epsilon^2 - 1}} \approx \frac{2L_p^2}{L_\epsilon} \frac{1}{2L_p^2/L_\epsilon^2} = L_\epsilon, \quad (8-59)$$

т. е. длина затягивания равна длине дрейфа

$$L_1 = L_\epsilon = \tau_p \mu_p \mathcal{E} = \tau_p v \quad (8-60)$$

и распределение избыточной концентрации дырок определяется выражением

$$\Delta p_1 = \Delta p(0) e^{-x/L_1} = \Delta p(0) e^{-x/\tau_p \mu_p \mathcal{E}}, \quad (8-61)$$

т. е. так же, как и в случае только диффузии, избыточная концентрация дырок спадает с ростом x экспоненциально, но только теперь с постоянной спада $L_1 = L_\epsilon$. Поскольку постоянная спада $L_1 > L_p$, то L_1 называется также «д и ф ф у з и о н н о й д л и н о й п о п о л ю».

Из сравнения соотношений (8-61) и (8-54) следует, что при $L_\epsilon^2 \gg 4L_p^2$ в области полупроводника при $x > 0$ избыточная концентрация носителей заряда больше (правая часть рис. 8-5, б), чем в отсутствие внешнего электрического поля. Следовательно, при прохождении тока, вызванного сильным электрическим полем, т. е. когда скорость дрейфа намного больше диффузионной скорости ($v \gg v_D$), при $\mathcal{E} > 0$ избыточные дырки в электронном полупроводнике затягиваются полем в область полупроводника $x > 0$ и полупроводник обогащен неосновными носителями заряда в большем количестве, чем при наличии только диффузии в отсутствие внешнего электрического поля ($\mathcal{E} = 0$). Это явление носит название и н ж е к ц и и неравновесных носителей заряда.

Для дырочного полупроводника инжекция электронов будет наблюдаться при $\mathcal{E} < 0$.

Рассмотрим теперь распределение избыточных дырок для области $x < -l$ в случае $\mathcal{E} > 0$. Теперь длина затягивания

$$L_2 = \frac{2L_p^2}{L_\varepsilon} \frac{1}{\sqrt{1 + 4L_p^2/L_\varepsilon^2 + 1}} \approx \frac{2L_p^2}{L_\varepsilon} \frac{1}{2 + 2L_p^2/L_\varepsilon^2} \approx \frac{L_p^2}{L_\varepsilon} \quad (8-62)$$

и на основании (8-52) получим:

$$\Delta p_2 = \Delta p(0) e^{-x/L_2} = \Delta p(0) e^{-\frac{L_\varepsilon}{L_p^2} x} \quad (8-63)$$

Очевидно, что с ростом напряженности поля L_ε увеличивается, а L_2 уменьшается. Поскольку $L_2 < L_p$, то L_2 называют «диффузионной длиной против поля» и для области $x < -l$ во всех точках полупроводника на расстоянии L_p $\Delta p_2 < \Delta p$. Это значит, что при $\mathcal{E} > 0$ с ростом напряженности электрического поля объем электронного полупроводника для области $x < -l$ обедняется неосновными носителями заряда (левая часть рис. 8-5, б). Это явление носит название эксклюзии носителей заряда.

Для дырочного полупроводника эксклюзия электронов будет иметь место при $\mathcal{E} < 0$.

При изменении направления внешнего электрического поля $\mathcal{E} < 0$ в области полупроводника $x > 0$ будет иметь место уменьшение концентрации избыточных носителей заряда, а в области $x < -l$ — их увеличение (рис. 8-5, в). Эти явления соответственно называются экстракцией и аккумуляцией неравновесных носителей заряда. В дырочном полупроводнике они будут наблюдаться при $\mathcal{E} > 0$.

8-6. ДИФФУЗИЯ И ДРЕЙФ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКЕ С ПРОВОДИМОСТЬЮ, БЛИЗКОЙ К СОБСТВЕННОЙ

Рассмотрим теперь диффузию и дрейф неравновесных носителей заряда в полупроводнике, в котором имеется сравнимое количество равновесных электронов и дырок, т. е. в полупроводнике с проводимостью, близкой к собственной. Будем считать, что избыточные носители заряда возникают в результате ионизации основного вещества поэтому $\Delta n = \Delta p$.

Пусть в таком однородном полупроводнике существует неоднородное вдоль оси x распределение неравновесных носителей заряда $n(x) = n_0 + \Delta n(x)$ и $p(x) = p_0 + \Delta p(x)$, вызванное, например, локальным освещением образца. Будем считать, что уровень возбуждения низкий, так что $\Delta n = \Delta p \ll (n_0 + p_0)$. Благодаря наличию внешнего электрического поля и градиента концентраций носителей заряда возникнут диффузионные и дрейфовые токи элек-

тронов и дырок, которые будут описываться уравнениями (8-21) и (8-22), и полный ток будет равен:

$$j = j_n + j_p = e(n\mu_n + p\mu_p)\mathcal{E} + e\left(D_n \frac{dn}{dx} - D_p \frac{dp}{dx}\right). \quad (8-64)$$

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 (8-64), если положить $j = 0$:

$$\mathcal{E}_{ст} = \frac{D_p \frac{dp}{dx} - D_n \frac{dn}{dx}}{n\mu_n + p\mu_p} = \frac{D_p \frac{dp}{dx} - D_n \frac{dn}{dx}}{(n_0 + \Delta n)\mu_n + (p_0 + \Delta p)\mu_p}. \quad (8-65)$$

Из выражения (8-65) видно, что при отсутствии электрического тока ($j = 0$) возникающее в результате диффузии носителей заряда статическое электрическое поле $\mathcal{E}_{ст}$ будет тем меньше, чем больше равновесные концентрации электронов n_0 и дырок p_0 . Если электропроводность полупроводника достаточно велика и концентрации избыточных носителей заряда Δn и Δp значительно меньше концентрации равновесных носителей заряда, т. е. $\Delta n \ll n_0$ и $\Delta p \ll p_0$, то в первом приближении можно считать, что статическое электрическое поле отсутствует и имеет место условие электронейтральности, которое обеспечивается тем, что в каждой точке полупроводника $\Delta n = \Delta p$. Это значит, что диффундирующие носители заряда увлекают с собой в процессе диффузии и носители заряда противоположного знака в равном количестве. В этом случае процесс диффузии избыточных электронов и дырок происходит свободно, как диффузия избыточных нейтральных пар электрон—дырка, характеризующихся одним временем жизни τ .

Пусть образец, в котором созданы неравновесные электронно-дырочные пары, 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 уравнение для плотности токов должны быть записаны как для электронов, так и для дырок:

$$\left. \begin{aligned} \frac{\partial \Delta n}{\partial t} &= \frac{1}{e} \frac{\partial j_n}{\partial x} - \frac{\Delta n}{\tau}, \\ \frac{\partial \Delta p}{\partial t} &= -\frac{1}{e} \frac{\partial j_p}{\partial x} - \frac{\Delta p}{\tau}; \end{aligned} \right\} \quad (8-66)$$

$$\left. \begin{aligned} j_n &= \sigma_n \mathcal{E} + eD_n \frac{d\Delta n}{dx}; \\ j_p &= \sigma_p \mathcal{E} - eD_p \frac{d\Delta p}{dx}, \end{aligned} \right\} \quad (8-67)$$

где $\sigma_n = en\mu_n$ и $\sigma_p = er\mu_p$ — электронная и дырочная составляющие удельной проводимости.

Под воздействием внешнего электрического поля пары электрон—дырка будут дрейфовать с постоянной скоростью. При этом совместная диффузия и дрейф электронов и дырок при условии 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 (8-66), подставив в них значения токов j_n и j_p из (8-67):

$$\left. \begin{aligned} \frac{\partial \Delta n}{\partial t} &= D_n \frac{\partial^2 \Delta n}{\partial x^2} + \mu_n \mathcal{E} \frac{\partial \Delta n}{\partial x} - \frac{\Delta n}{\tau}; \\ \frac{\partial \Delta p}{\partial t} &= D_p \frac{\partial^2 \Delta p}{\partial x^2} + \mu_p \mathcal{E} \frac{\partial \Delta p}{\partial x} - \frac{\Delta p}{\tau}. \end{aligned} \right\} \quad (8-68)$$

Умножим уравнения (8-68) соответственно на σ_p и σ_n и сложим оба уравнения. В результате, учитывая, что $\Delta n = \Delta p$, получаем:

$$\frac{\partial \Delta n}{\partial t} = \frac{D_n \sigma_p + D_p \sigma_n}{\sigma_n + \sigma_p} \frac{\partial^2 \Delta n}{\partial x^2} + \frac{\mu_n \sigma_p - \mu_p \sigma_n}{\sigma_n + \sigma_p} \mathcal{E} \frac{\partial \Delta n}{\partial x} - \frac{\Delta n}{\tau}. \quad (8-69)$$

Для стационарного случая, когда $\frac{\partial \Delta n}{\partial t} = 0$, уравнение (8-69) запишется в виде

$$\frac{D_n \sigma_p + D_p \sigma_n}{\sigma_n + \sigma_p} \frac{\partial^2 \Delta n}{\partial x^2} + \frac{\mu_n \sigma_p - \mu_p \sigma_n}{\sigma_n + \sigma_p} \mathcal{E} \frac{\partial \Delta n}{\partial x} - \frac{\Delta n}{\tau} = 0. \quad (8-70)$$

Уравнение (8-70) по форме совпадает с выражением (8-43) для диффузии и дрейфа неосновных носителей заряда. Разница между ними лишь в том, что вместо коэффициента диффузии при второй производной и подвижности при первой производной в уравнении (8-70) стоят сложные величины, которые соответственно могут быть обозначены через D и $\mu_{\mathcal{E}}$.

Учитывая, что $n \approx p_0$ и $p \approx p_0$, а это справедливо, когда $\Delta n \ll n_0$ и $\Delta p \ll p_0$, и используя соотношение Эйнштейна для

электронов и дырок ($\mu_n/D_n = \mu_p/D_p = e/kT$), коэффициент биполярной диффузии можно записать в виде

$$D = \frac{D_n \sigma_p + D_p \sigma_n}{\sigma_n + \sigma_p} = \frac{n_0 + p_0}{\frac{n_0}{D_p} + \frac{p_0}{D_n}} = \frac{kT}{e} \frac{n_0 + p_0}{\frac{n_0}{\mu_n} + \frac{p_0}{\mu_p}}, \quad (8-71)$$

а биполярную дрейфовую подвижность — следующим образом:

$$\mu_\varepsilon = \frac{\mu_n \sigma_p - \mu_p \sigma_n}{\sigma_n + \sigma_p} = \frac{p_0 - n_0}{\frac{n_0}{\mu_p} + \frac{p_0}{\mu_n}}. \quad (8-72)$$

Если воспользоваться соотношением Эйнштейна, то коэффициент биполярной диффузии D можно представить в виде

$$D = \frac{kT}{e} \mu_D. \quad (8-73)$$

Из сопоставления с равенством (8-73) следует, что в уравнении (8-71) роль подвижности играет величина, равная:

$$\mu_D = \frac{\frac{n_0 + p_0}{\frac{n_0}{\mu_p} + \frac{p_0}{\mu_n}}}{\frac{n_0}{\mu_p} + \frac{p_0}{\mu_n}} \quad (8-74)$$

и называемая биполярной диффузионной подвижностью.

Из уравнения (8-72) ясно, что величина μ_ε 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:

$$D = 2 \frac{D_n D_p}{D_n + D_p} = 2 \frac{kT}{e} \frac{\mu_n \mu_p}{\mu_n + \mu_p}; \quad (8-75)$$

$$\mu_D = 2 \frac{\mu_n \mu_p}{\mu_n + \mu_p}; \quad (8-76)$$

$$\mu_\varepsilon = 0. \quad (8-77)$$

Из соотношений (8-75) и (8-77) следует, что для собственного полупроводника диффузия избыточных носителей заряда определяется коэффициентом, который зависит от коэффициентов диффузии (или подвижностей) электронов и дырок. Равенство нулю биполярной дрейфовой подвижности свидетельствует о том, что внешнее электрическое поле в собственном полупроводнике не влияет на пространственное распределение носителей заряда.

Для примесных полупроводников, когда концентрация основных носителей заряда значительно превышает концентрацию неосновных носителей заряда ($n_0 \gg p_0$ или $p_0 \gg n_0$), коэффициент биполярной диффузии D , биполярная диффузионная подвижность

μ_D и биполярная дрейфовая подвижность μ_ϵ равны соответственно коэффициенту диффузии и подвижности неосновных носителей заряда. Например, для полупроводника n -типа ($n_0 \gg p_0$)

$$D = D_p,$$

а

$$\mu_D = |\mu_\epsilon| = \mu_p. \quad (8-78)$$

Следовательно, в примесном полупроводнике диффузия и дрейф избыточных носителей заряда определяются коэффициентом диффузии и подвижностью неосновных носителей заряда.

Глава девятая

КОНТАКТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

9-1. ПОЛУПРОВОДНИК ВО ВНЕШНЕМ ЭЛЕКТРИЧЕСКОМ ПОЛЕ

В основе контактных явлений лежат физические процессы, протекающие в полупроводнике под действием электрического поля, возникающего на контакте. Для понимания этих явлений необходимо знание свойств полупроводника, находящегося во внешнем электрическом поле. С этой целью мы проведем рассмотрение свойств электронного полупроводника, внесенного в однородное электрическое поле конденсатора (рис. 9-1).

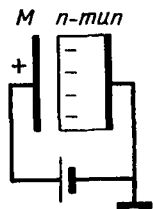


Рис. 9-1. Полупроводник в однородном электрическом поле.

В полупроводнике в отсутствие внешнего электрического поля объемный заряд равен нулю. При наличии внешнего электрического поля в полупроводнике произойдет перераспределение носителей заряда, в результате чего в нем появится объемный заряд, плотность которого ρ (г), и электрическое поле \mathcal{E} (г). Изменение распределения концентрации носителей заряда, обуславливающее появление объемного заряда, будет иметь место в области, прилегающей к поверхности полупроводника.

Объемный заряд будет экранировать внешнее электрическое поле, вследствие чего оно проникнет только в приповерхностный слой полупроводника.

При таком подключении внешнего источника питания, как это изображено на рис. 9-1, в приповерхностной области будет повышенная концентрация электронов (рис. 9-2, а) и, следовательно, возникнет отрицательный объемный заряд (рис. 9-2, б). Избыточная концентрация электронов, а значит, и объемный заряд будут уменьшаться с увеличением расстояния от поверхности в глубь полупроводника. Отрицательный объемный заряд породит электри-

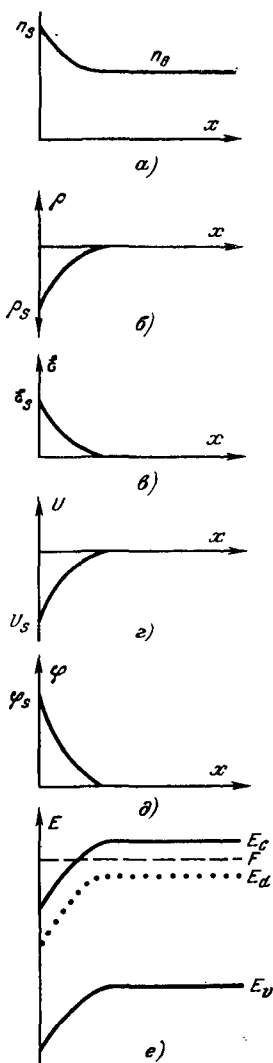


Рис. 9-2.

Рис. 9-2. Изменение концентрации свободных электронов (а), объемного заряда (б), напряженности поля (в), потенциальной энергии электронов (г), потенциала поля (д) и искривления зон донорного полупроводника (е) во внешнем электрическом поле.

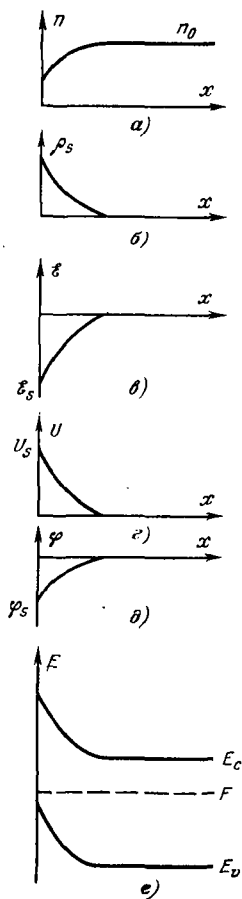


Рис. 9-3.

Рис. 9-3. Концентрация свободных электронов (а), плотность объемного заряда (б), напряженность электрического поля (в), потенциальная энергия электронов (г), потенциал поля (д) и положение зон донорного полупроводника (е) при изменении направления внешнего электрического поля.

ческое поле, напряженность \mathcal{E}_s , которого будет максимальна на поверхности полупроводника (рис. 9-2, в). Это электрическое поле изменит потенциальную энергию электрона (рис. 9-2, г) на величину, равную $U(\mathbf{r}) = -e\varphi(\mathbf{r})$, где $\varphi(\mathbf{r})$ — электростатический потенциал контактного поля (рис. 9-2, д). Следовательно, электрическое поле вызовет искривление зон энергии полупроводника таким образом, что

$$\left. \begin{aligned} E_c(\mathbf{r}) &= E_c - U(\mathbf{r}); \\ E_v(\mathbf{r}) &= E_v - U(\mathbf{r}). \end{aligned} \right\} \quad (9-1)$$

При этом 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. § 5-2), поэтому расстояние между уровнем Ферми и разрешенными зонами энергии изменится. Если это расстояние без поля было:

$$E_c - F \quad \text{и} \quad F - E_v, \quad (9-2)$$

то при наличии поля оно будет:

$$E_c - U(\mathbf{r}) - F \quad \text{и} \quad F - [E_v - U(\mathbf{r})]. \quad (9-3)$$

Из сравнения (9-3) и (9-2) следует, что, если расстояние между E_c и F уменьшается на величину $U(\mathbf{r})$, то между F и E_v увеличивается на ту же величину. Изменение расстояния между F и зонами энергии должно привести к изменению распределения электронов по уровням. В рассматриваемом случае, как это изображено на рис. 9-2, е, вдали от поверхности полупроводник обладает электронной электропроводностью, невырожден, и на уровнях донорной примеси имеются электроны (поскольку уровень Ферми лежит выше уровня примеси), а в приповерхностной области полупроводник, оставаясь электронным, становится вырожденным. Здесь уровень Ферми расположен в зоне проводимости.

При изменении направления приложенного поля в приповерхностной области электронного полупроводника концентрация электронов будет меньше, а дырок больше, чем в объеме полупроводника (рис. 9-3, а), а объемный заряд (рис. 9-3, б) и изменение потенциальной энергии (рис. 9-3, г) по знаку будут положительные. В данном случае, как это изображено на рис. 9-3, е, в приповерхностной области произошло даже изменение типа основных носителей заряда — полупроводник из электронного превратился в дырочный, т. е. возник и н в е р с н ы й, или о б р а щ е н н ы й, с л о й. У такого полупроводника на некотором расстоянии от поверхности будет слой с собственной электропроводностью или i -слой с уровнем Ферми, лежащим в середине запрещенной зоны. Эта область полупроводника, расположенная около i -слоя, в которой меняется тип электропроводности, носит название физического p - n перехода. Он исчезает при снятии внешнего электрического поля.

Рассмотрение влияния внешнего электрического поля проведем для одномерного невырожденного полупроводника n -типа, энергетическая схема которого представлена на рис. 9-3, *е*. Как известно, напряженность электрического поля \mathcal{E} связана с объемным зарядом ρ уравнением Пуассона

$$\frac{d\mathcal{E}}{dx} = \frac{1}{\varepsilon\varepsilon_0} \rho(x). \quad (9-4)$$

Если напряженность поля выразить через градиент потенциала $\mathcal{E} = -\frac{d\varphi}{dx}$, то уравнение Пуассона будет иметь вид:

$$\frac{d^2\varphi}{dx^2} = -\frac{1}{\varepsilon\varepsilon_0} \rho(x). \quad (9-5)$$

Пусть концентрация электронов в объеме полупроводника ($x \rightarrow \infty$) равна n_0 , а в его приповерхностной области n . Поскольку полупроводник невырожден, то

$$n = N_c e^{-\frac{E_c + U - F}{kT}} = n_0 e^{-\frac{U}{kT}}. \quad (9-6)$$

В приповерхностном слое объемный заряд определяется положительными ионами донорной примеси и свободными электронами. Будем считать, что донорная примесь полностью ионизирована, т. е. $N_d^+ = n_0$. Для такого полупроводника объемный заряд в приповерхностном слое будет равен:

$$\bar{\rho} = e(N_d^+ - n) = e(n_0 - n) = en_0 \left(1 - e^{-\frac{U}{kT}}\right). \quad (9-7)$$

Ограничимся случаем малого искривления зон под действием поля, т. е. когда $|U| \ll kT$. Разлагая выражение $e^{-U/kT}$ в ряд и ограничиваясь первым членом разложения, на основании соотношения (9-7) будем иметь:

$$\bar{\rho} = \frac{en_0 U}{kT} = -\frac{e^2 n_0}{kT} \varphi. \quad (9-8)$$

Обозначим $l_s^2 = \frac{\varepsilon\varepsilon_0 kT}{e^2 n_0}$, тогда уравнение (9-5) примет вид:

$$\frac{d^2\varphi}{dx^2} - \frac{1}{l_s^2} \varphi = 0. \quad (9-9)$$

Это уравнение имеет решение

$$\varphi = A_1 e^{-x/l_s} + A_2 e^{x/l_s}. \quad (9-10)$$

Так как при $x \rightarrow \infty$ $\varphi \rightarrow 0$, значит $A_2 = 0$, а в точке $x = 0$ $\varphi = -\varphi_s$ и $A_1 = -\varphi_s$, поэтому в приповерхностной области потенциал

$$\varphi(x) = -\varphi_s e^{-x/l_s}, \quad (9-11)$$

напряженность поля

$$\mathcal{E}(x) = -\frac{d\varphi}{dx} = -\frac{\varphi_s}{l_s} e^{-x/l_s} = \mathcal{E}_s e^{-x/l_s}, \quad (9-12)$$

потенциальная энергия электронов

$$U(x) = -e\varphi(x) = e\varphi_s e^{-x/l_s} = U_s e^{-x/l_s}, \quad (9-13)$$

а плотность объемного заряда на поверхности будет равна:

$$\rho_s = \frac{en_0}{kT} U_s. \quad (9-14)$$

Таким образом, при внесении полупроводника в электрическое поле в приповерхностной области его происходят искривление зон энергии и изменение концентрации электронов и дырок. При $U_s > 0$ зоны смещаются вверх и увеличивается концентрация дырок (рис. 9-3, е). В этом случае у полупроводника *n*-типа приповерхностный слой обогащается неосновными носителями заряда, а у *p*-типа — основными носителями заряда. При $U_s < 0$ зоны смещаются вниз (рис. 9-2, е) и увеличивается концентрация электронов, т. е. у полупроводника *n*-типа имеет место обогащение основными носителями заряда, а у *p*-типа — обогащение неосновными носителями заряда.

В формулах (9-10) — (9-13) l_s есть дебаевская длина экранирования. Следовательно, l_s — это то расстояние, на котором напряженность электрического поля в веществе уменьшается в $e \approx 2,7$ раза. Для металла, у которого $n_0 = 10^{22} \text{ см}^{-3}$ и $\epsilon = 1$, при комнатной температуре дебаевская длина экранирования составляет примерно 10^{-8} см. В чистый германий ($n_0 = 10^{14} \text{ см}^{-3}$ и $\epsilon = 16$) электрическое поле проникает на расстояние 4 мкм.

9-2. РАБОТА ВЫХОДА

Как известно, чтобы электрон оторвать от тела, необходимо затратить некоторую энергию. Следовательно, на границе твердого тела существует энергетический барьер, препятствующий выходу электронов из кристалла. Его могут покинуть лишь электроны, имеющие энергию, достаточную для преодоления этого барьера. Очевидно, что чем выше температура, тем больше будет электронов, способных преодолеть потенциальный барьер. *Явление выхода из вещества электронов за счет тепловой энергии называется термоэлектронной эмиссией.*

Найдем плотность тока термоэлектронной эмиссии электронов из невырожденного полупроводника. Для этого необходимо подсчитать число электронов, способных преодолеть потенциальный барьер и выйти из полупроводника. Обозначим через $E_a = E_{\text{вак}}$ энергию электрона, вышедшего из полупроводника в вакуум и по-

коящегося относительно образца. Тогда для перевода электрона со дна зоны проводимости, где он имеет нулевую скорость, в вакуум без сообщения ему скорости потребуется энергия χ (рис. 9-4), равная:

$$\chi = E_a - E_c. \quad (9-15)$$

Энергия χ , отделяющая край зоны проводимости от уровня вакуума ($E_{\text{вак}} = E_a$), есть энергия электронного средства. Иногда ее называют истинной, или внешней, работой выхода. Численно она равна работе, необходимой для перевода покоящегося в твердом теле электрона в вакуум без сообщения ему кинетической энергии.

Для преодоления прямоугольного потенциального барьера высотой χ кинетическая энергия электрона, имеющего скорость v_x , должна быть не меньше глубины потенциального барьера, т. е.

$$\frac{m^* v_x^2}{2} \geq \chi. \quad (9-16)$$

Число квантовых состояний для кристалла единичного объема в интервале скоростей от v до $v + dv$ определяется выражением (см. § 4-1)

$$dZ = 2 \left(\frac{m^*}{h} \right)^3 dv_x dv_y dv_z. \quad (9-17)$$

Поскольку для электронов, способных покинуть полупроводник, $E - F \gg kT$, то их число будет:

$$dn = \int_0^\infty dZ = 2 \left(\frac{m^*}{h} \right)^3 e^{-(E-F)/kT} dv_x dv_y dv_z. \quad (9-18)$$

Будем считать, что все электроны, преодолевшие потенциальный барьер, не возвращаются в полупроводник. Если к поверхности кристалла движется поток электронов со скоростью v_x , равной dnv_x , то плотность тока, текущего из вакуума в полупроводник, будет равна:

$$j = e \int_{v_{x \text{ мин}}}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty dnv_x = 2e \left(\frac{m^*}{h} \right)^3 e^{F/kT} \times \\ \times \int_{v_{x \text{ мин}}}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty e^{-E/kT} v_x dv_x dv_y dv_z. \quad (9-19)$$

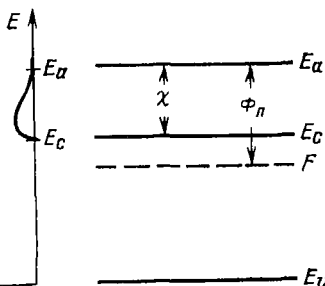


Рис. 9-4. Работа выхода электронов из полупроводника.

Учитывая, что полная энергия электрона $E = E_c + \frac{m^*}{2}(v_x^2 + v_y^2 + v_z^2)$, уравнение (9-19) можно записать в виде

$$j = \frac{2em^*}{h^3} e^{-(E_c - F)/kT} \int_{v_x \text{ мин}}^{\infty} e^{-m^* v_x^2/2kT} v_x dv_x \int_{-\infty}^{\infty} e^{-m^* v_y^2/2kT} dv_y \times \\ \times \int_{-\infty}^{\infty} e^{-m^* v_z^2/2kT} dv_z. \quad (9-20)$$

Для вычисления j используем соотношение

$$\int_{-\infty}^{\infty} e^{-\alpha y^2} dy = \sqrt{\frac{\pi}{\alpha}}, \quad (9-21)$$

а интегрирование по v_x с учетом выражения (9-16) дает:

$$\int_{v_x \text{ мин}}^{\infty} e^{-m^* v_x^2/2kT} v_x dv_x = \frac{kT}{m^*} e^{-m^* v_{x \text{ мин}}^2/2kT} = \frac{kT}{m^*} e^{-\chi/kT}. \quad (9-22)$$

На основании равенств (9-20) — (9-22) получим выражение для плотности тока термоэлектронной эмиссии:

$$j = \frac{4\pi em^* k^2}{h^3} T^2 e^{-(E_a - F)/kT} = A e^{-\Phi/kT}, \quad (9-23)$$

где

$$A = \frac{4\pi em^* k^2}{h^3};$$

$$\Phi = \chi + E_c - F = E_a - F. \quad (9-24)$$

Величина Φ называется термодинамической работой выхода электрона из полупроводника. Формула (9-23) справедлива и для металла. Разница в том, что у металла электроны находятся на уровне Ферми, а в невырожденном полупроводнике на уровне Ферми нет электронов. Как у металлов, так и у полупроводников Φ составляет обычно несколько электрон-вольт.

Поскольку в полупроводнике положение уровня Ферми зависит от температуры, природы и концентрации примеси, то и значение термодинамической работы выхода также будет определяться этими параметрами. Выразим работу выхода для различных полупроводников, используя формулу (9-24) и соответствующие соотношения для уровня Ферми, которые были получены в § 4-6 и 4-7.

Для собственного полупроводника работа выхода Φ_i равна:

$$\Phi_i = \chi + \frac{1}{2} E_g + \frac{kT}{2} \ln \left(\frac{m_n^*}{m_p^*} \right)^{3/2}. \quad (9-25)$$

Из приведенного выражения видно, что работа выхода электрона из собственного полупроводника зависит от ширины запрещенной зоны, температуры и соотношения эффективных масс электронов и дырок.

Для донорного полупроводника при слабой ионизации примеси

$$\Phi_n = \chi + \frac{E_c - E_d}{2} + \frac{kT}{2} \ln \frac{gN_c}{N_d}, \quad (9-26)$$

а при сильной ионизации примеси

$$\Phi_n = \chi + kT \ln \frac{N_c}{N_d}. \quad (9-27)$$

Если при сильной ионизации донорной примеси работа выхода определяется концентрацией примеси и температурой полупроводника, то при слабой ионизации она, кроме того, зависит еще и от глубины залегания уровня донорной примеси в запрещенной зоне и степени его вырождения.

Для акцепторного полупроводника при слабой и сильной ионизации примеси

$$\Phi_p = \chi + E_g + \frac{E_v - E_a}{2} - \frac{kT}{2} \ln \frac{gN_v}{N_a}; \quad (9-28)$$

$$\Phi_p = \chi + E_g - kT \ln \frac{N_v}{N_a}. \quad (9-29)$$

Из сравнения формул (9-26) — (9-29) следует, что работа выхода электронов из дырочного полупроводника больше, чем из электронного.

9.3. КОНТАКТ МЕТАЛЛ—МЕТАЛЛ. КОНТАКТНАЯ РАЗНОСТЬ ПОТЕНЦИАЛОВ

Пусть мы имеем два различных изолированных металла M_1 и M_2 . Предположим, что у металла M_1 работа выхода Φ_1 и энергия Ферми F_1 больше, чем Φ_2 и F_2 второго металла M_2 .

Приведем эти металлы в контакт так, что одними концами они будут соединены вместе, а два других будут разделены вакуумным промежутком (рис. 9-5).

Рассмотрим вначале процессы, происходящие в контакте C . В момент соприкосновения металлы начнут обмениваться электронами. Поскольку $F_1 > F_2$, то преимущественным будет переход электронов из металла M_1 в металл M_2 , поэтому на металле M_1 появится положительный заряд, а на металле M_2 — отрицательный. Следовательно, между металлами перпендикулярно плоскости контакта C возникнет электрическое поле, которое будет локализовано в слое толщиной порядка 10^{-8} см. Направленный поток электронов из металла M_1 в металл M_2 прекратится в тот момент, когда уровень Ферми для

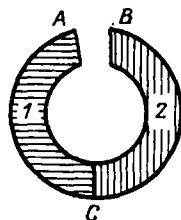


Рис. 9-5. Контакт двух разнородных металлов.

всей системы будет одинаков. Разность потенциалов, установившаяся при равновесном состоянии на жестком контакте между двумя металлами, равная φ_i , называется внутренней контактной разностью потенциалов.

Для определения величины φ_i необходимо подсчитать число электронов, перешедших вдоль оси x из металла M_1 в металл M_2 и наоборот. В соответствии с выражением (9-18) из металла M_1 перейдут электроны, имеющие скорость u_{x1} , в количестве

$$dN_1 = 2 \frac{1}{m_1^* h^3} p_{x1} dp_{x1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dp_{y1} dp_{z1}}{e^{(E_1 - F_1)/kT} + 1}, \quad (9-30)$$

а количество электронов, перешедших из второго в первый металл, составит величину

$$dN_2 = 2 \frac{1}{m_2^* h^3} p_{x2} dp_{x2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dp_{y2} dp_{z2}}{e^{(E_2 - F_2)/kT} + 1}. \quad (9-31)$$

Так как металл M_1 на границе раздела S заряжен положительно, то его электроны приобретут дополнительную отрицательную энергию, равную по величине $e\varphi_i$ (рис. 9-6, a). Но при термодинамическом равновесии полная энергия электронов по всей системе одинакова, поэтому

$$E_1 - e\varphi_i = E_2. \quad (9-32)$$

Поскольку

$$\frac{\partial E_1}{\partial p_{x1}} dp_{x1} = \frac{\partial E_2}{\partial p_{x2}} dp_{x2}, \quad (9-33)$$

то

$$\frac{1}{m_1^*} p_{x1} dp_{x1} = \frac{1}{m_2^*} p_{x2} dp_{x2}. \quad (9-34)$$

В условиях термодинамического равновесия число прямых переходов равно числу обратных переходов, т. е.

$$dN_1 = dN_2. \quad (9-35)$$

Это осуществимо лишь при условии равенства двойных интегралов в (9-30) и (9-31). Их равенство возможно в том случае, если

$$E_1 - F_1 = E_2 - F_2 = E_1 - e\varphi_i - F_2. \quad (9-36)$$

Отсюда следует, что

$$e\varphi_i = F_1 - F_2. \quad (9-37)$$

Таким образом, внутренняя контактная разность потенциалов определяется разностью энергий Ферми изолированных металлов.

Для металлов уровень Ферми связан с концентрацией электронов выражением

$$F_0 = \frac{h^2}{2m^*} \left(\frac{3n_0}{8\pi} \right)^{2/3}, \quad (9-38)$$

поэтому

$$e\varphi_i = \frac{h^2}{2} \left(\frac{3}{8\pi} \right)^{2/3} \left(\frac{n_1^{2/3}}{m_1^*} - \frac{n_2^{2/3}}{m_2^*} \right). \quad (9-39)$$

Следовательно, внутренняя контактная разность потенциалов определяется концентрацией электронов в изолированных металлах и их эффективными массами. Для металлов φ_i невелико и составляет 10^{-3} — 10^{-2} В.

Рассмотрим теперь процессы, происходящие между металлами в вакуумном зазоре (на участке AB рис. 9-5). Поскольку работа выхода электронов из металла M_1 больше, чем из металла M_2 , то в первый момент ток термоэлектронной эмиссии из металла M_2 будет больше, чем из металла M_1 , т. е. $j_2 > j_1$. Это приведет к тому, что на металле M_2 (в плоскости B) появится положительный заряд, а на металле M_1 (в плоскости A) — отрицательный, а следовательно, возникнет поле, препятствующее эмиссии электронов из металла M_2 . Направленный поток электронов будет существовать до тех пор, пока в системе не наступит динамическое равновесие, характеризующееся постоянством уровня Ферми (рис. 9-6, б).

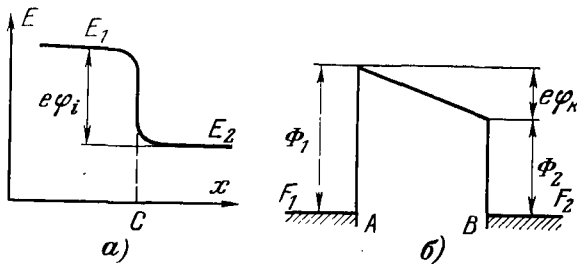


Рис. 9-6. Внутренняя (а) и внешняя (б) контактная разность потенциалов.

Контактная разность потенциалов φ_k , возникшая между металлами в вакуумном зазоре, носит название внешней контактной разности потенциалов.

Поскольку при термодинамическом равновесии имеет место равенство токов эмиссии

$$j_{10} = j_{20} \quad (9-40)$$

или

$$AT^2 e^{-\Phi_1/kT} = AT^2 e^{-(\Phi_2 + e\varphi_k)/kT}, \quad (9-41)$$

то

$$e\varphi_k = \Phi_1 - \Phi_2. \quad (9-42)$$

Таким образом, внешняя контактная разность потенциалов, возникающая между металлами в вакуумном зазоре, определяется разностью работ выхода электронов из этих металлов.

9-4. КОНТАКТ МЕТАЛЛ—ПОЛУПРОВОДНИК

Допустим, что между металлом и электронным полупроводником создан надежный контакт. Если уровень Ферми изолированного металла F_M лежит ниже уровня Ферми полупроводника F_{II} , т. е. $\Phi_M > \Phi_{II}$, то в первый момент их соприкосновения поток электронов из полупроводника превышает поток электронов из металла. Металл заряжается отрицательно, а полупроводник положительно, и возникшее между контактирующими образцами электрическое поле $\mathcal{E}_{ст}$ будет препятствовать переходу электронов из полупроводника в металл. Направленный поток электронов будет иметь место

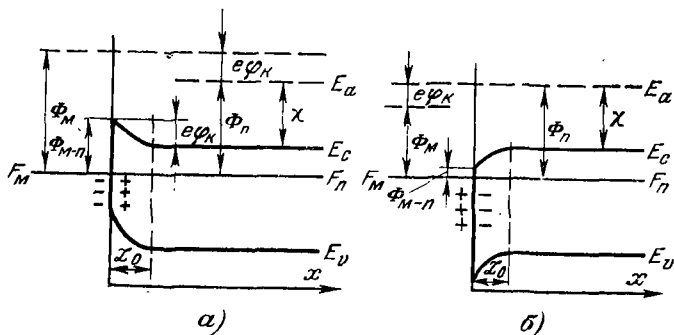


Рис. 9-7. Контакт металл—электронный полупроводник в случае $\Phi_M > \Phi_{II}$ (а) и $\Phi_M < \Phi_{II}$ (б).

до тех пор, пока уровни Ферми в системе не выравняются (рис. 9-7, а) и установится динамическое равновесие, характеризующееся равенством токов термоэлектронной эмиссии

$$i_{M0} = i_{II0}. \quad (9-43)$$

Отсюда найдем, что контактная разность потенциалов, возникающая на контакте металла с полупроводником

$$e\varphi_k = \Phi_M - \Phi_{II}. \quad (9-44)$$

Так как напряженность электрического поля в слое объемного заряда $\mathcal{E}_{ст} = \int \frac{1}{\epsilon\epsilon_0} \rho dx = \frac{1}{\epsilon\epsilon_0} Q$, то при равенстве объемных зарядов ($Q = \text{const}$) разность потенциалов в соприкасающихся телах пропорциональна толщине слоя объемного заряда ($V = - \int_0^l \mathcal{E}_{ст} dx = - \frac{Q}{\epsilon\epsilon_0} l$). Но толщина слоя объемного заряда в металле не превышает 10^{-8} — 10^{-7} см, поэтому разность потенциалов на слое объемного заряда в металле ничтожно мала по сравнению с разностью потенциалов в полупроводнике. Поэтому можно считать, что контактная разность потенциалов φ_k практически полностью падает в приконтактной области полупроводника. На-

пряженность электрического поля в приповерхностном слое, вызванного контактной разностью потенциалов, не превышает 10^6 В/см, а напряженность поля, обусловленного ионами кристаллической решетки, составляет 10^8 В/см. Поэтому небольшая добавка за счет контактного поля не может изменить ширину запрещенной зоны полупроводника, но искривит его зоны энергии относительно уровня Ферми, а это вызовет изменение концентрации электронов и дырок в его приповерхностном слое.

При этом, если работа выхода электронов из полупроводника меньше работы выхода из металла, полупроводник заряжается положительно и зоны энергии в приконтактной области искрив-

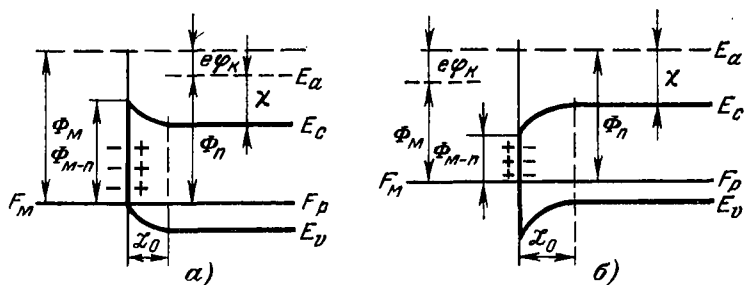


Рис. 9-8. Контакт металл—дырочный полупроводник в случае $\Phi_M > \Phi_P$ (а) и $\Phi_M < \Phi_P$ (б).

ляются кверху. Поэтому вблизи контакта число электронов в зоне проводимости убывает, а число дырок в валентной зоне возрастает по сравнению с их числом в объеме полупроводника. Таким образом, в случае, когда $\Phi_M > \Phi_P$, в электронном полупроводнике концентрация дырок у контакта будет больше, чем в объеме полупроводника (рис. 9-7, а), т. е. возникает слой с пониженной удельной проводимостью. Слой с пониженной удельной проводимостью (обогащенный неосновными носителями заряда) называется запирающим. У дырочного полупроводника концентрация дырок у контакта будет больше, чем в объеме образца (рис. 9-8, а), т. е. возникает слой с повышенной удельной проводимостью. Слой с повышенной удельной проводимостью (обогащенный основными носителями заряда) называется антизапирающим.

Если работа выхода из полупроводника больше работы выхода из металла, т. е. $\Phi_P > \Phi_M$, полупроводник заряжается отрицательно и зоны энергии в приконтактной области искривляются книзу, поэтому вблизи контакта число электронов в зоне проводимости увеличивается, а число дырок в валентной зоне убывает. Это значит, что в электронном полупроводнике возникнет антизапирающий слой (рис. 9-7, б), а в дырочном — запирающий слой (рис. 9-8, б).

При сильном обогащении приконтактной области неосновными носителями заряда, приводящего к инверсии электропроводности,

возникнет физический p - n переход. Сильное обогащение приконтактной области основными носителями заряда может привести к вырожденности полупроводника.

У собственного полупроводника как при $\Phi_M > \Phi_n$, так и при $\Phi_M < \Phi_n$ искривление зон энергии сопровождается повышением удельной проводимости в приконтактной области (рис. 9-9).

Преобразуем выражение (9-44) следующим образом:

$$e\varphi_K = (\Phi_M - \chi) - (\Phi_n - \chi) = \Phi_{M-n} - \Phi_{n-m}. \quad (9-45)$$

Здесь $\Phi_{M-n} = \Phi_M - \chi$ есть работа, которую необходимо совершить для перевода электрона с уровня Ферми металла на дно зоны

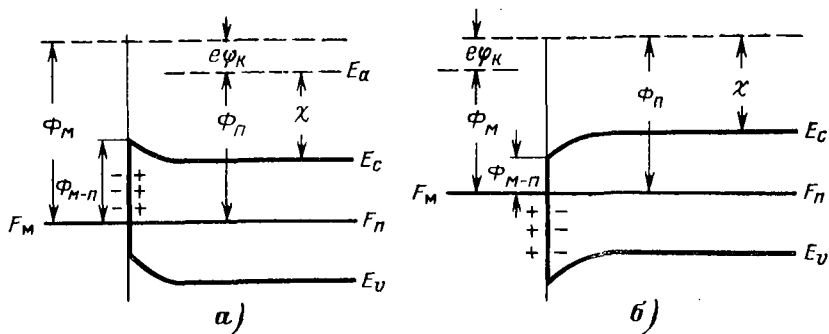


Рис. 9-9. Контакт металл — собственный полупроводник в случае $\Phi_M > \Phi_n$ (а) и $\Phi_M < \Phi_n$ (б).

проводимости полупроводника в контакте (рис. 9-7 — 9-9); $\Phi_{n-m} = \Phi_n - \chi$ — энергия, которую необходимо затратить для перевода электрона, находящегося на дне зоны проводимости в объеме полупроводника, непосредственно в металл.

Предположим, что электрическое поле полностью проникает в электронный полупроводник на глубину L_0 (рис. 9-7, а). Ее определение проведем для полупроводника, у которого имеет место полная ионизация донорной примеси, т. е. $n_0 = N_d$. Поскольку в приконтактном слое энергия электронов на дне зоны проводимости равна $E_C = e\varphi(x)$, объемный заряд в этом слое согласно (9-7) определяется соотношением вида

$$\rho = en_0 [1 - e^{-e\varphi(x)/kT}]. \quad (9-46)$$

Так как контактная разность потенциалов полностью падает в приконтактной области полупроводника, можно считать, что $|e\varphi(x)| \gg kT$. Для этого случая объемный заряд

$$\rho = en_0. \quad (9-47)$$

Это означает, что из области полупроводника, протяженностью L_0 , свободные электроны вытеснены электрическим полем и в приконтактной области сосредоточен положительный заряд, обусловленный ионами донорной примеси.

Для области объемного заряда уравнение Пуассона с учетом равенства (9-47) запишется в виде

$$\frac{d^2\varphi}{dx^2} + \frac{e}{\varepsilon\varepsilon_0} n_0 = 0. \quad (9-48)$$

Общим решением этого уравнения является функция

$$\varphi(x) = -\frac{en_0}{2\varepsilon\varepsilon_0} (\mathcal{L}_0 - x)^2 + A(\mathcal{L}_0 - x) + B. \quad (9-49)$$

Поле проникает в полупроводник только на глубину \mathcal{L}_0 , а поэтому уравнение (9-49) должно удовлетворять граничным условиям

$$\begin{aligned} \varphi(\mathcal{L}_0) &= 0; \\ \mathcal{E}(\mathcal{L}_0) &= -\left. \frac{d\varphi}{dx} \right|_{x=\mathcal{L}_0} = 0. \end{aligned} \quad (9-50)$$

Подставляя граничные условия в уравнение (8-49), получаем:

$$\left. \begin{aligned} \varphi(\mathcal{L}_0) = B &= 0; \\ \left. \frac{d\varphi}{dx} \right|_{x=\mathcal{L}_0} = -A &= 0. \end{aligned} \right\} \quad (9-51)$$

Следовательно, в приконтактной области электронного полупроводника (рис. 9-7, а) электростатический потенциал в зависимости от координаты x меняется как

$$\varphi(x) = -\frac{en_0}{2\varepsilon\varepsilon_0} (\mathcal{L}_0 - x)^2. \quad (9-52)$$

Для определения величины \mathcal{L}_0 используем граничное условие в точке $x = 0$. В этом случае

$$\varphi(0) = -\varphi_K = -\frac{1}{e} (\Phi_M - \Phi_n). \quad (9-53)$$

Это условие позволяет на основании уравнения (9-52) определить глубину проникновения поля в полупроводник:

$$\mathcal{L}_0 = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_K}{en_0}} = \sqrt{\frac{2\varepsilon\varepsilon_0(\Phi_M - \Phi_n)}{e^2n_0}}. \quad (9-54)$$

Из этой формулы следует, что чем меньше степень легирования полупроводника и чем больше разность работ выхода электронов из металла и полупроводника, тем больше глубина проникновения в полупроводник электрического поля, вызванного контактной разностью потенциалов. При разности работ выхода около 1 эВ глубина проникновения поля в полупроводник примерно в 10 раз превышает длину экранирования в случае контакта металл — вакуум — полупроводник.

Согласно рис. 9-7, а и 9-8, б приконтактный слой полупроводника, когда он является запирающим, обеднен основными носи-

телями заряда, а поэтому обладает повышенным сопротивлением по сравнению с толщиной полупроводника. По существу система металл — запирающий слой повышенного сопротивления, заключенный между металлом и сравнительно хорошо проводящим полупроводником, является плоским конденсатором. Таким образом, контакт металла с полупроводником обладает емкостью на единицу площади, равной:

$$C = \frac{\varepsilon \varepsilon_0}{\mathcal{L}} = \sqrt{\frac{\varepsilon \varepsilon_0 e n_0}{2\varphi_K}}. \quad (9-55)$$

9-5. ВЫПРЯМЛЕНИЕ НА КОНТАКТЕ МЕТАЛЛ—ПОЛУПРОВОДНИК

Вернемся к рассмотренному контакту металла с донорным полупроводником, когда выполняется условие $\Phi_m > \Phi_n$. В равновесном состоянии в приконтактной области имеет место изгиб зон и возникает запирающий слой, обедненный электронами (рис. 9-7, а). В глубине полупроводника ($x > \mathcal{L}_0$) концентрация свободных носителей

$$n_0 = N_c e^{-(E_c - F)/kT} = N_c e^{\xi_0/kT}, \quad (9-56)$$

где E_c — положение дна зоны проводимости в толще полупроводника; $\xi_0 = F - E_c$ — положение уровня Ферми в глубине полупроводника, отсчитываемого от нижнего края зоны проводимости. При термодинамическом равновесии положение уровня Ферми для всей структуры одинаково:

$$F = E_c + \xi_0 = E_c + e\varphi(x) + \xi(x) = \text{const}. \quad (9-57)$$

Здесь $e\varphi(x)$ — потенциальная энергия электронов в приконтактной области; $\xi(x)$ — положение уровня Ферми в приконтактной области, отсчитываемого от дна зоны проводимости.

При подключении источника напряжения через контакт металл—полупроводник будет протекать ток. Так как состояние полупроводника становится неравновесным, то концентрация электронов проводимости будет определяться квазиуровнем Ферми (см. § 7-1), равным:

$$F_n = E_c + e\varphi(x) + \xi_1(x), \quad (9-58)$$

где $\xi_1(x)$ — положение квазиуровня Ферми в приконтактной области, отсчитываемого от дна зоны проводимости. Поэтому концентрацию свободных электронов в приконтактной области полупроводника, т. е. в области $0 < x < \mathcal{L}$, по аналогии с (9-56) можно записать в виде

$$n(x) = N_c e^{\xi_1(x)/kT}. \quad (9-59)$$

Поскольку в приконтактной области концентрация электронов зависит от x , то квазиуровень Ферми непостоянен во всей области

полупроводника. Чтобы определить его изменение, вычислим ток, протекающий через полупроводник. Плотность тока

$$j = en\mu_n \mathcal{E} + eD_n \frac{dn}{dx} \quad (9-60)$$

можно выразить через потенциал и концентрацию электронов, используя для этого формулу (9-59). В результате получим:

$$j = n\mu_n \frac{d(e\varphi)}{dx} + \frac{eD_n}{kT} \frac{d\xi_1}{dx}. \quad (9-61)$$

Это выражение с учетом соотношения Эйнштейна преобразуется к виду

$$j = n\mu_n \frac{d(e\varphi + \xi_1)}{dx}. \quad (9-62)$$

Сравнивая равенства (9-62) и (9-58), можно написать:

$$j = n\mu_n \frac{dF_n}{dx}. \quad (9-63)$$

Найдем, как изменится квазиуровень Ферми между двумя точками 1 и 2 полупроводника, в которых плотность тока одинакова:

$$\int_{(1)}^{(2)} dF_n = \int_{(1)}^{(2)} \frac{j dx}{n\mu_n} = ejs \int_{(1)}^{(2)} \frac{dx}{en\mu_n s} = eI \int_{(1)}^{(2)} dR. \quad (9-64)$$

Здесь $dR = \frac{dx}{\sigma s}$ — сопротивление участка полупроводника длиной dx с поперечным сечением s и удельной проводимостью $\sigma = en\mu_n$. Интегрируя уравнение (9-64), будем иметь:

$$\Delta F_n = F_{n2} - F_{n1} = eI (R_2 - R_1) = e(V_2 - V_1). \quad (9-65)$$

Таким образом, при наличии внешнего источника напряжения изменение квазиуровня Ферми определяется приложенной разностью 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 V падает в приповерхностном слое полупроводника, обладающем высоким сопротивлением. В этом случае положение уровня Ферми в объеме полупроводника относительно уровня E_c останется постоянным, но по отношению к положению уровня Ферми в металле будет смещено на величину eV . Следовательно, контактная разность потенциалов между металлом и полупроводником уменьшится на V и станет равной $(\varphi_k - V)$, а в приконтактной области

положение квазиуровня Ферми изменится от величины F_n до F_m (рис. 9-10, а). В результате снижения потенциального барьера со стороны полупроводника равновесие в системе нарушается и потечет ток, равный разности тока термоэлектронной эмиссии, направленного от металла к полупроводнику и в обратном направлении:

$$j = j_1 - j_2. \quad (9-66)$$

Снижение потенциального барьера со стороны полупроводника вызовет более интенсивный переход электронов из полупроводника в металл по сравнению с равновесным состоянием. Этот переход

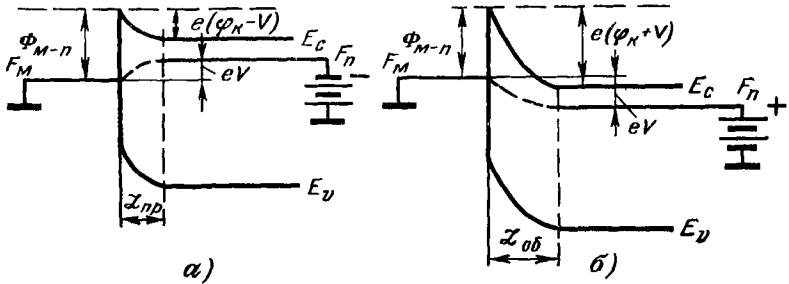


Рис. 9-10. Прямое (а) и обратное (б) подключение батареи к контакту металл—полупроводник.

электронов обусловит ток термоэлектронной эмиссии, направленный от металла к полупроводнику и равный:

$$j_1 = AT^2 e^{-(\Phi_n + e\phi_K - eV)/kT} = AT^2 e^{-(\Phi_m - eV)/kT}. \quad (9-67)$$

При подключении к контакту прямого смещения высота потенциального барьера со стороны металла не изменилась, поэтому поток электронов из металла в полупроводник при неравновесном состоянии будет тот же, что и в равновесном состоянии. Он создает ток термоэлектронной эмиссии, направленный от полупроводника к металлу и равный:

$$j_2 = AT^2 e^{-\Phi_m/kT}. \quad (9-68)$$

Результирующая плотность тока, текущего через контакт, будет определяться выражением вида

$$j = j_1 - j_2 = AT^2 e^{-\Phi_m/kT} (e^{eV/kT} - 1) = j_s (e^{eV/kT} - 1). \quad (9-69)$$

Величина

$$j_s = AT^2 e^{-\Phi_m/kT} \quad (9-70)$$

носит название т о к а п а с ы щ е н и я.

В случае обратного подключения батареи (рис. 9-10, б) со стороны полупроводника потенциальный барьер увеличится на приложенную разность потенциалов и выражение для тока, текущего через контакт, примет вид:

$$j = AT^2 e^{-(\Phi_{\text{п}} - e\varphi_{\text{к}} - eV)/kT} - AT^2 e^{-\Phi_{\text{м}}/kT} = j_s (e^{-eV/kT} - 1). \quad (9-71)$$

Если при прямом подключении батарей величину V считать положительной, а при обратном подключении отрицательной, равенства (9-69) и (9-71) можно записать как

$$j = j_s (e^{eV/kT} - 1). \quad (9-72)$$

Из (9-72) следует, что при прямом смещении ($V > 0$) ток через контакт в зависимости от напряжения растет по экспоненциальному закону, а при обратном смещении ($V < 0$) ток стремится к величине j_s . Таким образом, контакт металл — полупроводник обладает выпрямляющими свойствами.

Область полупроводника вблизи контакта представляет собой слой объемного заряда, толщина которого при отсутствии внешнего напряжения выражается формулой (9-54). В случае подключения батарей толщина слоя объемного заряда

$$\mathcal{L} = \sqrt{\frac{2\epsilon\epsilon_0 (\varphi_{\text{к}} - V)}{en_0}}, \quad (9-73)$$

где, как и в равенстве (9-72), знак V зависит от полярности приложенного напряжения. При прямом смещении ($V > 0$) толщина слоя $\mathcal{L}_{\text{пр}}$ объемного заряда меньше по сравнению с толщиной запирающего слоя \mathcal{L}_0 при термодинамическом равновесии, а в случае обратного смещения ($V < 0$) $\mathcal{L}_{\text{об}} > \mathcal{L}_0$.

9-6. ДИОДНАЯ ТЕОРИЯ ВЫПРЯМЛЕНИЯ

Для расчета вольт-амперной характеристики контакта металл—полупроводник используются две теории — диодная и диффузионная.

Если длина свободного пробега электронов l значительно меньше толщины запирающего слоя \mathcal{L}_0 , носители заряда, проходя через этот слой, многократно рассеиваются. В этом случае необходимо учитывать как диффузионную, так и дрейфовую составляющие тока. Использование закона Ома для нахождения дрейфового тока возможно в том случае, если электрическое поле слабое, условием чего является:

$$\mathcal{E}_s \ll \frac{kT}{el}. \quad (9-74)$$

Максимальная напряженность поля \mathcal{E}_s будет на границе полупроводника с металлом ($x = 0$). Она с учетом соотношений (9-52) и (9-54) равна:

$$\mathcal{E}_s = \mathcal{E}_{\text{макс}} = - \left. \frac{d\varphi}{dx} \right|_{x=0} = - \frac{en_0 \mathcal{L}_0}{\epsilon \epsilon_0} = - \frac{2\varphi_K}{\mathcal{L}_0}. \quad (9-75)$$

Подставляя найденное значение \mathcal{E}_s в неравенство (9-74), получаем условие применимости диффузионной теории выпрямления:

$$\mathcal{L}_0 \gg \frac{2e\varphi_K}{kT} l. \quad (9-76)$$

Если соотношение длины свободного пробега электронов l и толщины запирающего слоя \mathcal{L}_0 таково, что выполняется обратное неравенство

$$l \gg \frac{kT}{2e\varphi_K} \mathcal{L}_0, \quad (9-77)$$

то можно считать, что носители заряда пролетают запирающую область, почти не испытывая столкновений. В этом случае запирающий слой подобен вакуумному промежутку между электродами электронной лампы. В связи с этим теория, используемая в этом случае для расчета вольт-амперной характеристики контакта металл — полупроводник, получила название диодной.

Согласно неравенству (9-77) диодная теория применима в том случае, если

$$\frac{kT \mathcal{L}_0}{2e\varphi_K l} \ll 1. \quad (9-78)$$

Это условие будет выполняться тем лучше, чем меньше \mathcal{L}_0 и больше l . Но толщина запирающего слоя, как следует из соотношения (9-54), тем меньше, чем больше равновесная концентрация свободных электронов n_0 , а длина свободного пробега электронов l согласно формуле (6-29) тем больше, чем больше их подвижность. Следовательно, *диодная теория выпрямления применима для полупроводников с большой концентрацией носителей заряда, обладающих высокой подвижностью.*

Например, для германия электронного типа $\mu_n = 3900 \text{ см}^2/(\text{В} \cdot \text{с})$; $\epsilon = 16,5$; $m_n^* = 0,3 m_0$. При $n_0 = 10^{14} \text{ см}^{-3}$ $e\varphi_K = 0,3 \text{ эВ}$, $\mathcal{L}_0 = = 2,3 \cdot 10^{-4} \text{ см}$, а $l = 1,5 \cdot 10^{-5} \text{ см}$ и отношение $\frac{kT \mathcal{L}_0}{2e\varphi_K l} = 0,6 < 1$.

Если же повысить концентрацию электронов до $n_0 = 10^{16} \text{ см}^{-3}$, то $kT \mathcal{L}_0 / 2e\varphi_K l = 0,06$. Применение диодной теории выпрямления в этом случае полностью оправдано.

В полупроводниках с малым значением концентрации и подвижности носителей заряда дело обстоит по-иному. Например, в закиси меди p -типа $\mu_p = 60 \text{ см}^2/(\text{В} \cdot \text{с})$; $\epsilon = 8,75$. При $n_0 = 10^{14} \text{ см}^{-3}$ $e\varphi_K = = 0,5 \text{ эВ}$; $\mathcal{L}_0 = 2,2 \cdot 10^{-4} \text{ см}$, а $l = 4 \cdot 10^{-7} \text{ см}$ и $kT \mathcal{L}_0 / 2e\varphi_K l = 15 \gg 1$.

В данном случае для расчета вольт-амперной характеристики контакта металл — полупроводник должна применяться диффузионная теория выпрямления.

При рассмотрении контакта металла с полупроводником (§ 9-4) 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

$$D = e^{-\frac{2\delta}{\hbar} [2m^*(\Phi_0 - E_x)]^{1/2}} \quad (9-79)$$

Здесь $E_x = \frac{m^* v_x^2}{2}$, где v_x — скорость электронов в направлении x поверхности образца. Если принять, что $\Phi_0 \approx 1$ эВ; $\delta = 10^{-8}$ см, то при условии $E_x \ll \Phi_0$ получим $D = 1/3$. Это означает, что столь тонкие барьеры практически прозрачны для электронов. Поэтому в дальнейшем мы будем полагать, что все электроны, энергия которых превышает величину Φ_{m-p} , способны перейти в полупроводник.

Для перехода из полупроводника в металл электроны должны иметь в направлении x достаточную скорость, чтобы преодолеть потенциальный барьер высотой $e(\varphi_k - V)$. Ток, обусловленный такими электронами, согласно уравнению (9-23) будет выражаться соотношением

$$j_1 = e \int_{v_{x \text{ мин}}}^{\infty} \int_{-\infty}^{\infty} v_x dn = \frac{4\pi e m^* (kT)^2}{h^3} e^{(F - E_c)/kT} e^{-e(\varphi_k - V)/kT} \quad (9-80)$$

Подставляя сюда выражения для равновесной концентрации электронов в толще полупроводника

$$n_0 = 2 \left(\frac{2\pi m^* kT}{h^2} \right)^{3/2} e^{(F - E_c)/kT} \quad (9-81)$$

и их средней тепловой скорости

$$v_0 = \left(\frac{8kT}{\pi m^*} \right)^{1/2}, \quad (9-82)$$

получаем:

$$j_1 = \frac{1}{4} e n_0 v_0 e^{-e(\varphi_k - V)/kT} = \frac{1}{4} e n_s v_0 e^{eV/kT} \quad (9-83)$$

Здесь

$$n_s = n_0 e^{-e\varphi_k/kT} \quad (9-84)$$

есть концентрация электронов на поверхности полупроводника при $V = 0$.

Со стороны металла высота потенциального барьера остается та же самая (см. рис. 9-10), поэтому поток электронов из металла в полупроводник не зависит от величины внешнего смещения и равен потоку электронов из полупроводника в состоянии термодинамического равновесия:

$$j_2 = j_1|_{V=0} = \frac{1}{4} en_s v_0. \quad (9-85)$$

Суммарный ток через контакт будет определяться выражением

$$j = j_1 - j_2 = \frac{1}{4} en_s v_0 (e^{eV/kT} - 1) = j_s (e^{eV/kT} - 1), \quad (9-86)$$

где

$$j_s = \frac{1}{4} en_s v_0 \quad (9-87)$$

есть ток насыщения для диодной теории выпрямления, который, как это следует из (9-87), определяется тепловой скоростью и концентрацией электронов на поверхности полупроводника.

9-7. ДИФFUЗИОННАЯ ТЕОРИЯ ВЫПРЯМЛЕНИЯ

В диффузионной теории выпрямления при определении тока, текущего через контакт металл — полупроводник, учитываются как дрейфовая, так и диффузионная составляющие тока. В этом случае для расчета вольт-амперной характеристики поступим следующим образом. Умножим обе части уравнения (9-60) на величину $e^{e\varphi(x)/kT}$. Используя соотношение Эйнштейна и выражая напряженность поля через потенциал, будем иметь:

$$\begin{aligned} j e^{e\varphi(x)/kT} &= n \mu_n e^{e\varphi(x)/kT} \frac{d(e\varphi)}{dx} + \mu_n kT e^{e\varphi(x)/kT} \frac{dn}{dx} = \\ &= \mu_n kT \frac{d}{dx} [n e^{e\varphi(x)/kT}]. \end{aligned} \quad (9-88)$$

Принтегрируем обе части равенства (9-88) по x в пределах от 0 до \mathcal{L} :

$$j \int_0^{\mathcal{L}} e^{e\varphi(x)/kT} dx = \mu_n kT \int_0^{\mathcal{L}} \frac{d}{dx} [n e^{e\varphi(x)/kT}] dx. \quad (9-89)$$

Выразим отсюда плотность тока j , которая не зависит от x :

$$j = \frac{\mu_n kT [n(\mathcal{L}) e^{e\varphi(\mathcal{L})/kT} - n(0) e^{e\varphi(0)/kT}]}{\int_0^{\mathcal{L}} e^{e\varphi(x)/kT} dx}. \quad (9-90)$$

Для вычисления числителя соотношения (9-90) используем граничные условия. В объеме полупроводника поле отсутствует, поэтому

$$\varphi(\mathcal{L}) = 0; \quad n(\mathcal{L}) = n_0, \quad (9-91)$$

а на границе раздела с металлом ($x = 0$) потенциал равен:

$$\varphi(0) = \varphi_K - V. \quad (9-92)$$

Рассмотрим случай, когда величина внешнего смещения сравнительно невелика и в приконтактной области всегда имеется запирающий слой

$$\varphi_K - V \gg \frac{kT}{e}. \quad (9-93)$$

Можно показать, что при этом условии концентрация электронов на поверхности полупроводника практически не зависит от величины смещения, так что можно без большой ошибки положить:

$$n(0) = n_s = n_0 e^{-e\varphi_K/kT}. \quad (9-94)$$

С учетом граничных условий (9-91), (9-92) и (9-94) выражение (9-90) для тока принимает вид:

$$j = \frac{n_0 \mu_n kT (1 - e^{-eV/kT})}{\int_0^{\mathcal{L}} e^{e\varphi(x)/kT} dx}. \quad (9-95)$$

Для вычисления знаменателя дроби (9-95) произведем замену переменной интегрирования:

$$\int_0^{\mathcal{L}} e^{e\varphi(x)/kT} dx = \int_{\varphi_K - V}^0 e^{e\varphi(x)/kT} \left(\frac{d\varphi}{dx} \right)^{-1} d\varphi. \quad (9-96)$$

Наибольший вклад в величину интеграла будет вносить та область, где значение $\varphi(x)$ максимально, т. е. участок вблизи $x = 0$. С увеличением x функция, стоящая под знаком интеграла, в результате уменьшения $\varphi(x)$ будет резко падать и вкладом ее можно пренебречь. Поэтому для приближенного вычисления интеграла можно вынести за знак интеграла значение $\left(\frac{d\varphi}{dx} \right)^{-1}$ в точке $x = 0$:

$$\left(\frac{d\varphi}{dx} \right)^{-1} = -\mathcal{E}_s^{-1}. \quad (9-97)$$

На основании (9-97) будем иметь:

$$\begin{aligned} \int_{\varphi_K - V}^0 e^{e\varphi(x)/kT} \left(\frac{d\varphi}{dx} \right)^{-1} d\varphi &= -\mathcal{E}_s^{-1} \frac{kT}{e} [1 - e^{e(\varphi_K - V)/kT}] = \\ &= \mathcal{E}_s^{-1} \frac{kT}{e} e^{e(\varphi_K - V)/kT}. \end{aligned} \quad (9-98)$$

Здесь мы пренебрегли единицей, стоящей в скобках, так как в соответствии с условием (9-93)

$$e^{e(\varphi_k - V)/kT} \gg 1. \quad (9-99)$$

Подставляя в (9-95) выражение (9-98), находим:

$$j = en_0\mu_n\mathcal{E}_s e^{-e\varphi_k/kT} (e^{eV/kT} - 1). \quad (9-100)$$

Используя соотношение (9-94), получаем выражение вольт-амперной характеристики контакта металл — полупроводник для диффузионной теории выпрямления:

$$j = en_s\mu_n\mathcal{E}_s (e^{eV/kT} - 1) = j_s (e^{eV/kT} - 1). \quad (9-101)$$

Величина $j_s = en_s\mu_n\mathcal{E}_s = en_s v_{сдр}$, определяющая значение тока при больших обратных смещениях, в отличие от тока насыщения

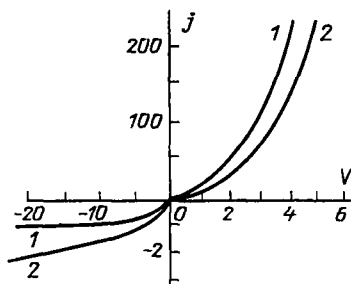


Рис. 9-11. Вольт-амперные характеристики контакта металл — полупроводник.

1 — диодная теория; 2 — диффузионная теория (в относительных единицах).

в диодной теории, зависит от приложенного напряжения, т. е. выражается через дрейфовую скорость электрона на поверхности полупроводника.

Вольт-амперные характеристики контакта металл — полупроводник для диодной и диффузионной теорий приведены на рис. 9-11 (кривые 1 и 2 соответственно).

9-8. КОНТАКТ ЭЛЕКТРОННОГО И ДЫРОЧНОГО ПОЛУПРОВОДНИКОВ

Рассмотрим явления, возникающие на контакте электронного и дырочного полупроводников. Для создания такого контакта в полупроводник вводятся как донорная, так и акцепторная примеси. При этом концентрации доноров и акцепторов меняются так, что в одной части образец содержит доноры и обладает электронной электропроводностью, а в другой части содержит акцепторы и обладает дырочной электропроводностью, а следовательно, в некоторой области кристалла происходит смена электропроводности с электронной на дырочную. Такой переход между материалами с электропроводностью противоположного типа называется *p-n переходом*.

Рассмотрим резкий p - n переход и предположим, что акцепторная область полупроводника легирована сильнее, чем электронная, т. е. $N_a > N_d$ (рис. 9-12, *a*). Пусть в p -области концентрация основных носителей заряда — дырок p_p и неосновных носителей заряда — электронов n_p , а в n -области концентрация электронов n_n и дырок p_n . Предположим также, что энергетические уровни доноров и акцепторов расположены так близко к зоне проводимости и валентной зоне, что все они полностью ионизированы. Тогда $p_p = N_a$, $n_n = N_d$ и для состояния термодинамического равновесия в случае отсутствия вырождения

$$p_p n_p = n_n p_n = n_i^2. \quad (9-102)$$

При образовании контакта между полупроводниками с разным типом электропроводности в области p - n перехода будет существовать большой градиент концентрации электронов и дырок. В результате начнется диффузия электронов из n -области в p -область и дырок из p -области в n -область. Возникновение диффузионных потоков приведет к разделению зарядов, вследствие чего появится положительный объемный заряд в n -области, примыкающей к переходу, и отрицательный — в p -области около перехода (рис. 9-12, *b*, *в*). Эти объемные заряды в области контакта создадут сильное электрическое поле, направленное от n -области к p -области и препятствующее движению электронов и дырок. В результате установится равновесное состояние, которое будет характеризоваться постоянством уровня Ферми для всего полупроводника, а в области перехода, где имеется электрическое поле, зоны энергии будут искривлены (рис. 9-12, *г*). Искривление зон энергии вызовет перераспределение концентрации электронов и дырок (рис. 9-12, *д*) и изменит ход электростатического потенциала в области p - n перехода (рис. 9-12, *е*).

Как следует из рис. 9-13, основные носители заряда при переходе через контакт должны преодолевать потенциальный барьер высотой $e\phi_K$. Переход неосновных носителей заряда совершается под действием электрического поля p - n перехода. В состоянии термодинамического равновесия диффузионный ток основных носителей заряда j_{op} и j_{on} уравновешен дрейфовым током неосновных носителей заряда j_{np} и j_{nn} , и суммарный ток через p - n переход равен нулю.

Предположим, что Φ_n есть работа выхода электронов из электронного полупроводника, а Φ_p — из дырочного. Тогда величина потенциального барьера на p - n переходе при термодинамическом равновесии определяется уравнением

$$e\phi_K = \Phi_p - \Phi_n = (\chi - E_c + F_p) - (\chi - E_c + F_n) = F_p - F_n. \quad (9-103)$$

Так как донорные и акцепторные примеси полностью ионизированы, то, используя соотношения (9-27) и (9-29), получаем:

$$e\phi_K = E_g - kT \ln \frac{N_v}{N_a} - kT \ln \frac{N_c}{N_d}. \quad (9-104)$$

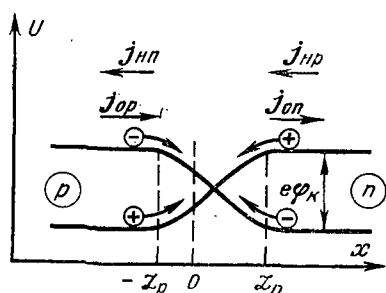
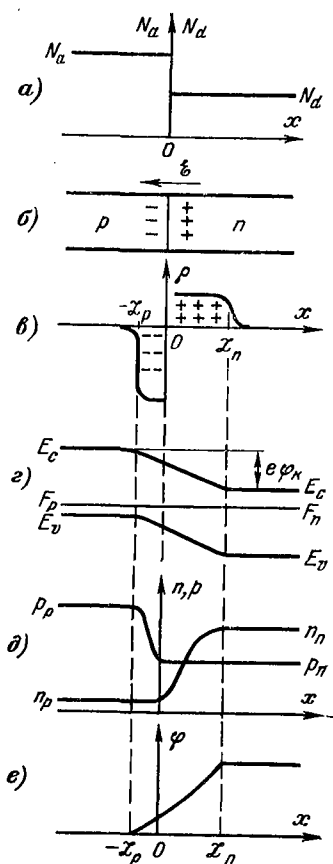


Рис. 9-13. Потенциальные барьеры для дырок и электронов в p - n переходе.

Рис. 9-12. Распределение примесей (а), разделение зарядов и возникновение электрического поля (б), распределение объемного заряда (в), зонная структура (г), распределение концентрации электронов и дырок (д) и изменение потенциала (е) в контакте электронного и дырочного полупроводников.

Учитывая, что

$$n_i^2 = N_c N_v e^{-E_g/kT}, \quad (9-105)$$

из (9-104) получим:

$$e\varphi_k = kT \ln \frac{n_n p_p}{n_i^2} \quad (9-106)$$

или

$$\frac{p_n}{p_p} = \frac{n_p}{n_n} = e^{-e\varphi_k/kT}. \quad (9-107)$$

Таким образом, контактная разность потенциалов на p - n переходе тем больше, чем сильнее легированы n - и p -области полупроводника; ее максимальное значение для невырожденного полупроводника равно:

$$\varphi_{k, \text{ макс}} = \frac{1}{e} E_g. \quad (9-108)$$

На основании (9-107) можно получить следующие соотношения для концентрации основных и неосновных носителей заряда:

$$p_n = p_p e^{-e\varphi_k/kT}; \quad n_p = n_n e^{-e\varphi_k/kT}. \quad (9-109)$$

Работа выхода электронов из полупроводника довольно велика (несколько электронвольт), и при комнатной температуре практически нет электронов, обладающих достаточной энергией, чтобы покинуть кристалл. Однако электроны вполне могут переходить из n -области непосредственно в p -область, преодолевая возникший потенциальный барьер. Например, для германия при $T = 300$ К и значениях концентрации $N_d = n_n = 10^{16}$ см $^{-3}$; $p_n = 10^{10}$ см $^{-3}$; $N_a = p_p = 10^{14}$ см $^{-3}$; $n_p = 10^{12}$ см $^{-3}$ величина потенциального барьера составляет:

$$e\varphi_k = kT \ln \frac{p_p}{p_n} = 0,25 \text{ эВ.}$$

Поскольку p -область легирована сильнее, чем n -область, то $p_p \gg n_n$ и, следовательно, толщина слоя объемного заряда низкоомной дырочной области \mathcal{L}_p будет меньше, чем электронной \mathcal{L}_n , т. е. $\mathcal{L}_n > |\mathcal{L}_p|$ (рис. 9-12, в), а вся область объемного заряда составит:

$$\mathcal{L}_0 = \mathcal{L}_n + \mathcal{L}_p. \quad (9-110)$$

В интервале $-\mathcal{L}_p \leq x < 0$ объемный заряд отрицательный, определяется концентрацией ионов акцепторной примеси:

$$\rho = -eN_a^- = -ep_p \quad (9-111)$$

и уравнение Пуассона для этой области будет иметь вид:

$$\frac{d^2\varphi}{dx^2} = \frac{ep_p}{\epsilon\epsilon_0}. \quad (9-112)$$

В интервале $0 < x < \mathcal{L}_n$ объемный заряд положительный, определяется ионами донорной примеси:

$$\rho = eN_d^+ = en_n \quad (9-113)$$

и уравнение Пуассона запишется в форме

$$\frac{d^2\varphi}{dx^2} = -\frac{en_n}{\epsilon\epsilon_0}. \quad (9-114)$$

Граничными условиями в этом случае являются:

$$\varphi(-\mathcal{L}_p) = 0; \quad \left. \frac{d\varphi}{dx} \right|_{x=-\mathcal{L}_p} = 0; \quad (9-115)$$

$$\varphi(\mathcal{L}_n) = \varphi_k; \quad \left. \frac{d\varphi}{dx} \right|_{x=\mathcal{L}_n} = 0. \quad (9-116)$$

Этим условиям удовлетворяют следующие выражения для φ , являющиеся согласно соотношению (9-52) решением уравнений (9-112) и (9-114):
при $-\mathcal{L}_p \leq x < 0$

$$\varphi_p = \frac{e p_p}{2 \varepsilon \varepsilon_0} (\mathcal{L}_p + x)^2; \quad (9-117)$$

при $0 < x \leq \mathcal{L}_n$

$$\varphi_n = \varphi_k - \frac{e n_n}{2 \varepsilon \varepsilon_0} (\mathcal{L}_n - x)^2. \quad (9-118)$$

При $x = 0$ потенциал (рис. 9-12, e) и его производная непрерывны, поэтому

$$\varphi_p(0) = \varphi_n(0); \quad (9-119)$$

$$\left. \frac{d\varphi_p}{dx} \right|_{x=0} = \left. \frac{d\varphi_n}{dx} \right|_{x=0}. \quad (9-120)$$

Применяя условие (9-120) к выражениям (9-117) и (9-118), получаем:

$$n_n \mathcal{L}_n = p_p \mathcal{L}_p. \quad (9-121)$$

Следовательно, в обеих областях полупроводника, прилегающих к p - n переходу, объемные заряды равны. Это является условием сохранения электронейтральности.

Из уравнения (9-121) нетрудно получить следующие соотношения:

$$\frac{\mathcal{L}_n}{\mathcal{L}_0} = \frac{p_p}{n_n + p_p}; \quad \frac{\mathcal{L}_p}{\mathcal{L}_0} = \frac{n_n}{n_n + p_p}. \quad (9-122)$$

Используя эти условия, на основании равенства (9-119) получаем:

$$\varphi_k = \frac{e}{2 \varepsilon \varepsilon_0} (n_n \mathcal{L}_n^2 + p_p \mathcal{L}_p^2) = \frac{e}{2 \varepsilon \varepsilon_0} \mathcal{L}_0 \frac{n_n p_p}{n_n + p_p}. \quad (9-123)$$

Отсюда следует выражение для полной толщины слоя объемного заряда p - n перехода:

$$\mathcal{L}_0 = \sqrt{\frac{2 \varepsilon \varepsilon_0}{e} \varphi_k \frac{n_n + p_p}{n_n p_p}}. \quad (9-124)$$

Из этой формулы видно, что чем выше степень легирования полупроводника, тем меньше толщина области объемного заряда \mathcal{L}_0 . Если одна из областей легирована значительно сильнее другой, то большая часть падения электростатического потенциала приходится на высокоомную область (рис. 9-12, e).

Как следует из рис. 9-12, d , в области p - n перехода имеет место значительное уменьшение концентрации носителей заряда. В силу этого сопротивление перехода велико по сравнению с сопротивлением слоя полупроводника n - или p - типа той же площади и толщины, что и область объемного заряда. Следовательно, *электронно-дырочный переход представляет собой слой низкой удельной проводимости, заключенный между областями высокой удельной проводимости, поэтому обладает свойствами конденсатора*. Емкость на единицу площади, называемая *барьерной емкостью*, может быть определена по формуле

$$C = \frac{\epsilon\epsilon_0}{\mathcal{L}_0} = \sqrt{\frac{\epsilon\epsilon_0}{2\varphi_K} \frac{n_n p_p}{n_n + p_p}}. \quad (9-125)$$

9-9. ВЫПРЯМЛЕНИЕ НА p - n ПЕРЕХОДЕ

Рассмотрим теперь p - n переход, к которому приложено внешнее напряжение V , минус — к электронному полупроводнику, а плюс — дырочному, т. е. приложено напряжение в прямом направлении (рис. 9-14, b). Поскольку сопротивление слоя объемного заряда перехода высокое, то падение напряжения будет в основном в этой области. Вследствие этого высота потенциального барьера снизится на величину eV по сравнению с равновесным состоянием (рис. 9-14, a), соответственно изменится и толщина запирающего слоя. Для ее вычисления нужно вместо формулы (9-124) использовать выражение

$$\mathcal{L}_{np} = \sqrt{\frac{2\epsilon\epsilon_0(\varphi_K - V)}{e} \frac{n_n + p_p}{n_n p_p}}. \quad (9-126)$$

Понижение потенциального барьера приведет к увеличению потока основных носителей заряда, так как большее число электронов из n -области будет переходить в p -область и большее число дырок из p -области в n -область, а поток неосновных носителей заряда через переход останется практически неизменным. В результате этого во внешней цепи будет протекать ток, равный разности токов основных и неосновных носителей заряда и направленный от p -области к n -области (рис. 9-14, b), т. е.

$$j = j_{он} - j_{ин}. \quad (9-127)$$

В n -области появившиеся избыточные неосновные носители заряда — дырки Δp создадут в первый момент вблизи контакта положительный объемный заряд. Однако через очень короткое время, определяемое максвелловским временем релаксации (см. § 7-3), этот заряд будет скомпенсирован объемным зарядом основных носителей заряда — электронов, которые под действием электрического поля, созданного избыточными дырками, будут подтянуты в количестве Δn из глубины n -области, а в n -область электроны

поступят из внешней цепи. Таким образом, во всех частях электронного полупроводника будет соблюдаться электронейтральность, но в приконтактной области к p - n переходу концентрация электронов и дырок будет повышена на величину $\Delta n = \Delta p$ по сравнению с равновесным состоянием. Введение в полупроводник носителей заряда с помощью p - n перехода при подаче на него прямого смещения в область, где эти носители заряда являются неосновными, называется инжекцией. Теперь концентрация дырок в n -области вблизи контакта будет равна:

$$p = p_n + \Delta p. \quad (9-128)$$

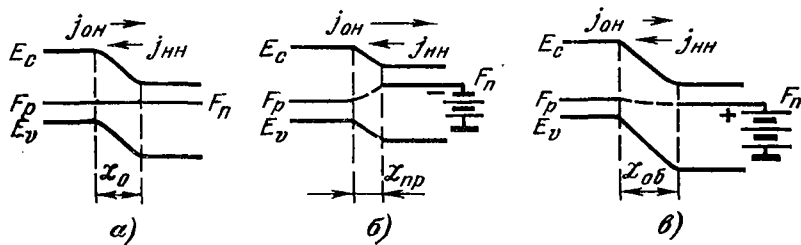


Рис. 9-14. Энергетическая диаграмма p - n перехода при термодинамическом равновесии (а), при подаче на переход прямого (б) и обратного (в) смещения.

Для ее нахождения в стационарном случае при $x = \mathcal{L}_n$ и отсутствии вырождения нужно в равенстве (9-109) вместо $e\varphi_k$ использовать значение

$$p(\mathcal{L}_n) = p_p e^{-e(\varphi_k - V)/kT} = p_n e^{eV/kT}, \quad (9-129)$$

откуда следует, что концентрация избыточных дырок в n -области при $x = \mathcal{L}_n$ равна:

$$\Delta p(\mathcal{L}_n) = p - p_n = p_n (e^{eV/kT} - 1). \quad (9-130)$$

Аналогичные явления происходят в p -области: сюда из n -области инжектируются электроны, и концентрация избыточных электронов при $x = -\mathcal{L}_p$ составит:

$$\Delta n(-\mathcal{L}_p) = n_p (e^{eV/kT} - 1). \quad (9-131)$$

Из выражений (9-130) и (9-131) следует, что с увеличением прямого смещения на p - n переходе концентрация инжектируемых неосновных носителей заряда резко возрастает, что приводит к сильному росту тока через контакт в прямом направлении.

Если внешнее напряжение приложено в обратном направлении (рис. 9-14, в), т. е. плюс к электронному полупроводнику, потенциальный барьер повышается и увеличивается толщина запирающего слоя. Чтобы найти ее значение, нужно в формулу (9-129) подставить значение напряжения V с отрицательным знаком. Чем

сильнее смещен переход в обратном направлении, тем меньшее количество основных носителей заряда способно преодолеть возросший потенциальный барьер. В соответствии с этим количество основных носителей заряда в приконтактной области уменьшается по сравнению с равновесным состоянием, уменьшается также количество неосновных носителей заряда вследствие соблюдения электронейтральности. Это явление носит название *экстракции носителей заряда*. Избыточная концентрация электронов в p -области по-прежнему также будет определяться равенством (9-131), но значение V теперь берется с отрицательным знаком. Таким образом, при обратном подключении батареи к p - n переходу ток основных носителей заряда будет меньше, чем при равновесном состоянии, а ток неосновных носителей заряда практически не изменяется. Поэтому суммарный ток через p - n переход будет направлен от n -области к p -области и с увеличением обратного напряжения вначале будет незначительно расти, а затем стремиться к некоторой величине, называемой током насыщения. Следовательно, p - n переход имеет нелинейную вольт-амперную характеристику (рис. 9-15).

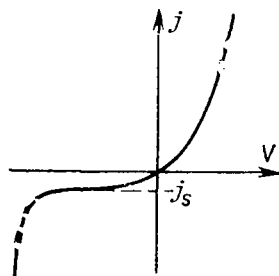


Рис. 9-15. Вольт-амперная характеристика p - n перехода.

9-10. 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 заряда пролетают через слой объемного заряда без рекомбинации. Это означает, что толщина заборного слоя значительно меньше диффузионной длины:

$$L_0 \ll L_D. \quad (9-132)$$

2. Обе области полупроводника сильно легированы, т. е. $p_p \gg n_i$ и $n_n \gg n_i$. Благодаря этому падением напряжения на образце можно пренебречь.

3. Контакты полупроводника с металлом, к которым приложена внешняя разность потенциалов, сделаны антизапорными и находятся далеко от перехода, так что неравновесные носители заряда рекомбинируют полностью, не доходя до них. Падением напряжения на этих контактах также можно пренебречь и считать, что вся внешняя разность потенциалов приложена к p - n переходу.

4. Ловушки захвата и поверхностная рекомбинация отсутствуют, и уменьшение избыточной концентрации носителей заряда происходит только в результате рекомбинации их в объеме полупроводника, которую будем считать линейной.

Чтобы рассчитать вольт-амперную характеристику p - n перехода, нужно найти закон изменения концентрации свободных носителей заряда, решив уравнения непрерывности для дырок и электронов, которые имеют вид:

$$\frac{\partial p}{\partial t} = -\frac{1}{e} \frac{\partial j_p}{\partial x} - \frac{\Delta p}{\tau_p}; \quad (9-133)$$

$$\frac{\partial n}{\partial t} = \frac{1}{e} \frac{\partial j_n}{\partial x} - \frac{\Delta n}{\tau_n}. \quad (9-134)$$

Рассмотрим дырочный ток в p - и n -областях, считая, что значения коэффициента диффузии и подвижности дырок в обеих областях полупроводника одинаковы (что, строго говоря, не совсем верно). Полный ток дырок определяется дрейфовой и диффузионной составляющими:

$$j_p = e p \mu_p \mathcal{E} - e D_p \frac{dp}{dx}, \quad (9-135)$$

где \mathcal{E} — напряженность внешнего электрического поля.

В p -области при прямом смещении p - n перехода концентрация неравновесных дырок $p = p_p + \Delta p$. Здесь $\Delta p = \Delta n$ — количество дырок, подтянувшихся к контакту из глубины p -области для нейтрализации заряда электронов, инжектированных из электронного полупроводника. Поскольку p -область сильно легирована, то $p_p \gg \Delta p$ и $p \approx p_p$, поэтому дрейфовая составляющая тока $j_{pдр}^{(p)}$ значительно превосходит диффузионную $j_{pдиф}^{(p)}$, т. е. $j_{pдр}^{(p)} \gg j_{pдиф}^{(p)}$. Следовательно, в p -области ток дырок $j_p^{(p)}$ примерно будет равен дрейфовой составляющей тока дырок:

$$j_p^{(p)} \approx j_{pдр}^{(p)} = e p \mu_p \mathcal{E}. \quad (9-136)$$

В n -области $n_n \gg p_n$, $\Delta p \gg p_n$ и концентрация неравновесных дырок вблизи p - n перехода, равная $p = p_n + \Delta p$, в основном определяется количеством избыточных дырок Δp , инжектированных из p -области. Поэтому диффузионная составляющая дырочного тока превышает дрейфовую, т. е. $j_{pдиф}^{(n)} \gg j_{pдр}^{(n)}$, и

$$j_p^{(n)} = j_{pдиф}^{(n)} = -e D_p \frac{dp}{dx}. \quad (9-137)$$

С учетом равенства (9-137) уравнение непрерывности для дырок в n -области для стационарного случая ($\frac{\partial p}{\partial t} = 0$) запишется в виде

$$D_p \frac{d^2 p}{dx^2} - \frac{p - p_n}{\tau_p} = 0. \quad (9-138)$$

Используя соотношение $L_p^2 = D_p \tau_p$, получаем:

$$\frac{d^2 \Delta p}{dx^2} - \frac{\Delta p}{L_p^2} = 0. \quad (9-139)$$

Общее решение такого уравнения имеет вид:

$$\Delta p = A e^{-x/L_p} + B e^{x/L_p}. \quad (9-140)$$

Так как при движении избыточных дырок в объем полупроводника ($x \rightarrow \infty$) концентрация их спадает до нуля, константа B должна быть принята равной нулю. Тогда

$$p = p_n + \Delta p = p_n + A e^{-x/L_p}. \quad (9-141)$$

Концентрация неравновесных дырок на границе запорного слоя при $x = \mathcal{L}_n$ согласно (9-129) равна:

$$p(\mathcal{L}_n) = p_n e^{eV/kT}. \quad (9-142)$$

Учитывая это, из (9-141) при $x = \mathcal{L}_n$ находим:

$$A = p_n (e^{eV/kT} - 1) e^{\mathcal{L}_n/L_p}. \quad (9-143)$$

Окончательно закон изменения концентрации неравновесных дырок в n -области при $x > \mathcal{L}_n$ принимает вид:

$$p(x) = p_n + p_n (e^{eV/kT} - 1) e^{-(x-\mathcal{L}_n)/L_p}, \quad (9-144)$$

и для дырочного тока на основании (9-137) получаем:

$$j_p^{(n)} = \frac{e D_p p_n}{L_p} (e^{eV/kT} - 1) e^{-(x-\mathcal{L}_n)/L_p}. \quad (9-145)$$

Проведя аналогичные расчеты, найдем, что в p -области изменение концентрации неравновесных электронов при $x < -\mathcal{L}_p$ определяется соотношением

$$n(x) = n_p + n_p (e^{eV/kT} - 1) e^{(x+\mathcal{L}_p)/L_n}, \quad (9-146)$$

а электронная составляющая тока имеет вид:

$$j_n^{(p)} = \frac{e D_n n_p}{L_n} (e^{eV/kT} - 1) e^{(x+\mathcal{L}_p)/L_n}. \quad (9-147)$$

В любом сечении полупроводника сумма электронного и дырочного токов постоянна:

$$j = j_p^{(p)} + j_n^{(p)} = j_n^{(n)} + j_p^{(n)} = \text{const}. \quad (9-148)$$

Поскольку слой объемного заряда достаточно узок и внутри его нет рекомбинации носителей заряда, то дырочные токи в p - и n -областях на границе запорного слоя одинаковы:

$$j_p^{(p)} \Big|_{x=-\mathcal{L}_p} = j_p^{(n)} \Big|_{x=\mathcal{L}_n}. \quad (9-149)$$

Принимая это во внимание, для общего тока, текущего через p - n переход, можно использовать формулу

$$j = j_p^{(p)} \Big|_{x=-\mathcal{L}_p} + j_n^{(p)} \Big|_{x=-\mathcal{L}_p} = j_p^{(n)} \Big|_{x=\mathcal{L}_n} + j_n^{(p)} \Big|_{x=-\mathcal{L}_p}. \quad (9-150)$$

Эти токи на основании (9-145) и (9-147) соответственно равны:

$$j_p^{(n)} \Big|_{x=x_n} = \frac{eD_p p_n}{L_p} (e^{eV/kT} - 1); \quad (9-151)$$

$$j_n^{(p)} \Big|_{x=-x_p} = \frac{eD_n n_p}{L_n} (e^{eV/kT} - 1). \quad (9-152)$$

Следовательно, вольт-амперная характеристика тонкого p - n перехода описывается уравнением вида

$$j = e \left(\frac{D_n n_p}{L_n} + \frac{D_p p_n}{L_p} \right) (e^{eV/kT} - 1) = j_s (e^{eV/kT} - 1), \quad (9-153)$$

где ток насыщения j_s равен:

$$\begin{aligned} j_s = j_{sn} + j_{sp} &= \frac{eD_n n_p}{L_n} + \frac{eD_p p_n}{L_p} = e n_i^2 \left(\frac{D_n}{L_n p_p} + \frac{D_p}{L_p n_n} \right) = \\ &= e \left(\frac{n_p L_n}{\tau_n} + \frac{p_n L_p}{\tau_p} \right). \end{aligned} \quad (9-154)$$

Кривые распределения концентрации равновесных и избыточных носителей заряда и токов через p - n переход, построенные в соответствии с формулами (9-144) — (9-152) для прямого и обратного смещений, изображены на рис. 9-16.

Из соотношения (9-153) видно, что при прямом смещении ток, текущий через p - n переход, экспоненциально возрастает, а с ростом отрицательного напряжения стремится к j_s . Выпрямляющие свойства p - n перехода тем лучше, чем меньше ток насыщения. Согласно (9-154) он уменьшается с ростом концентрации основных носителей заряда n_n и p_p (т. е. с увеличением степени легирования) и увеличением времени жизни неосновных носителей заряда τ_n и τ_p . Повышение температуры приводит к росту собственной концентрации n_i , а значит и значения тока насыщения j_s .

При дальнейшем увеличении обратного напряжения в некоторый момент происходит пробой p - n перехода и ток резко возрастает (рис. 9-15). Пробой p - n перехода может быть обусловлен тем, что сильное электрическое поле вызывает такой наклон энергетических зон, при котором в результате туннельного перехода сильно увеличивается концентрация свободных носителей заряда 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 заряда могут создаваться в результате разрыва валентных связей за счет кинетической энергии носителя заряда, полученной им от электрического поля при прохождении запирающего слоя. Возникающие свободные носители заряда 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.

При подключении к p - n переходу источника тока состояние полупроводника будет неравновесным, характеризующимся квазиуровнями Ферми F_n и F_p . Если внешняя разность потенциалов V и не слишком велика, то избыточные концентрации носителей заряда будут существовать справа и слева от перехода лишь на расстоянии нескольких диффузионных длин L_n и L_p и в этих областях квазиуровни Ферми будут зависеть от координаты x (рис. 9-17).

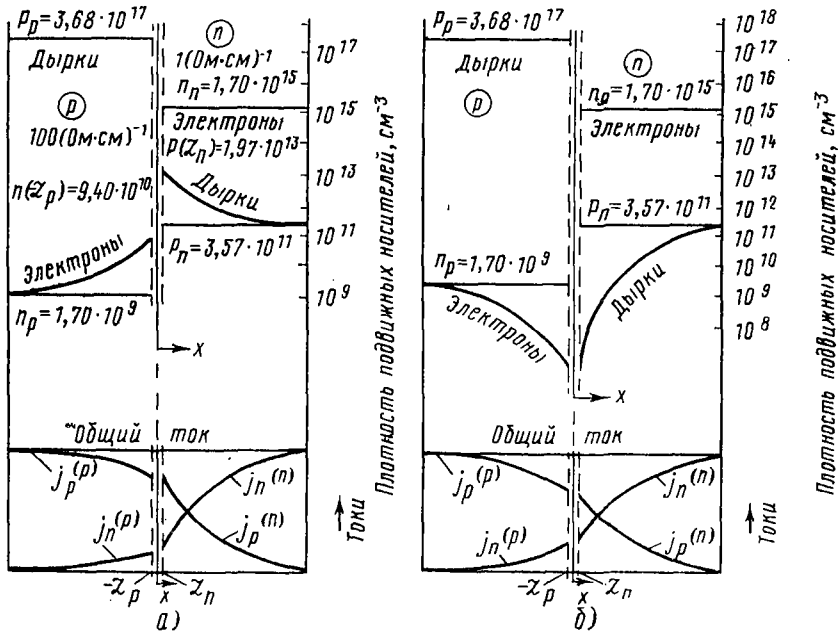


Рис. 9-16. Концентрация носителей заряда и токи в p - n переходе при прямом (а) и обратном (б) смещении.

Чтобы найти изменение положения квазиуровней Ферми, определим неравновесную концентрацию электронов в точке $x = -L_p$. Она будет равна:

$$n(-L_p) = n_p e^{eV/kT} = N_c e^{-\frac{E_c - F_n(-L_p)}{kT}}, \quad (9-155)$$

где $F_n(-L_p)$ — квазиуровень Ферми для электронов при $x = -L_p$.

На расстоянии трех диффузионных длин ($3L_n$) от контакта концентрация равновесных электронов равна n_p и определяется соотношением вида

$$n(3L_n) = n_p = N_c e^{-(E_c - F_p)/kT}. \quad (9-156)$$

Подставляя значение n_p (9-156) в выражение (9-155), находим:

$$F_p = F_n(-L_p) - eV. \quad (9-157)$$

Таким образом, при прямом смещении p - n перехода квазиуровень Ферми для электронов $F_n(x)$ в p -области на расстоянии $(2 \div 3) L_n$ понижается на величину eV , а квазиуровень Ферми для дырок $F_p(x)$ в n -области на расстоянии $(2 \div 3) L_p$ повышается на величину eV (рис. 9-17, а).

Следовательно, при прямом смещении p - n перехода дырочный квазиуровень Ферми в n -области приближается к верхнему краю валентной зоны, а квазиуровень Ферми для электронов в p -области поднимается ко дну зоны проводимости. Такое изменение квазиуровней Ферми и определяет возрастание концентраций неосновных носителей заряда у контакта.

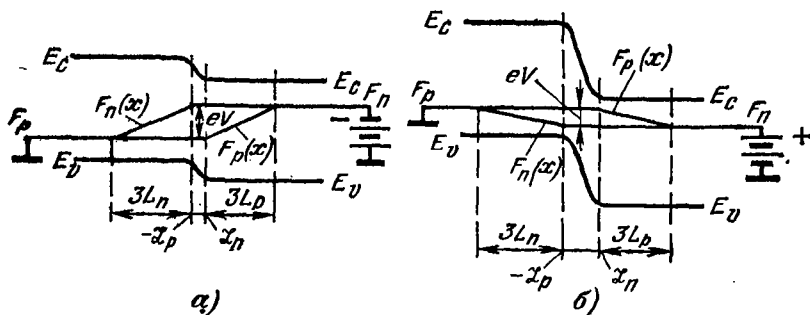


Рис. 9-17. Квазиуровень Ферми в p - n переходе при прямом (а) и обратном (б) смещении.

При обратном смещении p - n перехода имеет место уменьшение количества неосновных носителей заряда вблизи контакта и квазиуровни Ферми $F_n(x)$ и $F_p(x)$ в приконтактных областях удаляются от соответствующих зон (рис. 9-17, б).

Рассмотрим теперь соотношение токов, текущих через p - n переход. В соответствии с формулами (9-151) и (9-152) отношение электронной и дырочной составляющих токов через p - n переход равно:

$$\frac{j_n}{j_p} = \frac{j_{sn}}{j_{sp}} = \frac{D_n n_p L_p}{D_p p_n L_n} = \frac{\mu_n n_n L_p}{\mu_p p_p L_n} = \frac{\sigma_n L_p}{\sigma_p L_n}. \quad (9-158)$$

Из этого равенства следует, что отношение электронного и дырочного токов, текущих через p - n переход, определяется главным образом отношением концентраций основных носителей заряда в n - и p -областях. Если степень легирования обеих частей полупроводника приблизительно одинакова ($\sigma_n \approx \sigma_p$), то такой p - n переход инжектирует в равной степени электроны и дырки и общий ток определяется выражением

$$j = (j_{sn} + j_{sp}) (e^{eV/kT} - 1). \quad (9-159)$$

В том случае, когда p -область легирована значительно сильнее n -областью ($\sigma_p \gg \sigma_n$), то $j_{sp} \gg j_{sn}$ и через p - n переход течет в основном дырочный ток. При $\sigma_n \gg \sigma_p$, напротив, основной вклад в общий ток вносит электронная составляющая тока.

Глава десятая

ПОВЕРХНОСТНЫЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

10-1. ПРИРОДА ПОВЕРХНОСТНЫХ УРОВНЕЙ

В гл. 2 при рассмотрении энергетического спектра электрона в твердом теле в приближении сильносвязанных электронов мы предполагали, что везде в кристалле сохраняется строгая периодичность кристаллического потенциала, а ограниченность объема кристалла учитывалась введением циклических граничных условий Борна—Кармана. Последние приводили к дискретности энергетического спектра в пределах разрешенных зон энергий. Однако более детальное рассмотрение показывает, что влияние конечных размеров кристалла наличием границ не исчерпывается. Действительно, поверхность есть естественное нарушение периодичности потенциала и, следовательно, можно ожидать появления особенностей в спектре разрешенных состояний электрона кристалла. Таммом было показано, что обрыв периодичности кристаллического потенциала на поверхности приводит к появлению локализованных состояний, энергетические уровни которых располагаются в запрещенной зоне. Эти состояния называют **поверхностными состояниями** или **уровнями Тамма**. Поскольку обрыв потенциала происходит в каждой цепочке атомов, нормальной к поверхности, то, очевидно, плотность таммовских уровней равна плотности поверхностных атомов, т. е. 10^{15} см².

К такой же плотности состояний приводят поверхностные уровни, теоретически предсказанные Шокли и интерпретирующиеся как ненасыщенные валентности поверхностных атомов кристалла.

Оба типа состояний относятся к идеализированной модели поверхности и могут рассматриваться лишь применительно к так называемой атомарно-чистой поверхности. Однако даже в случае атомарно-чистой поверхности структура поверхности существенно отличается от идеализированной модели из-за перегруппировки поверхностных атомов, вызванной взаимным насыщением свободных валентных связей поверхностных атомов кристалла.

Реальная поверхность полупроводника, с которой приходится иметь дело, весьма далека от идеальной, ибо на ней практически всегда имеют место различного рода макроскопические и микроскопические структурные дефекты, связанные с условиями обработки (резка, полировка, шлифовка и т. д.) и роста кристалла (огранка

кристалла и другого рода особенности рельефа поверхности макроскопических и микроскопических размеров). Кроме того, реальная поверхность полупроводника находится в постоянном контакте с окружающей средой, различного рода химическими соединениями, применяемыми в качестве травителей, в результате чего на поверхности возможна адсорбция посторонних атомов и молекул из этих источников, появление окисных пленок как результат окислительно-растворительных реакций при травлении и т. д. Все это приводит к появлению локализованных на поверхности полупроводника состояний, которые в зависимости от степени сродства к электрону и дырке, положения уровня Ферми на поверхности могут проявлять себя как донорные или акцепторные ловушки захвата или рекомбинационные ловушки электронно-дырочных пар. При наличии, например, донорных состояний на поверхности, которые, как известно, могут быть либо нейтральными, либо заряженными положительно при отдаче электрона в зону проводимости, поверхность полупроводника будет заряжена положительно. При наличии акцепторных состояний поверхность полупроводника будет заряжена отрицательно, поскольку акцепторные состояния могут быть либо нейтральны (когда пусты), либо отрицательно заряжены (когда заполнены электронами). Величина заряда Q_{ss} на поверхности зависит от концентрации поверхностных состояний N_s и функций распределения для ловушек захвата, которые определяются значением электрохимического потенциала на поверхности полупроводника по отношению к энергетическому положению уровня ловушки захвата.

В условиях термодинамического равновесия полупроводник в целом электронейтрален. В приповерхностной области в присутствии заряда в поверхностных состояниях электронейтральность обеспечивается тем, что электрическое поле вблизи поверхности, вызванное зарядом Q_{ss} , приводит к перераспределению подвижных носителей заряда в приповерхностной области полупроводника, в результате чего в ней возникает пространственный заряд Q_{sp} , равный по величине и противоположный по знаку заряду в поверхностных состояниях. Таким образом, возникает двойной электрический слой, экранирующий объем полупроводника от действия поля.

В металле, где концентрация свободных электронов составляет 10^{22} см⁻³, нейтрализация поверхностного заряда происходит на расстоянии 10^{-8} — 10^{-7} см, и такой тонкий слой пространственного заряда не может существенно сказываться на свойствах всего кристалла. У полупроводников концентрация свободных носителей заряда значительно меньше, поэтому область пространственного заряда достаточно обширна. Например, для германия, у которого концентрация поверхностных уровней составляет 10^{11} см⁻², при концентрации носителей заряда 10^{15} см⁻³ нейтрализация поверхностного заряда происходит на расстоянии 10^{-4} см, а в собствен-

ном 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 , как это показано на рис. 10-1. При заполнении акцепторных поверхностных состояний

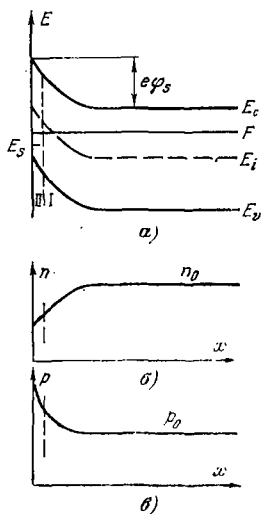


Рис. 10-1.

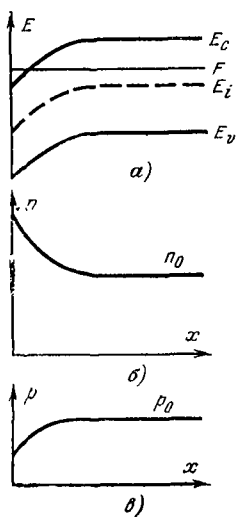


Рис. 10-2.

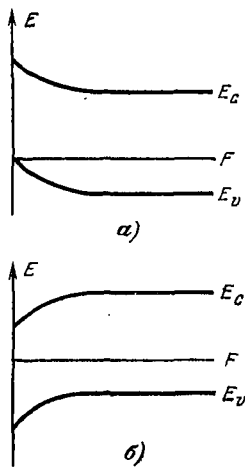


Рис. 10-3.

Рис. 10-1. Энергетическая диаграмма (а) и распределение концентрации электронов (б) и дырок (в) в приповерхностной области полупроводника n -типа.

I — область обеднения; II — область инверсии.

Рис. 10-2. Энергетическая диаграмма (а) и распределение концентрации электронов (б) и дырок (в) в приповерхностной области полупроводника n -типа в случае обогащения.

Рис. 10-3. Энергетическая диаграмма приповерхностной области полупроводника p -типа в случае обогащения (а) и обеднения (б).

электронами на поверхности полупроводника возникает отрицательный заряд, а в его приповерхностном слое при этом появится обедненный электронами слой, обладающий положительным пространственным зарядом. Наличие двойного электрического слоя приводит к появлению электрического поля, направленного к поверхности полупроводника, которое вызывает в приповерхностном слое полупроводника изгиб его энергетических зон вверх. Если через $e\phi_s$ обозначить величину изгиба зоны проводимости в поверхностной области, то ϕ_s — электростатический поверхностный потенциал. В приповерхностном слое такого полупроводника в за-

зависимости от положения границы его энергетических зон относительно уровня Ферми в общей сложности может быть область обеднения, для которой проводимость меньше, чем в объеме полупроводника, и область, в которой имеет место изменение типа проводимости, так называемый инверсионный слой (рис. 10-1, а, область I и II соответственно). Образование инверсионного слоя зависит от степени легирования образца.

В том случае, если у электронного полупроводника на поверхности имеются донорные состояния, то его энергетические зоны в приповерхностной области изгибаются вниз и образуется область обогащения. В приповерхностном слое такого полупроводника концентрация основных носителей заряда больше, чем в объеме (рис. 10-2).

У акцепторных полупроводников обеднение наблюдается в том случае, когда поверхностные состояния захватывают дырки и зоны изгибаются вниз, а при захвате электронов имеет место обогащение, при котором зоны изгибаются вверх (рис. 10-3).

10-2. ТЕОРИЯ СЛОЯ ПРОСТРАНСТВЕННОГО ЗАРЯДА

Поскольку в приповерхностной области полупроводника имеет место изгиб энергетических зон, то концентрация подвижных носителей заряда зависит от координаты в направлении, нормальном к поверхности (рис. 10-1, б, в). Для выяснения конкретного вида этих зависимостей необходимо решить уравнение Пуассона. В § 9-4 подобное решение проведено для случая обедненных слоев на контакте металл — полупроводник в пренебрежении вкладом неоснов-

ных носителей заряда в так называемом приближении обедненного слоя Шоттки, которое приводит к квадратичной зависимости потенциала от координаты (9-52). Это решение, однако, не описывает $\varphi(x)$ для случая обогащенных и инверсионных слоев. Поэтому необходимо провести решение уравнения Пуассона в общем виде, не пренебрегая вкладом в плотность объемного заряда подвижных носителей заряда.

Прежде чем перейти к решению уравнения Пуассона, введем обозначения, принятые при рассмотрении поверхностных явлений и смысл которых поясняется рис. 10-4. Удобно вести отсчет всех величин, характеризующих полупроводник в области пространственного заряда и вне ее, от F_i — уровня Ферми для собственного полупроводника [обычно F_i называют серединой запрещенной зоны, что, однако, полностью соответствует

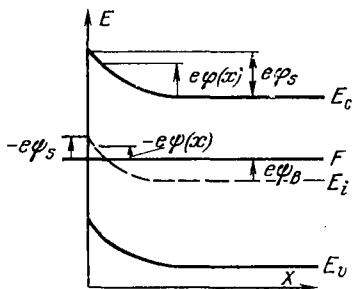


Рис. 10-4. Энергетическая диаграмма приповерхностной области полупроводника *n*-типа.

поверхностных явлений и смысл которых поясняется рис. 10-4. Удобно вести отсчет всех величин, характеризующих полупроводник в области пространственного заряда и вне ее, от F_i — уровня Ферми для собственного полупроводника [обычно F_i называют серединой запрещенной зоны, что, однако, полностью соответствует

действительности лишь в случае равенства эффективных масс плотностей состояния m_{dn}^* и m_{dp}^* (см. соотношение 4-88)]. Потенциал Ψ определяется в общем случае:

$$e\Psi = F - F_i, \quad (10-1)$$

где F — уровень Ферми для данного полупроводника. Электростатический потенциал в объеме полупроводника обозначим через Ψ_B , на его поверхности — через Ψ_s . Потенциал в произвольной точке пространственного заряда $\varphi(x) = \Psi(x) - \Psi_B$, на поверхности его значение $\varphi_s = \Psi_s - \Psi_B$. Концентрации электронов и дырок в области пространственного заряда могут быть выражены через φ и Ψ следующим образом:

$$\begin{aligned} n &= N_c e^{-(E_c + U - F)/kT} = n_0 e^{e\varphi/kT} = n_i e^{e\Psi_B/kT} e^{e(\Psi - \Psi_B)/kT} = \\ &= n_i e^{e\Psi/kT}. \end{aligned} \quad (10-2)$$

Так как $U = -e\varphi$ (в соответствии с рис. 10-1 и 10-4 $\varphi < 0$), то

$$p = p_0 e^{-e\varphi/kT} = n_i e^{-e\Psi_B/kT} e^{e(\Psi_B - \Psi)/kT} = n_i e^{-e\Psi/kT}. \quad (10-3)$$

Поверхностные концентрации электронов и дырок

$$n_s = n_0 e^{e\varphi_s/kT} = n_i e^{e\Psi_s/kT}; \quad (10-4)$$

$$p_s = p_0 e^{-e\varphi_s/kT} = n_i e^{-e\Psi_s/kT}. \quad (10-5)$$

Из соотношений (10-2) — (10-5) очевидно, что $\Psi < 0$ при изгибе зон вверх и $\Psi > 0$ при изгибе зон вниз. Если объемный Ψ_B и поверхностный Ψ_s электростатические потенциалы имеют одинаковые знаки, то приповерхностный слой обогащен основными носителями заряда, если же Ψ_B и Ψ_s имеют разные знаки, то имеют место слои обеднения или инверсии; последний будет иметь место в том случае, когда электростатический потенциал Ψ_s меняет знак и концентрация неосновных носителей заряда на поверхности больше, нежели основных (см. рис. 10-4 и 10-1).

Уравнение Пуассона имеет вид:

$$\frac{d^2\varphi}{dx^2} = -\frac{\rho(x)}{\epsilon\epsilon_0}, \quad (10-6)$$

где ϵ , ϵ_0 — диэлектрическая проницаемость полупроводника и вакуума соответственно.

Полагая, что при всех значениях $\varphi(x)$ примесь N_a и N_d являются полностью ионизированными и равномерно распределены по всему объему полупроводника, для $\rho(x)$ при произвольном значении x можно записать:

$$\rho(x) = -e(n - p + N_a - N_d). \quad (10-7)$$

Принимая во внимание условие электронейтральности для объема полупроводника

$$n_0 - p_0 + N_a - N_d = 0, \quad (10-8)$$

выражение (10-7) может быть трансформировано в

$$\rho(x) = -e[(n - n_0) - (p - p_0)]. \quad (10-9)$$

Используя (10-2) и (10-3), можно получить $\rho(x)$ в функции поверхностного электростатического потенциала:

$$\rho(x) = -e[n_0(e^{e\Phi/kT} - 1) - p_0(e^{-e\Phi/kT} - 1)]. \quad (10-10)$$

Введем обозначения

$$Y = \frac{e\Phi}{kT}; \quad \lambda = \frac{n_0}{n_i} = \frac{n_i}{p_0} = e^{e\Phi_B/kT}; \quad L_d^2 = \frac{\epsilon\epsilon_0 kT}{2e^2 n_i}. \quad (10-11)$$

Величина Y есть безразмерный электростатический потенциал; Y отрицателен при изгибах зон вверх и положителен при изгибах зон вниз; $\lambda > 1$ для полупроводника n -типа, этому соответствует $\Phi_B > 0$; $\lambda < 1$ для полупроводника p -типа, объемный электростатический потенциал является величиной отрицательной; L_d по физическому смыслу есть дебаевский радиус экранирования для собственного полупроводника.

Используя обозначения (10-11) и подставляя (10-10) в (10-6), получаем:

$$2 \frac{d^2 Y}{dx^2} = L_d^{-2} [\lambda (e^Y - 1) - \lambda^{-1} (e^{-Y} - 1)]. \quad (10-12)$$

Граничные условия для решения (10-12) следующие:
при $x=0$

$$Y = Y_s, \quad (10-13)$$

а при $x \rightarrow \infty$

$$Y = 0 \quad \text{и} \quad \frac{dY}{dx} = 0. \quad (10-14)$$

Первый интеграл (10-12) находим, используя тождество

$$\frac{d}{dx} \left(\frac{dY}{dx} \right)^2 = 2 \frac{d^2 Y}{dx^2} \frac{dY}{dx}. \quad (10-15)$$

Умножая правую и левую части уравнения (10-12) на $\frac{dY}{dx}$ и интегрируя, а затем извлекая квадратный корень, получаем:

$$\frac{dY}{dx} = \pm L_d^{-1} F(\lambda, Y) + C, \quad (10-16)$$

где

$$F(\lambda, Y) = [\lambda (e^Y - 1) - \lambda^{-1} (e^{-Y} - 1) + (\lambda^{-1} - \lambda) Y]^{1/2}. \quad (10-17)$$

Для отрицательных значений $Y \frac{dY}{dx} > 0$ и, следовательно, перед $F(\lambda, Y)$ необходимо выбрать плюс; а для положительных значений Y перед $F(\lambda, Y)$ необходимо выбрать знак минус. Использование граничного условия (10-14) приводит к $C = 0$. Для энергетической диаграммы рис. 10-4 $Y < 0$, поэтому

$$\frac{dY}{dx} = L_d^{-1} F(\lambda, Y). \quad (10-18)$$

Определим полный положительный заряд Q_{sp} в приповерхностной области полупроводника. Используя соотношение для L_d и (10-18), будем иметь:

$$\begin{aligned} Q_{sp} &= \int_0^{\infty} \rho(x) dx = -\frac{\epsilon\epsilon_0 kT}{e} \int_0^{\infty} \frac{d^2 Y}{dx^2} dx = \\ &= \frac{\epsilon\epsilon_0 kT}{e} \frac{dY}{dx} \Big|_{x=0} = 2en_i L_d F(\lambda, Y_s). \end{aligned} \quad (10-19)$$

Таким образом, величина заряда в приповерхностной области Q_{sp} определяется значением поверхностного потенциала Y_s и уровнем легирования полупроводника (рис. 10-5).

В ходе зависимости $Q_{sp}(Y_s)$, определяемой из (10-17), можно выделить следующие характерные участки. Для определенности будем полагать, что полупроводник n -типа, т. е. $\lambda \gg 1$, и достаточно сильно легированный, так что $\lambda \gg \lambda^{-1}$.

1. При больших положительных значениях Y в соотношении (10-17) преобладающим является член λe^Y , знак Q_{sp} отрицательный и, следовательно, концентрация основных носителей заряда в приповерхностной области полупроводника выше, чем в объеме ($n_s > n_0$). Это режим обогащения (см. рис. 10-2).

2. При $Y = 0$ и $Q_{sp} = 0$, а значит изгиб зон отсутствует. Это случай «плоских зон».

3. При малых отрицательных Y [диапазон изменений Y определяется условием $(1 + Y)\lambda > \lambda^{-1} e^{-Y}$] скорость убывания концентрации основных носителей заряда — электронов больше скорости нарастания концентрации дырок, вследствие чего приповерхностный слой полупроводника обедняется подвижными носителями заряда. Это режим обеднения (область I, рис. 10-1, а). Положительный пространственный заряд в этом случае создается в основном ионами донорной примеси.

Если вспомнить, что при выводе всех соотношений мы полагаем полную ионизацию примеси при любых значениях Y , то увеличение Q_{sp} в этом диапазоне означает одновременное увеличение ширины области пространственного заряда, которое происходит до тех пор, пока преимущественный вклад в Q_{sp} не будет вносить дырки. Под действием поля, соответствующего отрицательным зна-

чениям Y (направление этого поля из полупроводника к поверхности), подвижные дырки формируют слой положительного объемного заряда, непосредственно прилегающий к поверхности. Образуется инверсионный слой (область II, рис. 10-1, а).

4. Условием образования инверсионного слоя является преобладание члена $\lambda^{-1}e^{-Y}$ над всеми остальными членами в соотношении (10-17). Здесь заряд Q_{sp} , как и в случае обогащенного слоя, экспоненциально возрастает с ростом Y .

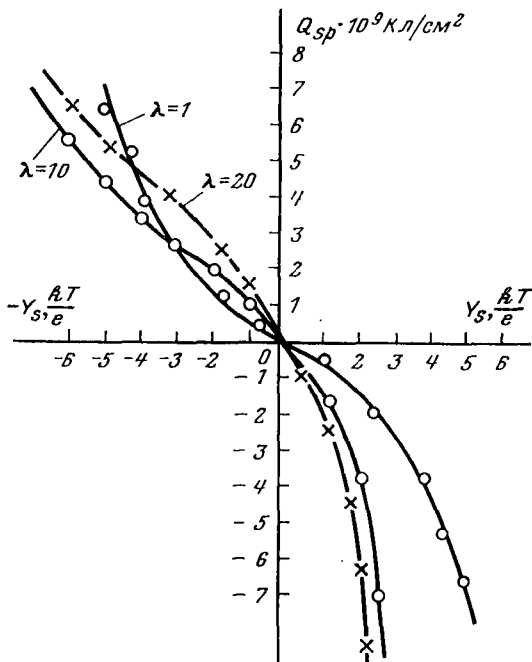


Рис. 10-5. Зависимость заряда в приповерхностном слое от безразмерного электростатического потенциала при различных уровнях легирования полупроводника.

График зависимости пространственного заряда Q_{sp} от величины безразмерного электростатического потенциала Y_s , рассчитанный с использованием (10-17), представлен на рис. 10-5 при λ в качестве параметра.

Как уже отмечалось, в случае свободной поверхности и отсутствия внешних полей заряд в области пространственного заряда Q_{sp} равен и противоположен по знаку заряду в поверхностных состояниях Q_{ss} . Величина и знак Q_{ss} определяются характером поверхностных состояний (акцепторные или донорные), числом состояний на каждом уровне и положением уровня Ферми на поверхности. В принятых обозначениях на рис. 10-4 $e\Psi_s = F - F_L$. Если считать, что E_s есть энергия поверхностного уровня, отсчитанная от F_L ,

то поверхностный заряд Q_{ss} для полупроводника, имеющего, например, поверхностные акцепторные уровни E_s с концентрацией N_s , равен:

$$Q_{ss} = - \frac{eN_s}{e^{(E_s - e\psi_s)/kT} + 1}. \quad (10-20)$$

Следовательно, используя (10-18) и (10-20), можно записать:

$$Q_{sp} = 2en_i L_d F(\lambda, Y_s) = -Q_{ss} - \frac{eN_s}{e^{(E_s - e\psi_s)/kT} + 1}. \quad (10-21)$$

Выражение (10-21) можно использовать для нахождения значения электростатического потенциала свободной поверхности, если известны уровень легирования полупроводника λ , концентрация N_s и энергия E_s поверхностного уровня. При изучении поверхностных свойств полупроводников чаще приходится решать обратную задачу — по известной величине Q_{ss} в функции Y_s ищутся (а точнее, подбираются) N_s и E_s . Для нахождения $Q_{ss}(Y_s)$ используется наиболее распространенный метод исследования параметров поверхности полупроводников — метод эффекта поля.

Для получения вида зависимости Y от x необходимо проинтегрировать уравнение (10-18) с граничными условиями $Y = Y_s$ при $x = 0$ и $Y = 0$ при $x \rightarrow \infty$:

$$\int_{Y_s}^0 \frac{dY}{\mp F(\lambda, Y)} = \frac{x}{L_d}. \quad (10-22)$$

Интеграл в (10-22) при $F(\lambda, Y)$, в общем виде определяемый соотношением (10-17), не может быть вычислен в квадратурах. Поэтому для выяснения хода зависимости рассматривают частные случаи, в которых возможно получить явный вид зависимости Y от x .

10-3. ЭФФЕКТ ПОЛЯ

Из предыдущего параграфа следует, что изменением поверхностного электростатического потенциала можно в широких пределах изменять концентрацию подвижных носителей заряда в слое пространственного заряда полупроводника. Обозначим через ΔN и ΔP поверхностные избытки электронов и дырок, которые определим как

$$\Delta N = \int_0^{\infty} [n(x) - n_0] dx; \quad (10-23)$$

$$\Delta P = \int_0^{\infty} [p(x) - p_0] dx. \quad (10-24)$$

Значения ΔN и ΔP могут быть вычислены в функции поверхностного электростатического потенциала Y_s с использованием (10-18) следующим образом:

$$\begin{aligned} \Delta N &= \int_0^{\infty} [n(x) - n_0] dx = n_0 \int_0^{\infty} (e^Y - 1) dx = \\ &= n_0 \int_{Y_s}^0 (e^Y - 1) \left(\frac{dY}{dx} \right)^{-1} dY = n_0 L_d \int_0^{Y_s} \frac{e^Y - 1}{F(\lambda, Y)} dY; \end{aligned} \quad (10-25)$$

$$\begin{aligned} \Delta P &= \int_0^{\infty} [p(x) - p_0] dx = p_0 \int_0^{\infty} (e^{-Y} - 1) dx = \\ &= p_0 \int_0^{Y_s} (e^{-Y} - 1) \left(\frac{dY}{dx} \right)^{-1} dY = -p_0 L_d \int_{Y_s}^0 \frac{e^{-Y} - 1}{F(\lambda, Y)} dY = \\ &= p_0 L_d \int_0^{Y_s} \frac{e^{-Y} - 1}{F(\lambda, Y)} dY. \end{aligned} \quad (10-26)$$

Избыточные электроны и дырки могут двигаться вдоль поверхности при приложении электрического поля, обуславливая тем самым изменение проводимости образца. Вызванная наличием ΔN и ΔP избыточная проводимость носит название поверхностной проводимости. Если предположить, что значение подвижностей электронов и дырок в области пространственного заряда равно их значениям в объеме полупроводника, то для поверхностной проводимости можно написать:

$$G_s = e\mu_n \Delta N + e\mu_p \Delta P. \quad (10-27)$$

Поскольку ΔN и ΔP являются функциями уровня легирования и поверхностного электростатического потенциала Y_s , то и G_s является функцией тех же величин. Однако характер этой зависимости не столь уже очевиден, если ΔN и ΔP определяются выражениями (10-25) и (10-26). Интегралы, входящие в (10-25) и (10-26), табулированы при различных Y и λ , что облегчает задачу вычисления G_s по (10-27).

На рис. 10-6 представлена зависимость $G_s(Y_s)$ с λ в качестве параметра, где $\lambda > 1$, что соответствует полупроводнику n -типа. На рисунке выделены характерные участки этой зависимости.

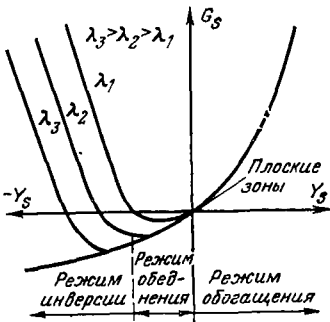


Рис. 10-6. Поверхностная проводимость в функции безразмерного электростатического потенциала при различных уровнях легирования полупроводника.

При $Y_s = 0$ — случай «плоских зон» и поверхностная проводимость $G_s = 0$.

Для положительных значений $Y_s > 0$ (изгиб зон «вниз», рис. 10-2) концентрация основных носителей заряда в области пространственного заряда растет, поверхностная проводимость $G_s > 0$ и монотонно возрастает с ростом Y_s (правая часть рис. 10-6). Это соответствует режиму обогащения.

При отрицательном Y_s поверхностная проводимость G_s отрицательна и с ростом $|Y_s|$ вплоть до $Y_s = 2e\psi_B/kT$ уменьшается, поскольку концентрация подвижных носителей заряда в слое пространственного заряда меньше, чем в случае $Y_s = 0$. Это соответствует режиму обеднения. Эта ситуация будет иметь место до тех пор, пока скорость нарастания неосновных носителей заряда — в данном случае дырок — с изменением Y_s не станет больше, нежели скорость убывания концентрации основных носителей заряда — электронов. Это соответствует условию

$$n_0 = p_s, \quad (10-28)$$

т. е. концентрация электронов в объеме полупроводника равна концентрации дырок на поверхности. Значение поверхностной проводимости, соответствующее этому условию, будет минимально и может быть найдено из условия экстремума

$$\frac{dG_s}{dY_s} = e \left[\mu_n n_0 L_d \frac{e^{Y_s} - 1}{F(\lambda, Y_s)} + \mu_p p_0 L_d \frac{e^{-Y_s} - 1}{F(\lambda, Y_s)} \right] = 0. \quad (10-29)$$

При достаточно высоком уровне легирования $\left| \frac{e\psi_B}{kT} \right| \gg 1$ в нашем случае минимум G_s будет наблюдаться при достаточно больших по абсолютной величине значениях поверхностного электростатического потенциала Y_s и, следовательно, выражение (10-29) может быть упрощено при использовании неравенств $e^{Y_s} \ll 1$ и $e^{-Y_s} \gg 1$ для полупроводника n -типа и $e^{Y_s} \gg 1$ и $e^{-Y_s} \ll 1$ для полупроводника p -типа (следует помнить о знаке Y_s). С учетом этого соотношение (10-29) примет вид:

$$\mu_p p_0 e^{-Y_s} - \mu_n n_0 = 0. \quad (10-30)$$

Поскольку $n_0 = n_i e^{e\psi_B/kT}$ и $p_0 = n_i e^{-e\psi_B/kT}$, то

$$\frac{\mu_n}{\mu_p} e^{\frac{2e\psi_B}{kT}} = e^{-Y_s}. \quad (10-31)$$

При $\mu_n = \mu_p$ условие (10-31) соответствует (10-28), последнее обычно используют в качестве критерия образования инверсионного слоя. При этом выражение (10-28) тем точнее, чем лучше для данного полупроводника выполняется условие $\mu_n = \mu_p$.

Последующее увеличение $|Y_s|$ приводит к резкому увеличению концентрации дырок, образованию инверсионного слоя и, следовательно, к увеличению поверхностной проводимости G_s . Все эти участки кривой $G_s(Y_s)$ приведены на рис. 10-6.

До этого мы полагали, что значения подвижностей носителей заряда в объеме и в слое пространственного заряда вблизи поверхности равны. Однако это не так в том случае, когда при соударении с поверхностью носитель заряда полностью или частично теряет дрейфовую составляющую скорости. Это так называемое диффузное рассеяние, в той или иной степени имеющее место в подавляющем числе практических случаев. Диффузное рассеяние приводит к уменьшению подвижности носителя заряда, движущегося в потенциальной яме (основного носителя заряда в случае слоев обогащения и неосновного — в случае слоев инверсии). Физической причиной снижения подвижности является то, что появление поля, нормального к поверхности, приводит к увеличению составляющей скорости к поверхности и, как результат этого, уменьшается время соударения носителя заряда с поверхностью.

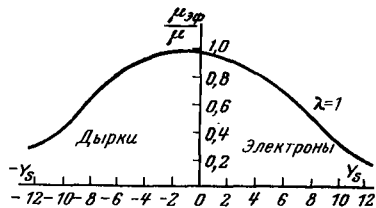


Рис. 10-7. Влияние поверхностного рассеяния на подвижность носителей заряда в области поверхностного заряда.

На рис. 10-7 представлена зависимость подвижности носителей заряда в слое пространственного заряда от поверхностного электростатического потенциала в случае полностью диффузного рассеяния.

Наиболее эффективным способом управления значением поверхностного электростатического потенциала (а следовательно, и значением поверхностной проводимости G_s и емкости пространственного заряда полупроводника C_{sp}) является приложение к полупроводнику электрического поля, нормального к поверхности полупроводника. Это сравнительно просто достигается в структуре, где полупроводник служит одной из обкладок конденсатора, отдаленной от второй обкладки — металлического электрода — слоем диэлектрика (так называемая МДП-структура). Эффектом поля называют явление изменения проводимости полупроводникового образца под действием электрического поля, нормального к его поверхности. На рис. 10-8 представлены МДП-структура, используемая обычно при измерениях эффекта поля, и соответствующая ей энергетическая диаграмма (рис. 10-9).

Экспериментально снимается зависимость изменения проводимости образца при вариации значения потенциала V_g , приложенного к металлическому электроду. Из диаграммы, представленной на рис. 10-9, следует:

$$V_g = V_{окс} + \varphi_s, \quad (10-32)$$

где $V_{окс}$ — падение потенциала в диэлектрике; $\varphi_s = \frac{kT}{e} Y_s$.

Если обозначить емкость диэлектрика через $C_{\text{окс}} = \frac{\epsilon \epsilon_{\text{окс}}}{d}$, где d — толщина окисла; $\epsilon_{\text{окс}}$ — диэлектрическая проницаемость окисла, то индуцированный в полупроводник заряд $Q_{\text{инд}} = C_{\text{окс}} V_{\text{окс}}$. В простейшем случае, когда контактной разностью потенциалов между металлом и полупроводником можно пренебречь в сравнении с V_q , а на границе раздела полупроводник — диэлектрик и в объеме диэлектрика отсутствуют состояния, способные захватить носители заряда, весь индуцированный заряд

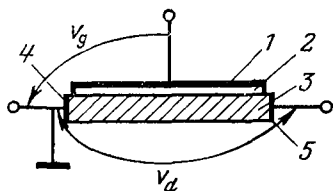


Рис. 10-8. Схематическое изображение структуры, используемой для измерения эффекта поля.

1 — металлический электрод (затвор);
2 — диэлектрик; 3 — полупроводник;
4 — омический контакт (исток); 5 — омический контакт (сток).

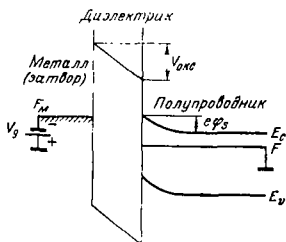


Рис. 10-9. Энергетическая диаграмма структуры, изображенной на рис. 10-8.

будет участвовать в создании дополнительной проводимости. Тогда можно записать:

$$Q_{\text{инд}} = C_{\text{окс}} V_{\text{окс}} = C_{\text{окс}} (V_q - \phi_s) = 2en_i L_d F(\lambda, Y_s). \quad (10-33)$$

Из соотношения (10-33) можно найти ϕ_s , соответствующий данному значению V_q , и, следовательно, значение G_s в функции Y_s . При выполнении вышеуказанных условий экспериментальная кривая $G_s(V_q)$ дает возможность при известном знаке V_q однозначно определить значение электростатического потенциала. Когда проводимость приповерхностного слоя носит монополярный характер (это имеет место в ярко выраженных слоях обогащения и инверсии), можно ввести в рассмотрение подвижность при эффекте поля $\mu_{\text{э.п}}$, которую определяют как отношение приращения проводимости к полному индуцированному заряду:

$$\mu_{\text{э.п}} = \frac{dG_s}{dQ_{\text{инд}}} = \frac{d(e\Delta N \mu_n)}{d(e\Delta N)}. \quad (10-34)$$

Здесь принято во внимание, что при $|Y_s| \gg 1$ $Q_{\text{инд}} = C_{\text{окс}} V_{\text{окс}} \approx e\Delta N$. При отсутствии явлений захвата $\mu_{\text{э.п}}$ равна омической подвижности носителей заряда в слое пространственного заряда.

В опытах по эффекту поля обычно стремятся выполнить условие $C_{\text{окс}} \ll C_{\text{сп}}$, что дает возможность величину индуцированного за-

ряда определять как $Q_{\text{инд}} = C_{\text{окс}} V_q$ (т. е. $V_q \gg \varphi_s$) и экспериментальную зависимость строить в координатах $G_s(Q_{\text{инд}})$. Теоретическая зависимость в тех же координатах при известных подвижностях электронов и дырок и уровне легирования полупроводника легко строится по схеме

$$C_{\text{окс}} V_q \rightarrow Q_{\text{инд}}(Y_s) \rightarrow V_s \rightarrow G_s(Y_s).$$

Реально наблюдаемая в экспериментах по эффекту поля зависимость поверхностной проводимости от индуцированного заряда существенно отличается от теоретической. Причины этого состоят в следующем: 1) между полупроводником и металлом существует контактная разность потенциала $e\varphi_k = \Phi_m - \Phi_p$; 2) на границе раздела полупроводник — диэлектрик и в самом диэлектрике практически всегда существуют состояния, способные нести нескомпен-

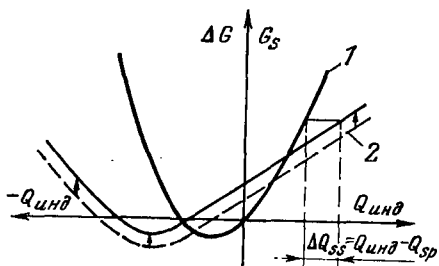


Рис. 10-10. Поверхностная проводимость в зависимости от индуцированного в полупроводнике заряда. (Стрелки поясняют методику расчета параметров поверхности из эксперимента).

1 — теоретическая кривая; 2 — экспериментальная кривая.

сированный заряд определенного знака и не меняющие своего зарядового состояния при всех достижимых в данном эксперименте значениях поверхностного электростатического потенциала Y_s ; 3) имеются такие состояния, заполнение которых электронами зависит от значения Y_s . Первые две из указанных причин приводят к тому, что в полупроводнике существует поверхностная проводимость ($G_s \neq 0$) в отсутствие напряжения на металлическом электроде ($V_q = 0$) и, следовательно, происходит параллельный сдвиг экспериментальной кривой $G_s(Q_{\text{инд}})$ относительно теоретической вдоль оси зарядов на величину $Q_{\text{с0}}$. Влияние третьей из указанных причин приводит к гораздо более существенной деформации экспериментальной кривой: кроме сдвига наблюдается изменение формы кривой (рис. 10-10).

Таким образом, количественное и качественное различие между теоретической и экспериментальной кривыми поверхностной проводимости от индуцированного заряда позволяет судить о концентрации и энергетическом спектре поверхностных состояний в структуре полупроводник — диэлектрик. Это обычно и является основной задачей метода эффекта поля. При наличии поверхностных состояний лишь часть индуцированного в полупроводник заряда

Q_{sp} участвует в создании поверхностной проводимости, другая часть, т. е. Q_{ss} , захватывается на поверхностные состояния:

$$Q_{\text{инд}} = C_{\text{окс}} V_{\text{окс}} = \mp Q_{sp} \mp Q_{ss} = \mp 2en_i L_d F(\lambda, Y_s) \mp e \sum_i N_{si} f_{si}(E_{si}), \quad (10-35)$$

где N_{si} — число состояний на уровне захвата с энергией E_{si} ; f_{si} — соответствующая функция распределения (знак минус во 2-м слагаемом относится к акцепторным, плюс — к донорным состояниям). Оставляя на дальнейшее вопросы кинетики захвата носителей заряда на поверхностные состояния (см. § 10-4), здесь лишь отметим, что мы будем предполагать такие условия измерения экспериментальной кривой $G_s(Q_{\text{инд}})$, при которых значения f_{si} однозначно определяются значением поверхностного потенциала φ_s (или уровня Ферми на поверхности), т. е. все уровни поверхностных состояний находятся в равновесии друг с другом и зоной проводимости и валентной зоной (так называемый квазиравновесный эффект поля).

Эффективность управления значениями Y_s и G_s с помощью поперечного электрического поля зависит от соотношения между изменением заряда в приповерхностной области ΔQ_{sp} и захваченного заряда ΔQ_{ss} при заданном ΔV_q : чем больше ΔQ_{ss} , тем меньше ΔY_s и ΔG_s . Поскольку при заданном знаке $\frac{dG_s}{dQ_{\text{инд}}} = \mu_{\text{эп}}$ величина по-

верхностной проводимости определяется только значением поверхностного электростатического потенциала Y_s , то очевидно, что при различной величине захваченного заряда Q_{ss} необходимо индуцировать в полупроводник различный по величине заряд для достижения одного и того же значения G_s (или Y_s). Разность в величинах индуцированного заряда для достижения данного значения G_s (или Y_s) в экспериментальной и теоретической кривой дает величину захваченного заряда для заданного значения поверхностного электростатического потенциала Y_s . Эту разность, как правило, и определяют при исследовании по методу эффекта поля.

Вышесказанное предполагает, что для расчета необходимо построить теоретическую кривую $G_s(Q_{\text{инд}})$, для чего должны быть известны параметры полупроводника — собственная концентрация n_i , уровень легирования N_d (N_a), подвижности электронов μ_n и дырок μ_p с учетом характера поверхностного рассеяния, а также емкость диэлектрика $C_{\text{окс}}$. Построение может быть проведено с помощью соотношений (10-25), (10-26) и (10-27). Обработка данных по эффекту поля возможна в том случае, когда экспериментальная кривая $G_s(Q_{\text{инд}})$ имеет, подобно теоретической, минимум. Это происходит тогда, когда плотность поверхностных состояний, способных менять заряд в условиях измерений, относительно мала, а величина поперечного поля достаточно велика. Значение G_s , соответствующее минимуму как в экспериментальной, так и теоретической кривых $G_s(Q_{\text{инд}})$, определяется условием (10-31) и может быть принято за точку отсчета абсолютных значений Y_s . Значение по-

верхностной проводимости также удобно отсчитывать от его значения в минимуме. Для этого минимум экспериментальной кривой совмещают с минимумом теоретической: этапы расчета ясны из рис. 10-10.

Данные исследования эффекта поля на недостаточно качественно приготовленных образцах МДП-структур на постоянном токе показали, что в процессе измерения после приложения постоянного напряжения на управляющий электрод поверхностная проводимость релаксирует к нулю при всех возможных значениях V_g любого знака. Это свидетельствует о большом количестве акцепторных и донорных состояний на границе раздела полупроводник — диэлектрик, захватывающих индуцированный заряд и экранирующую

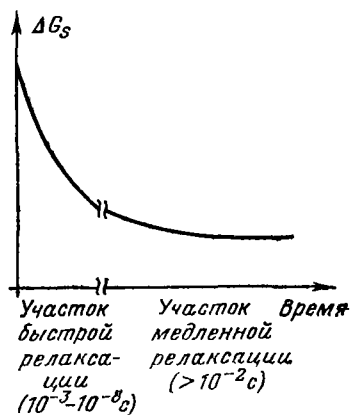


Рис. 10-11. Кривая релаксации поверхностной проводимости.

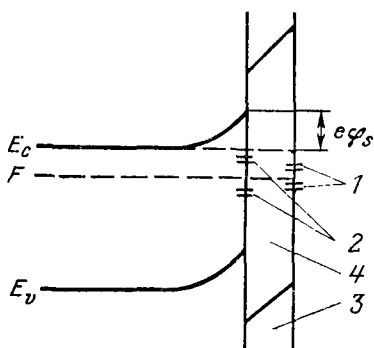


Рис. 10-12. Быстрые и медленные состояния на поверхности германия. 1 — медленные состояния; 2 — быстрые состояния; 3 — слой окиси германия; 4 — запрещенная зона окиси.

щих объем полупроводника от проникновения электрического поля. Более детальное изучение кинетики эффекта поля показало, что существуют два участка спада на кривой поверхностной проводимости, характеризующихся резко различными постоянными спада (рис. 10-11). Это позволило говорить о двух типах поверхностных состояний с различными временами захвата. К 1-му типу состояний относятся так называемые «быстрые состояния», которые характеризуются временами захвата порядка $10^{-8} - 10^{-3}$ с; полагают, что они расположены непосредственно на границе раздела полупроводник — диэлектрик. Ко 2-му типу состояний относятся «медленные состояния», расположенные на внешней стороне окисла и потому характеризующиеся большими временами захвата — от 10^{-3} с до нескольких часов (рис. 10-12). Применение в измерениях синусоидального напряжения определенной частоты на управляющем электроде позволяет исключить влияние «медленных» поверхностных состояний на результаты измерений в эффекте поля.

10.4. ПОВЕРХНОСТНАЯ РЕКОМБИНАЦИЯ

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

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

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

В условиях термодинамического равновесия поток дырок к поверхности полупроводника в соответствии с (9-85) равен $\frac{1}{4} v_0 p_0$, где v_0 — тепловая скорость дырок; p_0 — равновесная концентрация дырок. Часть этих дырок будет захвачена на уровне E_s , а часть дырок, равная $\frac{1}{4} r v_0 p_0$, отразится от поверхности, где r — средняя скорость отражения дырки от поверхности полупроводника. Количество дырок, захваченных поверхностными уровнями в единицу времени, в условиях равновесия равно числу дырок q_{s0} , генерируемых за это же время с поверхностных уровней в валентную зону. Но при термодинамическом равновесии поток дырок, подходящих к поверхности полупроводника, равен потоку дырок, движущихся от поверхности. Обратный поток включает как отра-

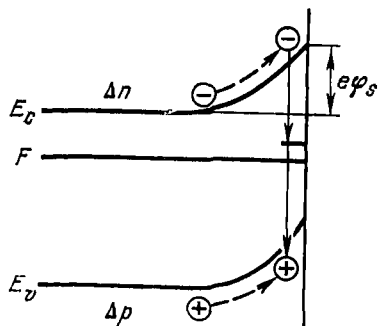


Рис. 10-13. Рекомбинация через поверхностные уровни.

женные от поверхности дырки, так и дырки, освобожденные с поверхностных уровней. Поэтому

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

откуда скорость ухода дырок с поверхностных уровней в состоянии равновесия

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

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

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

или

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

где

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

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

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

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

где γ_{ns} и γ_{ps} — всроятность захвата поверхностным уровнем электронов и дырок из соответствующих зон в течение одной секунды, имеющая размерность $\text{см}^2/\text{с}$; n_s и p_s — неравновесные концентра-

ции электронов и дырок на поверхности полупроводника; n_1 и p_1 — равновесные концентрации электронов и дырок, когда уровень Ферми совпадает с уровнем рекомбинационной ловушки E_s , поэтому $n_1 p_1 = n_0 p_0 = n_i^2$.

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

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

а в условиях термодинамического равновесия

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

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

$$R = \frac{\gamma_{ns}\gamma_{ps}(np - n_0p_0)}{\gamma_{ns}(n_s + n_1) + \gamma_{ps}(p_s + p_1)}. \quad (10-44)$$

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

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

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

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

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

$$Y_{\text{макс}} = \frac{1}{2} \ln \frac{\gamma_{ns} n_0}{\gamma_{ps} p_0}. \quad (10-47)$$

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

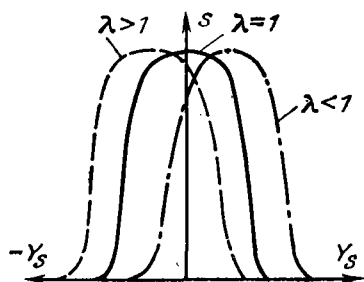


Рис. 10-14. Зависимость скорости поверхностной рекомбинации от безразмерного поверхностного электростатического потенциала.

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

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

10-5. ВЛИЯНИЕ ПОВЕРХНОСТНОЙ РЕКОМБИНАЦИИ НА ВРЕМЯ ЖИЗНИ НОСИТЕЛЕЙ ЗАРЯДА

Наличие поверхностных рекомбинационных ловушек, естественно, приводит к увеличению общего числа рекомбинационных центров в образце и, следовательно, к уменьшению времени жизни носителей заряда.

Рассмотрим распределение избыточной концентрации носителей заряда вблизи поверхности, на которой $s \neq 0$ в случае, когда генерация избыточных электронно-дырочных пар производится равномерно по всему объему образца со скоростью G . Для определенности выбираем полупроводник n -типа, координата x нормальна к поверхности образца и возрастает от поверхности в глубь образца, внешнее электрическое поле отсутствует.

Избыточная концентрация неосновных носителей — дырок $\Delta p(x)$ может быть получена из решения стационарного уравнения непрерывности

$$\frac{\partial p}{\partial t} = -\frac{1}{e} \frac{\partial j_p}{\partial x} - \frac{\Delta p}{\tau_p} + G = 0. \quad (10-48)$$

Поскольку напряженность электрического поля в области пространственного заряда невелика, дрейфовую составляющую тока можно не учитывать. Тогда $j_p = -eD_p \frac{dp}{dx}$ и уравнение (10-48) примет вид:

$$D_p \frac{d^2 \Delta p}{dx^2} - \frac{\Delta p}{\tau_p} + G = 0. \quad (10-49)$$

Граничное условие в точке $x = 0$ для потока дырок к поверхности может быть задано в виде

$$D_p \frac{d \Delta p}{dx} = -\frac{j_p}{e} = \Delta q_s = s \Delta p. \quad (10-50)$$

Поскольку потоки электронов и дырок, обусловленные поверхностной рекомбинацией, носят чисто диффузионный характер, то можно ожидать, что влияние поверхности на распределение избыточной концентрации дырок будет сказываться на расстояниях порядка нескольких диффузионных длин от поверхности. Следовательно, для $x \rightarrow \infty$ уравнение непрерывности (10-49) будет:

$$-\frac{\Delta p}{\tau_p} + G = 0 \quad \text{и} \quad \Delta p \Big|_{x > L_p} = G\tau_p. \quad (10-51)$$

Общее решение уравнения (10-49) имеет вид:

$$\Delta p(x) = A e^{-x/L_p} + G\tau_p. \quad (10-52)$$

Для нахождения A используем граничное условие (10-50) в точке $x = 0$

$$-\frac{A}{L_p} D_p = s(A + G\tau_p). \quad (10-53)$$

Отсюда

$$A = -\frac{Gs\tau_p L_p}{D_p + sL_p} = -\frac{Gs\tau_p^2}{L_p + s\tau_p}. \quad (10-54)$$

Окончательно имеем:

$$\Delta p(x) = G\tau_p \left[1 - \frac{s\tau_p}{L_p + s\tau_p} e^{-x/L_p} \right] = G\tau_p \left[1 - \frac{s/v_D}{1 + s/v_D} e^{-x/L_p} \right], \quad (10-55)$$

где $v_D = \frac{L_p}{\tau_p}$ есть диффузионная скорость.

Физический смысл полученного соотношения достаточно очевиден:

1) если $s \ll v_D$, то для обеспечения малого по величине диффузионного потока на поверхностную рекомбинацию необходимы малые градиенты концентраций избыточных носителей заряда, т. е. Δp мало меняется с изменением x и $\Delta p \approx G\tau_p$; 2) при $s \gg v_D$ будет

иметь место сильная зависимость Δp от x , поскольку необходимо иметь достаточный по величине диффузионный поток на поверхностную рекомбинацию. Характер зависимости $\Delta n = \Delta p$ от x при различных соотношениях s и v_D иллюстрируется рис. 10-15.

Концентрацию избыточных дырок на поверхности полупроводника Δp_s можно определить, если в выражении (10-55) положить $x = 0$. Она будет равна:

$$\Delta p_s = \frac{G\tau_p L_p}{L_p + s\tau_p}. \quad (10-56)$$

Рис. 10-15. Влияние скорости поверхностной рекомбинации на распределение избыточной концентрации носителей заряда.

Введем $\tau_s = \Delta p_s/G$, которое является временем жизни неосновных носителей заряда на поверхности полупроводника. Из (10-56) следует, что

$$\tau_s = \tau_p \frac{L_p}{L_p + s\tau_p}. \quad (10-57)$$

Тогда

$$1 - \frac{\tau_s}{\tau_p} = \frac{s\tau_p}{L_p + s\tau_p}. \quad (10-58)$$

В таком случае равенство (10-55) будет иметь вид:

$$\begin{aligned} \Delta p(x) &= G\tau_p \left[1 - \left(1 - \frac{\tau_s}{\tau_p} \right) e^{-x/L_p} \right] = \\ &= G\tau_p \left[1 + \left(\frac{\tau_s}{\tau_p} - 1 \right) e^{-x/L_p} \right]. \end{aligned} \quad (10-59)$$

Отсюда получим время жизни неосновных носителей заряда в приповерхностной области полупроводника:

$$\tau(x) = \frac{\Delta p(x)}{G} = \tau_p \left[1 + \left(\frac{\tau_s}{\tau_p} - 1 \right) e^{-x/L_p} \right]. \quad (10-60)$$

ПОГЛОЩЕНИЕ СВЕТА ПОЛУПРОВОДНИКАМИ

11-1. СПЕКТР ОТРАЖЕНИЯ И СПЕКТР ПОГЛОЩЕНИЯ

Когда пучок монохроматического излучения проходит через вещество, то вследствие отражения от поверхности и поглощения в объеме 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

$$R = \frac{I_R}{I_0}. \quad (11-1)$$

Зависимость коэффициента отражения от энергии падающего света $R(h\nu)$ или длины волны $R(\lambda)$ называется спектром отражения.

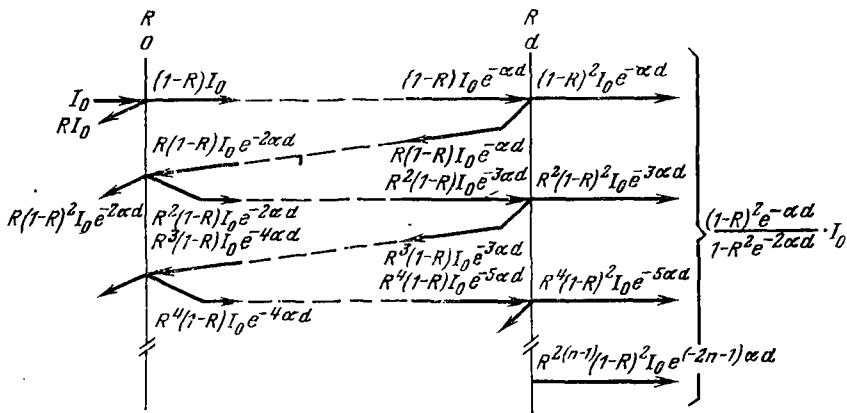


Рис. 11-1. Поток энергии в системе с многократным внутренним отражением.

Обозначим интенсивность света, падающего на слой толщиной d , как это изображено на рис. 11-1, через I_0 . Тогда с учетом отражения через первую поверхность пройдет излучение $(1-R) I_0$. Вследствие поглощения света в слое толщиной dx интенсивность излучения I уменьшится на величину dI . Количество поглощенной энергии dI пропорционально количеству падающего на слой энергии I и толщине поглощающего слоя dx :

$$-dI = \alpha I dx. \quad (11-2)$$

Коэффициент пропорциональности α , выражающий собой количество поглощенной энергии из пучка единичной интенсивности в слое единичной длины, носит название коэффициента поглощения.

Проинтегрировав уравнение (11-2), найдем интенсивность излучения $(1-R) I_0 e^{-\alpha d}$, достигшего второй поверхности. Интенсивность света, вышедшего из образца, будет равна $(1-R)(1-R) I_0 e^{-\alpha d}$. Свет, отраженный внутрь образца, как следует из рис. 11-1, выйдет из него значительно ослабленным. С учетом этих эффектов соотношение для коэффициента пропускания T , выражающего собой отношение интенсивности света $I = \frac{(1-R)^2 e^{-\alpha d}}{1-R^2 e^{-2\alpha d}} I_0$, прошедшего через образец толщиной d , к интенсивности падающего света I_0 , будет:

$$T = \frac{(1-R)^2 e^{-\alpha d}}{1-R^2 e^{-2\alpha d}}. \quad (11-3)$$

Если αd велико, то можно пренебречь вторым членом в знаменателе. В этом случае интенсивность света, прошедшего через образец толщиной d , с учетом отражения будет равна:

$$I = (1-R)^2 I_0 e^{-\alpha d}. \quad (11-4)$$

Здесь α является характеристикой поглощающей среды и зависит от длины волны излучения. Зависимость коэффициента поглощения от энергии падающего света $\alpha(h\nu)$ или длины волны $\alpha(\lambda)$ называется спектром поглощения вещества.

Пусть имеется N поглощающих центров. Обозначим через σ вероятность поглощения одного фотона на одном поглощающем центре в единицу времени. Другими словами, σ есть эффективное сечение поглощения одного фотона в единицу времени. Эффективное сечение σ зависит от энергии фотона и природы поглощающих центров.

Если через l_ϕ обозначить среднюю длину свободного пробега фотона в поглощающей среде, то согласно (5-51) она равна:

$$l_\phi = (\sigma N)^{-1}, \quad (11-5)$$

а коэффициент поглощения

$$\alpha = \frac{1}{l_\phi} = \sigma N \quad (11-6)$$

есть вероятность поглощения фотона на единичной длине.

Допустим, что в полупроводнике имеются поглощающие центры различной природы. Если центры поглощения N_i характеризуются эффективным сечением σ_i , то

$$\alpha_i(h\nu) = \frac{1}{l_{i\phi}} = \sigma_i N_i. \quad (11-7)$$

Полный коэффициент поглощения вещества α будет:

$$\alpha(h\nu) = \sum_i \sigma_i N_i = \sum_i \alpha_i, \quad (11-8)$$

т. е. он равен сумме парциальных коэффициентов поглощения. Следовательно, полный спектр поглощения вещества складывается из спектров поглощения различных поглощающих центров.

При взаимодействии электронов полупроводника с электромагнитным излучением должны выполняться два закона: закон сохранения энергии и закон сохранения квазиимпульса. Если электрон до взаимодействия с квантом света, имеющим энергию $h\nu = \hbar\omega$ и импульс $\hbar\eta$, обладал энергией E и квазиимпульсом \mathbf{p} , а после взаимодействия имел E' и \mathbf{p}' , то эти законы запишутся в виде

$$E' = E + \hbar\omega; \quad (11-9)$$

$$\mathbf{p}' = \mathbf{p} + \hbar\eta. \quad (11-10)$$

Поглощение излучения в полупроводниках может быть связано с изменением энергетического состояния свободных или связанных с атомами электронов, а также с изменением колебательной энергии атомов решетки. В связи с этим в полупроводниках различают пять основных типов оптического поглощения: собственное, экситонное, поглощение свободными носителями заряда, примесное и решеточное поглощение.

11-2. СОБСТВЕННОЕ ПОГЛОЩЕНИЕ ПРИ ПРЯМЫХ ПЕРЕХОДАХ

Если при поглощении полупроводником кванта излучения имеет место возбуждение электронов из валентной зоны в зону проводимости, то такое поглощение 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 $\mathbf{k}_{\text{мин}}$, и максимум энергии в валентной зоне, определяемый волновым вектором $\mathbf{k}_{\text{макс}}$, расположены в одной и той же точке зоны Бриллюэна (обычно в точке $\mathbf{k} = 0$). Другими словами, у этих полупроводников $\mathbf{k}_{\text{мин}} = \mathbf{k}_{\text{макс}}$ (рис. 11-2, а). В качестве примера полупроводника с таким построением зон можно указать антимирид индия (см. рис. 2-25).

У второго вида веществ экстремумы зоны проводимости и валентной зоны находятся при различных \mathbf{k} , так что в этом случае $\mathbf{k}_{\text{мин}} \neq \mathbf{k}_{\text{макс}}$ (рис. 11-2, б). К последнему типу веществ относится большинство полупроводников, в том числе германий и кремний (см. рис. 2-21).

Переходы электронов через запрещенную зону будут происходить прежде всего между энергетическими состояниями, соответствующими максимуму валентной зоны и минимуму зоны проводимости, т. е. при значениях квазиимпульса \mathbf{p} или волнового вектора \mathbf{k} , близких к нулю, как это показано на рис. 11-2, а.

Импульс фотона $\hbar\eta = 1/\lambda_\phi$ очень мал по сравнению с квазиимпульсом электрона, так как длина волны света λ_ϕ составляет $10^{-1}—10^{-5}$ см, а длина волны электрона в кристалле при энергии электрона, соответствующей 300 К, равна примерно $5 \cdot 10^{-7}$ см. В связи с этим в уравнении (11-10) импульсом фотона можно пренебречь, поэтому

$$\mathbf{p}' = \mathbf{p} \quad (11-11)$$

или

$$\mathbf{k}' = \mathbf{k}. \quad (11-12)$$

Соотношение (11-11) или (11-12), называемое правилом отбора для электронных переходов, показывает, что в процессе взаимодействия электрона кристалла с полем излучения возможны только такие переходы, при которых волновой вектор электрона сохраняется. Эти переходы получили название вертикальных, или прямых, переходов (рис. 11-2, а и переход 1 на рис. 11-2, б).

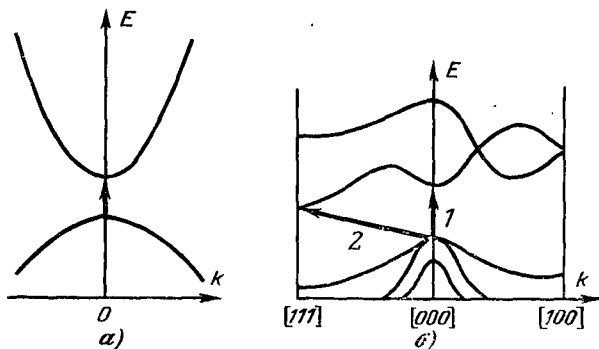


Рис. 11-2. Собственное поглощение при прямых межзонных переходах (а) и оптические переходы в полупроводнике со сложными энергетическими зонами (б).

Рассмотрим собственное поглощение при прямых межзонных переходах для полупроводника, у которого зона проводимости и валентная зона обладают сферической симметрией. Будем полагать, что валентная зона целиком занята электронами, а зона проводимости — полностью свободная. Это условие справедливо для нелегированных полупроводников при 0 К. Согласно рис. 11-3 каждому начальному состоянию E в валентной зоне при поглощении кванта света с энергией $h\nu$ отвечает такое конечное состояние E' в зоне проводимости, что

$$E' = h\nu + E. \quad (11-13)$$

В случае зон сферической симметрии в соответствии с (4-4) и (4-10)

$$E' = E_c + \frac{\hbar^2 k^2}{2m_n^*}; \quad E = E_v - \frac{\hbar^2 k^2}{2m_p^*}, \quad (11-14)$$

поэтому

$$h\nu - E_g = \frac{\hbar^2 k^2}{2m_r^*}, \quad (11-15)$$

где $E_g = E_c - E_v$ — ширина запрещенной зоны, а m_r^* — приведенная масса электрона дырки, определяемая соотношением

$$\frac{1}{m_r^*} = \frac{1}{m_n^*} + \frac{1}{m_p^*}. \quad (11-16)$$

Коэффициент поглощения $\alpha(h\nu)$ согласно (11-6) обратно пропорционален длине свободного пробега фотона l_ϕ , величина которой в свою очередь определяется скоростью движения фотона v_ϕ и временем его свободного пробега τ_ϕ . Следовательно,

$$\alpha(h\nu) = \frac{1}{l_\phi} = \frac{1}{v_\phi \tau_\phi} = g(h\nu) \frac{\bar{n}}{c}. \quad (11-17)$$

Здесь $g(h\nu) = \frac{1}{\tau_\phi}$ — вероятность поглощения фотона в единицу времени; $v_\phi = c/\bar{n}$ — скорость движения фотона в веществе; \bar{n} — показатель преломления вещества; c — скорость света в вакууме.

Вероятность поглощения $g(h\nu)$ кванта света с энергией, лежащей в интервале от $h\nu$ до $h(\nu + d\nu)$, пропорциональна вероятности перехода электрона $P(h\nu)$, числу квантовых состояний в валентной зоне в интервале энергии $|dE|$:

$$g(h\nu) d(h\nu) = 2P(h\nu) N(E) |dE|. \quad (11-18)$$

Здесь коэффициент 2 учитывает два возможных направления поляризации света. Плотность квантовых состояний в валентной зоне согласно выражению (4-11) имеет вид:

$$N(E) = \frac{(2m_p^*)^{3/2}}{2\pi^2 \hbar^3} (E_v - E)^{1/2}. \quad (11-19)$$

Используя соотношения (11-14) и (11-19), получаем:

$$g(h\nu) d(h\nu) = \frac{2}{\pi^2} P(h\nu) k^2 dk. \quad (11-20)$$

Но из выражения (11-15) следует, что

$$d(h\nu) = \frac{\hbar^2}{m_r^*} k dk, \quad (11-21)$$

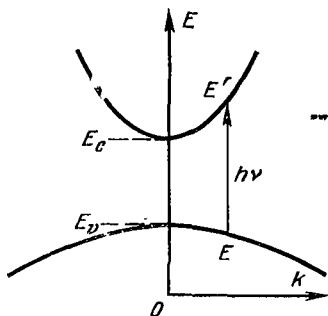


Рис. 11-3. Прямые переходы.

а

$$k = \frac{(2m_r^*)^{1/2}}{\hbar} (h\nu - E_g)^{1/2}, \quad (11-22)$$

поэтому

$$g(h\nu) = \frac{(2m_r^*)^{3/2}}{\pi^2 \hbar^3} P(h\nu) (h\nu - E_g)^{1/2}. \quad (11-23)$$

Для прямых переходов с сохранением квазиимпульса (11-11) вероятность перехода $P(h\nu)$ не зависит от энергии фотона. С учетом этого коэффициент собственного поглощения для прямых раз-

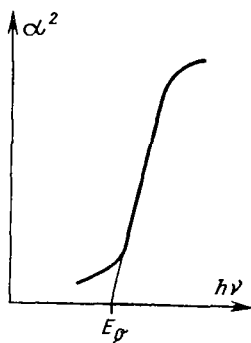


Рис. 11-4. Зависимость квадрата коэффициента поглощения при прямых разрешенных переходах от энергии.

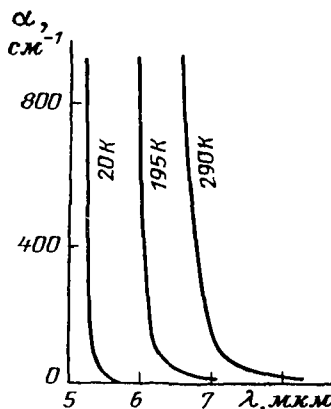


Рис. 11-5. Спектр поглощения InSb при разных температурах.

решенных переходов на основании соотношений (11-17) и (11-23) выражается формулой

$$\alpha = A (h\nu - E_g)^{1/2}, \quad (11-24)$$

где коэффициент A имеет вид:

$$A = \frac{e^2 \left(2 \frac{m_n^* m_p^*}{m_n^* + m_p^*} \right)^{3/2}}{\pi \hbar^2 c m_n^*}. \quad (11-25)$$

Соотношение (11-24) справедливо в ограниченном интервале изменения $(h\nu - E_g)$. Как это видно из рис. 11-4, α^2 линейно зависит от $h\nu$ в некотором интервале. Продолжение этой прямой до пересечения с осью абсцисс позволяет определить ширину запрещенной зоны E_g для прямых разрешенных переходов.

Из формулы (11-24) следует, что для прямых переходов не должно иметь место поглощение квантов с энергией, меньшей ши-

рины запрещенной зоны. В силу этого край собственного поглощения со стороны длинных волн (или малых энергий) должен быть очень резким. Действительно, как следует из рис. 11-5, у очень чистого монокристалла антимонида индия, имеющего прямые долины, наблюдается крутой рост поглощения.

В случае образования валентной зоны из s -состояния, а зоны проводимости на основе атомной d -функции переходы будут прямыми, но «запрещенными» (см. рис. 2-11). При запрещенных пере-

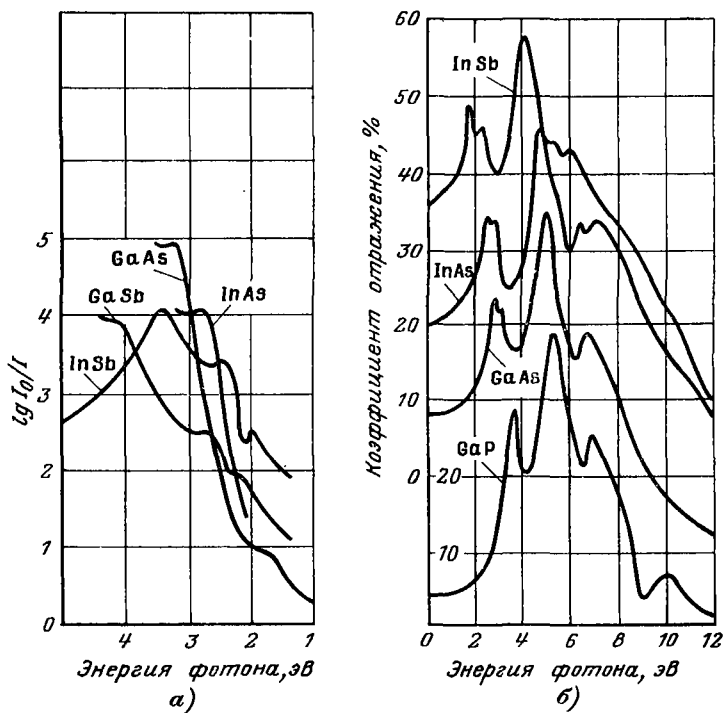


Рис. 11-6. Спектры поглощения при 80 К (а) и спектры отражения при 300 К (б) некоторых полупроводников типа $A^{III}B^V$.

ходах для всех $k \neq 0$ вероятность перехода $P(h\nu)$ пропорциональна k^2 . В силу этого для запрещенных прямых переходов вероятность перехода зависит от энергии как

$$P(h\nu) \approx k^2 = \text{const} (h\nu - E_g). \quad (11-26)$$

Поэтому коэффициент поглощения для прямых запрещенных межзонных переходов определяется соотношением вида

$$\alpha'(h\nu) = A' (h\nu - E_g)^{3/2}. \quad (11-27)$$

Зона проводимости полупроводников, как правило, включает несколько подзон (см. рис. 2-21). Поэтому при поглощении фотона

с энергией, большей ширины запрещенной зоны, прямые переходы из валентной зоны в зону проводимости могут происходить во всех точках пространства квазиимпульса, для которых выполняется правило отбора. В этом случае спектр поглощения будет сложным, отражая сложную структуру зоны проводимости. Для таких исследований используются тонкие пленки полупроводников, поскольку при энергиях фотона $E > E_g$ коэффициент поглощения α очень велик. В качестве примера на рис. 11-6, а представлены спектры поглощения тонких пленок некоторых полупроводников типа $A^{III}B^V$ при 80 К.

Прямые переходы при больших энергиях будут проявляться и в спектрах отражения. В них, как видно из рис. 11-6, б, наблюдается серия пиков, отражающая сложное строение зоны проводимости. Интенсивность и резкость этих пиков обусловлены высокой плотностью начальных и конечных состояний.

11-3. СОБСТВЕННОЕ ПОГЛОЩЕНИЕ ПРИ НЕПРЯМЫХ ПЕРЕХОДАХ

В некоторых случаях существуют факторы, «смягчающие» правила отбора, в результате чего оказываются допустимыми также неперпендикулярные переходы между непрямыми долинами (переход 2 на рис. 11-2, б). Однако вероятность таких переходов намного меньше вероятности прямых переходов (переход 1 на рис. 11-2, б).

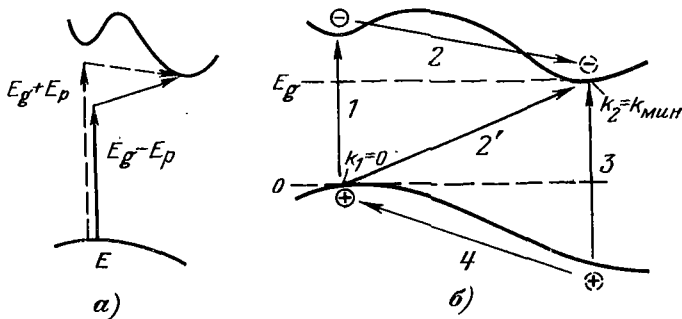


Рис. 11-7. Непрямые переходы в полупроводнике с непрямыми долинами.

В отличие от прямых переходов закон сохранения квазиимпульса при непрямах переходах обеспечивается взаимодействием с фононом, как это показано на рис. 11-7, а. Следовательно, неперпендикулярные, или неперпендикулярные, переходы осуществляются с испусканием или поглощением фононов. Эти переходы определяют поглощение, расположенное с длинноволновой стороны у границы собственного поглощения, обусловленного прямыми переходами. Прямые и неперпендикулярные переходы достаточно четко проявляются в спектре поглощения германия и кремния. Согласно рис. 11-8 пороговые значения энер-

гии фотонов в германии, соответствующие непрямым и прямым оптическим переходам, при комнатной температуре равны 0,64 и 0,81 эВ.

У полупроводника с непрямыми долинами в случае непрямых переходов возможны переходы из любого занятого состояния валентной зоны в любое свободное состояние зоны проводимости. При расчете коэффициента поглощения для непрямых переходов между непрямыми долинами, которые сопровождаются поглощением или испусканием одного фонона, можно поступить следующим образом. Непрямой переход будем рассматривать как переход, происходящий через ряд промежуточных «виртуальных» состояний, имеющих очень малое время жизни. В этом случае выполнение закона сохранения энергии имеет место только для всего перехода в целом, в то время как квазиимпульс электрона сохраняется при переходе его в виртуальное состояние.

Для простоты положим, что непрямой переход между непрямыми долинами происходит из начального состояния $k_1 = 0$ в конечное с $k_2 = k_{\text{мин}}$ (переход 2' на рис. 11-7, б). Для такого перехода, идущего с испусканием или поглощением фонона, возможны два пути. Электрон, находящийся вблизи потолка валентной зоны с $k_1 = 0$, возбуждаясь светом, переходит в зону проводимости, занимая там состояние с тем же волновым вектором (переход 1 на рис. 11-7, б). В валентной зоне в результате такого перехода остается дырка с $k_1 = 0$. Однако электрон проводимости с $k_1 = 0$ обладает большей энергией, чем та, которая соответствует дну зоны проводимости, поэтому он через очень малый промежуток времени совершит переход в состояние вблизи минимума энергии $k_2 = k_{\text{мин}}$, поглощая или испуская при этом фонон (переход 2 на рис. 11-7, б).

Кроме того, возможен и такой переход. Поглощая квант света, электрон из глубоко лежащего состояния в валентной зоне вертикально переходит в состояние зоны проводимости с $k_2 = k_{\text{мин}}$ (переход 3 на рис. 11-7, б). В результате в валентной зоне остается глубоко лежащая дырка, которая совершит переход в состояние $k_1 = 0$ вблизи потолка валентной зоны, испуская или поглощая при этом фонон (переход 4 на рис. 11-7, б).

Произведем расчет коэффициента поглощения для переходов электрона по первому варианту, т. е. когда вначале электрон со-

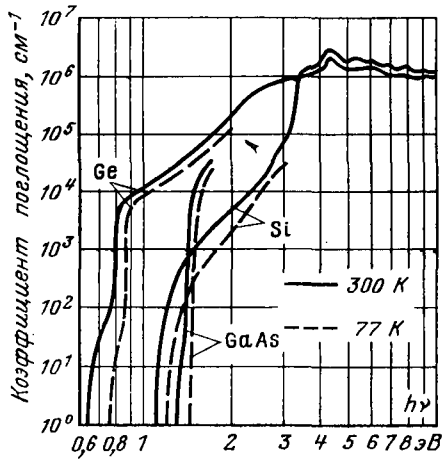


Рис. 11-8. Спектр поглощения чистых Ge, Si и GaAs.

вершает вертикальный переход 1, а затем он осуществляет переход 2. Пусть начальная энергия электрона в валентной зоне E , а энергия электронов в конечном состоянии в зоне проводимости E' . Энергию фонона, принимающего участие в оптических переходах, обозначим через E_p . Если начало отсчета энергии производится от дна валентной зоны, то энергия кванта света, необходимая электрону для совершения непрямого перехода (рис. 11-7, б) в случае поглощения фонона, будет определяться равенством

$$h\nu_\alpha = E' - E - E_p, \quad (11-28)$$

а в случае испускания фонона

$$h\nu_e = E' - E + E_p. \quad (11-29)$$

С учетом принятого отсчета энергии (рис. 11-7) плотность начальных состояний при энергии E имеет вид:

$$N(E) = \frac{(2m_p^*)^{3/2}}{2\pi^2\hbar^3} |E|^{1/2}, \quad (11-30)$$

а плотность состояний при энергии E' (в зоне проводимости)

$$N_c(E') = \frac{(2m_n^*)^{3/2}}{2\pi^2\hbar^3} (E' - E_g)^{1/2}. \quad (11-31)$$

Подставляя в это выражение E' из соотношений (11-28) и (11-29), получаем:

$$N(E') = \frac{(2m_n^*)^{3/2}}{2\pi^2\hbar^3} (h\nu - E_g \pm E_p + E)^{1/2}. \quad (11-32)$$

Коэффициент поглощения α пропорционален интегралу по всем возможным парам состояний, разделенных энергией $h\nu \pm E_p$, от произведения плотностей начальных $N(E)$ и конечных состояний $N(E')$, а также вероятности взаимодействия с фононами P , которая есть функция числа фононов с энергией E_p . Поскольку неравновесная функция распределения фононов N_p мало отличается от равновесной $\langle n \rangle$, то согласно (3-75) можно считать, что число фононов определяется формулой

$$N_p = \frac{1}{e^{E_p/kT} - 1}. \quad (11-33)$$

Следовательно,

$$\begin{aligned} \alpha(h\nu) &= AN_p \int_0^{-(h\nu - E_g \pm E_p)} |E|^{1/2} (h\nu - E_g \pm E_p + E)^{1/2} dE = \\ &= AN_p (h\nu - E_g \pm E_p)^2. \end{aligned} \quad (11-34)$$

Если подставить (11-33) в (11-34), то коэффициент поглощения для переходов с поглощением фонона приобретет вид:

$$\alpha_a(h\nu) = \frac{A(h\nu - E_g + E_p)^2}{e^{E_p/kT} - 1} \quad (11-35)$$

для $h\nu > E_g - E_p$.

Вероятность испускания фотона пропорциональна $N_p + 1$, поэтому коэффициент поглощения с испусканием фотона равен:

$$\alpha_e(h\nu) = \frac{A(h\nu - E_g - E_p)^2}{1 - e^{-E_p/kT}} \quad (11-36)$$

для $h\nu > E_g + E_p$.

Поскольку при $h\nu > E_g + E_p$ возможно поглощение света как с поглощением, так и с испусканием фононов, то коэффициент поглощения при неярких межзонных переходах будет иметь вид суммы:

$$\begin{aligned} \alpha(h\nu) &= \alpha_a(h\nu) + \alpha_e(h\nu) = \\ &= A \left[\frac{(h\nu - E_g + E_p)^2}{e^{E_p/kT} - 1} + \frac{(h\nu - E_g - E_p)^2}{1 - e^{-E_p/kT}} \right] \end{aligned} \quad (11-37)$$

для $h\nu > E_g + E_p$.

Соотношение (11-37) можно представить в виде

$$\alpha = A \frac{(h\nu - E_g + E_p)^2}{e^{E_p/kT} - 1} \quad \text{для } E_g - E_p < h\nu \leq E_g + E_p; \quad (11-38)$$

$$\alpha = 0 \quad \text{для } h\nu \leq E_g - E_p.$$

График зависимости $\alpha^{1/2}$ от $h\nu$ для германия и кремния, как следует из рис. 11-9, имеет два прямолинейных участка. Прямая 1

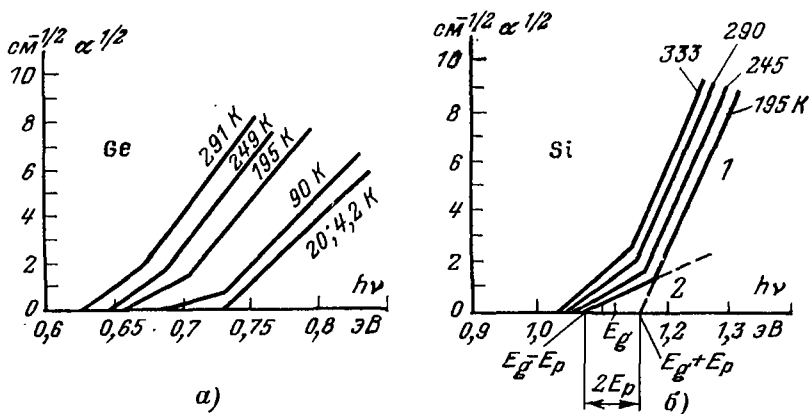


Рис. 11-9. Зависимость $\alpha^{1/2}$ от температуры для германия (а) и кремния (б).

соответствует зависимости $\alpha_e^{1/2}$, которая пересекается с осью абсцисс в точке $E_g + E_p$. Прямая 2 характеризует зависимость $\alpha_a^{1/2}$ и пересекается с осью абсцисс в точке $E_g - E_p$. Длина отрезка между точками пересечения прямых $\alpha_a^{1/2}$ и $\alpha_e^{1/2}$ с осью $h\nu$, соответ-

вующих одной и той же температуре, равна $2E_g$. Посередине этого отрезка находится точка $h\nu = E_g$. При этом, как следует из рис. 11-9, при понижении температуры исследования график зависимости $\alpha^{1/2}$ смещается в сторону больших значений $h\nu$, что отражает температурную зависимость ширины запрещенной зоны, а наклон прямой $\alpha_a^{1/2}$ стремится к нулю. Это происходит в силу того, что при низких температурах возбуждается очень мало фононов, а поэтому невозможны оптические переходы с поглощением фононов.

Если прямой переход при $k_1 = 0$ запрещен, то не прямые запрещенные межзонные переходы для этого случая будут характеризоваться коэффициентом поглощения, который согласно (11-26) будет пропорционален

$$\alpha' \sim \frac{(h\nu - E_g + E_p)^3}{e^{E_p/kT} - 1}, \quad (E_g - E_p < h\nu \leq E_g + E_p). \quad (11-39)$$

Для прямых разрешенных межзонных переходов коэффициент поглощения обычно достигает $10^4 - 10^5 \text{ см}^{-1}$, а для не прямых переходов между непрямыми долинами он составляет $10^{-1} - 10^3 \text{ см}^{-1}$ (см. рис. 11-8).

При рассмотрении не прямых межзонных переходов учитывался только один тип фононов. В действительности в кристалле может существовать много различных фононов. В оптических переходах принимают участие только те фононы, которые обеспечивают необходимое изменение квазимпульса электрона. Такими фононами обычно оказываются продольные и поперечные акустические фононы. В табл. 11-1 приведены для германия и кремния значения энергии фононов E_p , а также дана эквивалентная им температура $\theta = E_p/k$ (температура Дебая).

Таблица 11-1

Энергии фононов и дебаевские температуры для различных ветвей колебаний в германии и кремнии

Полупроводник	Тип колебаний	Величина энергии фононов $E_p = k\theta$, эВ	Эквивалентная температура θ , К
Германий	Продольные акустические LA	0,0275	320
	Поперечные акустические TA	0,0077	90
	Поперечные оптические TO	0,034	400
Кремний	Продольные акустические LA	0,016	230
	Поперечные акустические TA	0,055	700
	Поперечные оптические TO	0,119	1370
	Продольные оптические LO	0,083	1010

11-4. СОБСТВЕННОЕ ПОГЛОЩЕНИЕ СИЛЬНО ЛЕГИРОВАННОГО ПОЛУПРОВОДНИКА

У антимонида индия, имеющего прямую запрещенную зону, по мере роста концентрации свободных электронов край полосы собственного поглощения смещается в сторону коротких длин волн. Например, у материала n -типа при концентрации электронов $5 \cdot 10^{18} \text{ см}^{-3}$ край поглощения приходится на 3,5 мкм, а у нелегированного полупроводника длинноволновая граница поглощения при комнатной температуре соответствует 7,1 мкм (рис. 11-10). Это обусловлено тем, что у антимонида индия из-за малой величины

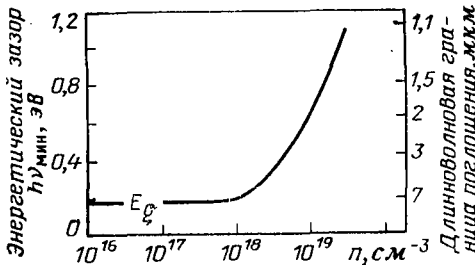


Рис. 11-10. Зависимость оптического энергетического зазора $h\nu_{\text{мин}}$ от концентрации электронов в InSb n -типа.

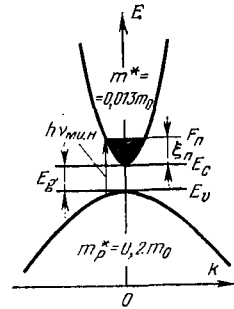


Рис. 11-11. Зависимость энергии от квазиимпульса в InSb .

эффективной массы электронов ($m_n^* = 0,013 m_0$) зона проводимости характеризуется большой крутизной ($\frac{1}{m^*} = \frac{1}{\hbar} \frac{\partial^2 E}{\partial k^2}$), поэтому плотность состояний в ней мала (рис. 11-11). Вследствие этого при увеличении концентрации электронов быстро наступает вырождение электронного газа и уровень Ферми F_n поднимается выше дна зоны проводимости на величину $\xi_n = F_n - E_c$. Поскольку состояния, расположенные ниже уровня Ферми, уже заполнены, поглощение, связанное с переходами в эти состояния, уже невозможно. В этом случае величина $h\nu_{\text{мин}}$ (прямой переход на рис. 11-11) будет больше ширины запрещенной зоны, а поэтому край поглощения должен быть смещен в сторону больших энергий на величину, примерно равную ξ_n .

В сильно легированном полупроводнике с непрямыми долинами также возможно смещение края поглощения, связанное с заполнением зон (рис. 11-12). В этом случае сохранение квазиимпульса электрона при непрямых переходах обеспечивается за счет электрон-электронного взаимодействия или за счет рассеяния на примеси. При этом вероятность рассеяния пропорциональна числу рассеивающих центров N и поскольку участия фононов не требуется, то коэффициент поглощения описывается формулой

$$\alpha = AN (h\nu - E_g - \xi_n)^2, \quad (11-40)$$

где $\xi_n = F_n - E_c$ — положение уровня Ферми в зоне проводимости сильно легированного полупроводника.

У полупроводников с прямыми долинами возможны и не прямые переходы (рис. 11-13). Такие переходы, будучи процессами второго порядка, обладают меньшей вероятностью, чем прямые переходы, поэтому они обуславливают незначительное поглощение, вызывая удлинение края собственного поглощения (рис. 11-5).

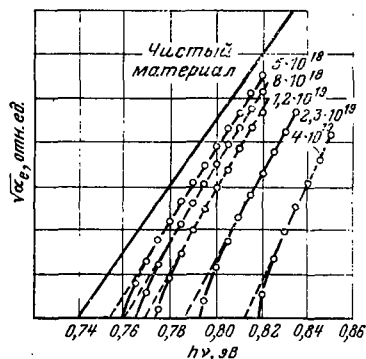


Рис. 11-12. Край поглощения германия, сильно легированного мышьяком, при 4,2 К.

проводимости. Поэтому коэффициент поглощения в этом случае определяется формулами (11-35) — (11-38), если закон сохранения квазиимпульса обеспечивается участием фононов, или формулой (11-40), если в акте поглощения фононы не принимают участия.

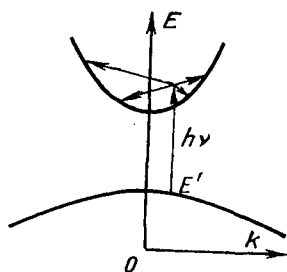


Рис. 11-13. Четыре из возможных не прямых переходов из начального состояния E' в зону проводимости полупроводника с прямой запрещенной зоной.

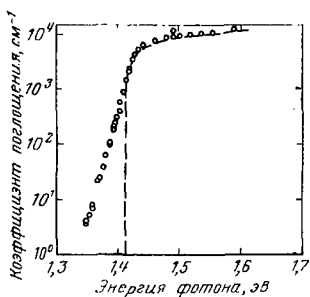


Рис. 11-14. Край поглощения GaAs при комнатной температуре.

Сплошные кружки — экспериментальные данные; пунктир — теоретическая кривая $\alpha = A(h\nu - E_g)^{1/2}$.

Кроме того, у некоторых сильно легированных полупроводников с прямыми долинами наблюдается поглощение вблизи края, которое возрастает по экспоненциальному закону. Как известно, увеличение концентрации примесей сопровождается образованием

вблизи разрешенных зон хвостов плотности состояний в пределах запрещенной зоны (см. § 4-9). Переходы между хвостами зон, форма и величина которых зависит от легирования материала, определяют, в частности, экспоненциальный край поглощения в арсениде галлия (рис. 11-14).

11-5. ВЛИЯНИЕ ВНЕШНИХ ВОЗДЕЙСТВИЙ НА СОБСТВЕННОЕ ПОГЛОЩЕНИЕ ПОЛУПРОВОДНИКОВ

Многочисленные опыты, проведенные на разных полупроводниках, показывают, что внешние воздействия, такие как температура, давление, электрические и магнитные поля, оказывают существенное влияние на их спектр собственного поглощения. Значение влияния их на энергетические уровни полупроводника необходимо как для решения теоретических задач, связанных с рассмотрением поведения электронов в кристалле, так и для решения задач практического значения, связанных с изготовлением и эксплуатацией полупроводниковых приборов.

1. Влияние температуры

У полупроводника с повышением температуры происходят расширение решетки и усиление колебаний атомов относительно положения равновесия. У некоторых полупроводников согласно рис. 2-7 увеличение постоянной решетки ведет к уменьшению ширины запрещенной зоны, а это вызовет смещение полосы фундаментального поглощения в сторону длинных волн (см. рис. 11-5). Для многих полупроводников в области температур, много меньших температуры Дебая, температурная зависимость ширины

Таблица 11-2

Значение E_g и параметров α , β

Полупроводник	Тип межзонных переходов ¹	$E_g(0)$, эВ	$\alpha \cdot 10^4$	β
Si	E_{gi}	1,1557	7,021	1108
Ge	E_{gi}	0,7412	4,561	210
Ge	E_{gd}	0,8893	6,842	398
6H·SiC	E_{gt}	3,024	-0,3055	-311
GaAs	E_{gd}	1,5216	8,871	572
InP	E_{gd}	1,4206	4,906	327
InAs	E_{gd}	0,426	3,158	93

¹ Здесь E_{gi} — не прямые переходы; E_{gd} — прямые переходы.

запрещенной зоны описывается эмпирическим соотношением вида *

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta}, \quad (11-41)$$

где $E_g(0)$ — ширина запрещенной зоны при 0 К; α и β — константы, значения которых для некоторых полупроводников приведены в табл. 11-2. При температурах, много больших температуры Дебая, зависимость ширины запрещенной зоны становится линейной:

$$E_g(t) = E_g(0) - \alpha T. \quad (11-42)$$

2. Влияние давления

При всестороннем сжатии кристалла, обеспечиваемом так называемым гидростатическим давлением, расстояние между его атомами уменьшается. Но уменьшение межатомного расстояния согласно рис. 2-7 вызовет увеличение ширины запрещенной зоны,

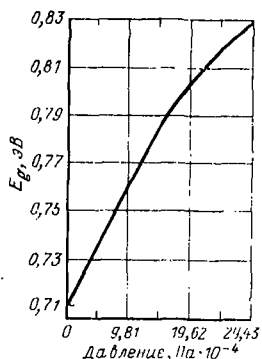


Рис. 11-15. Зависимость ширины запрещенной зоны германия при 350 К от давления.

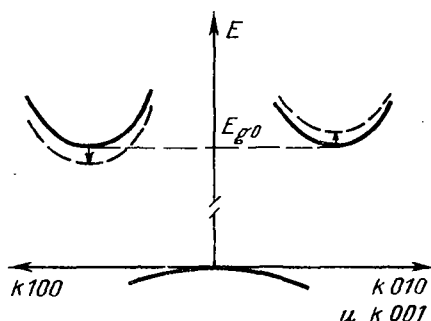


Рис. 11-16. Продольные и поперечные смещения долин зоны проводимости в кремнии под действием одноосного сжатия вдоль направления [100].

поэтому у большинства полупроводников с ростом гидростатического давления имеет место смещение края собственного поглощения в сторону коротких длин волн. У германия, как видно из рис. 11-15, ширина запрещенной зоны линейно меняется с изменением давления

$$E_g(P) = E_{g|P=0} + \gamma P \quad (11-43)$$

вплоть до давлений $14,7 \cdot 10^8 \text{ Па} = 1,5 \cdot 10^4 \text{ кг/см}^2$ с $\gamma = 4,7 \cdot 10^{-11} \text{ эВ/Па} = 4,6 \cdot 10^{-6} \text{ эВ} \cdot \text{см}^2/\text{кг}$, а в области более высоких давлений наблюдается нелинейная зависимость.

* Панков Ж. Оптические процессы в полупроводниках. М., «Мир», 1973, с. 36.

В кремнии увеличение ширины запрещенной зоны составляет $4,28 \cdot 10^{-11}$ эВ/Па = $4,2 \cdot 10^{-6}$ эВ · см²/кг. У арсенида галлия с ростом гидростатического давления минимум зоны проводимости для прямых переходов движется вверх быстрее, чем минимум для непрямых переходов. Поэтому при повышении давления он может стать полупроводником с непрямыми переходами.

Одностороннее сжатие или растяжение кристалла приводит к сдвигу положения различных минимумов энергии в зоне проводимости. Например, сжатие кристалла кремния вдоль оси [100] уменьшает ширину запрещенной зоны в направлении сжатия, но из-за одновременного растяжения в поперечном направлении минимумы зоны проводимости, находящиеся в плоскости, перпендикулярной направлению сжатия, смещаются в область больших энергий (рис. 11-16). В случае германия *n*-типа наибольший эффект наблюдается при сжатии в направлении [110]. Изменение энергетической структуры кристалла вызовет перераспределение электронов, которые переходят в состояния с меньшей энергией, возникшие в результате деформации сдвига. Это скажется на величине упругих констант, т. е. на коэффициенте пьезосопротивления.

3. Влияние сильного электрического поля

Приложение электрического поля вызывает наклон энергетических зон (см. § 2-10). В этом случае электрон валентной зоны может туннелировать через треугольный барьер (рис. 11-17, а).

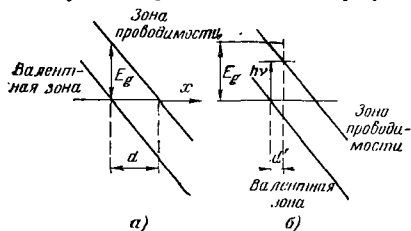
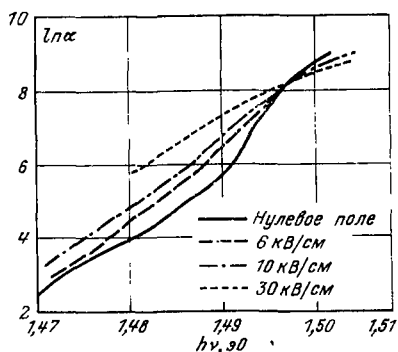


Рис. 11-17. Туннелирование электрона.

а — без изменения энергии; б — с поглощением фотона.

Рис. 11-18. Край поглощения GaAs → при разной напряженности электрического поля.



Высота этого барьера E_g , а его толщина d характеризуется величиной, равной согласно выражению (6-217):

$$d = \frac{E_g}{e\mathcal{E}}, \quad (11-44)$$

где \mathcal{E} — величина напряженности электрического поля. Как видно, с увеличением электрического поля толщина барьера уменьшается, а следовательно, увеличивается вероятность туннелирования.

В присутствии электрического поля участие фотона с энергией $h\nu$, как следует из рис. 11-17, б, эквивалентно уменьшению толщины барьера до величины

$$d' = \frac{E_g - h\nu}{e\mathcal{E}} \quad (11-45)$$

и туннельный переход становится более вероятным. Уменьшение толщины барьера равносильно уменьшению ширины запрещенной зоны в сильном электрическом поле. *Эффект туннелирования в присутствии электрического поля, сопровождаемый поглощением фотона, называется эффектом Франца—Келдыша.* В собственном полупроводнике он проявляется как сдвиг края поглощения в сторону меньших энергий (рис. 11-18).

4. Влияние магнитного поля

Рассмотрим вначале влияние магнитного поля на энергетический спектр кристалла. Эта задача может быть решена на основе уравнения Шредингера, в котором гамильтониан должен определять поведение электрона в магнитном поле.

Из теории электромагнитного поля известно, что вектор индукции магнитного поля \mathbf{B} связан с векторным потенциалом $\mathbf{A}(x, y, z)$ соотношением

$$\mathbf{B} = \text{rot } \mathbf{A}. \quad (11-46)$$

При этом электрон с зарядом $-e$, имеющий импульс \mathbf{p} , при прохождении в области действия поля, определенного векторным потенциалом $\mathbf{A}(x, y, z)$, будет обладать обобщенным импульсом

$$\mathbf{P} = \mathbf{p} + e\mathbf{A}. \quad (11-47)$$

В этом случае, поскольку кинетическая энергия $T = \frac{P^2}{2m_0}$, функция Гамильтона для электрона запишется в виде

$$H = T + U(x, y, z) = \frac{1}{2m_0} (\mathbf{p} + e\mathbf{A})^2 + U(x, y, z), \quad (11-48)$$

где $U(x, y, z)$ — потенциальная энергия.

Положим для определенности, что однородное магнитное поле направлено по оси z , т. е.

$$B_x = B_y = 0; \quad B_z = B. \quad (11-49)$$

Так как

$$\text{rot } \mathbf{A} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} = \mathbf{i} \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \mathbf{j} \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \mathbf{k} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right),$$

то на основании (11-49) имеем:

$$\left. \begin{aligned} B_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = 0; \\ B_y &= \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} = 0; \\ B_z &= \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} = B. \end{aligned} \right\} \quad (11-50)$$

Уравнения (11-50) удовлетворяются, если положить, например,

$$A_x = A_z = 0; \quad A_y = Bx. \quad (11-51)$$

С учетом (11-51) функция Гамильтона (11-48) определится соотношением

$$H = \frac{1}{2m_0} [p_x^2 + (p_y^2 + eA_y)^2 + p_z^2] + U(x, y, z). \quad (11-52)$$

Воспользовавшись приближением эффективной массы (см. § 2-14), запишем уравнение Шредингера $\hat{H}\Psi = E\Psi$, переходя от функции Гамильтона к оператору Гамильтона заменой в (11-52) p_x на $\frac{\hbar}{i} \frac{\partial}{\partial x}$, p_y на $\frac{\hbar}{i} \frac{\partial}{\partial y}$ и p_z на $\frac{\hbar}{i} \frac{\partial}{\partial z}$. В результате получим:

$$-\frac{\hbar^2}{2m^*} \left[\frac{\partial^2 \Psi}{\partial x^2} + \left(\frac{\partial}{\partial y} + \frac{ieBx}{\hbar} \right)^2 \Psi + \frac{\partial^2 \Psi}{\partial z^2} \right] = E\Psi. \quad (11-53)$$

В уравнении (11-53) z и y в явном виде не содержатся. К тому же в отсутствие магнитного поля уравнение (11-53) переходит в известное уравнение для волн де Бройля $-\frac{\hbar^2}{2m^*} \Delta \Psi_0 = E\Psi_0$, которое имеет решения:

$$\Psi_0 = Ce^{i(k_x x + k_y y + k_z z)} \quad \text{и} \quad E_0 = \frac{\hbar^2}{2m^*} (k_x^2 + k_y^2 + k_z^2). \quad (11-54)$$

Отсюда решение уравнения (11-53) можно искать в виде суперпозиции

$$\Psi(x, y, z) = \varphi(x) e^{i(k_y y + k_z z)}, \quad (11-55)$$

где $\varphi(x)$ — некоторая функция, зависящая только от x .

Подставив (11-55) в уравнение Шредингера (11-53) и сократив на экспоненту, получим:

$$-\frac{\hbar^2}{2m^*} \frac{d^2 \varphi}{dx^2} + \frac{\hbar^2 k_y^2}{2m^*} \varphi + \frac{\hbar e k_y B x}{m^*} \varphi + \frac{e^2 B^2 x^2}{2m^*} \varphi + \frac{\hbar^2 k_z^2}{2m^*} \varphi = E\varphi. \quad (11-56)$$

Введем обозначения

$$\omega_0 = \frac{eB}{m^*}; \quad x_0 = -\frac{\hbar}{eB} k_y; \quad E_B = E - \frac{\hbar^2 k_z^2}{2m^*} \quad (11-57)$$

и перепишем в них уравнение (11-56). В результате получим:

$$-\frac{\hbar^2}{2m^*} \frac{d^2\varphi}{dx^2} + \frac{1}{2} m^* \omega_0^2 (x - x_0)^2 \varphi = E_B \varphi. \quad (11-58)$$

Уравнение (11-58) представляет собой уравнение гармонического осциллятора. Таким образом, движение электрона по направлению x описывается уравнением линейного гармонического осциллятора с массой m^* и собственной частотой ω_0 , совершающего колебания около положения равновесия x_0 . Но энергия гармонического осциллятора

$$E_B = \hbar\omega_0 \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2 \dots \quad (11-59)$$

и, следовательно, энергия электрона в магнитном поле равна:

$$E = E_c + \frac{\hbar^2 k_z^2}{2m^*} + E_B = E_c + \frac{\hbar^2 k_z^2}{2m^*} + \hbar\omega_0 \left(n + \frac{1}{2} \right). \quad (11-60)$$

Отсюда видно, что если вектор индукции магнитного поля направлен по оси z , то оно не оказывает никакого влияния на движение электрона в этом направлении и энергия, связанная с этим движением, не квантована, т. е. электрон вдоль магнитного поля движется со скоростью $v_z = \frac{\hbar(k_z - k_{0z})}{m_n^*}$. В плоскости $хоу$, перпенди-

кулярной полю \mathbf{B} , электрон совершает движение по круговой орбите с циклотронной частотой $\omega_0 = \omega_c = \frac{eB}{m^*}$ (см. § 2-13), причем для перехода его с одной круговой орбиты на другую требуется энергия $\hbar\omega_0$. Таким образом, энергия движения электрона в плоскости $хоу$ под действием магнитного поля становится квантованной и спектр поглощенной энергии, связанной с переходами электрона между разрешенными орбитами, будет линейчатый.

Из соотношения (11-60) видно, что спектр энергии $E(k_z)$ электронов кристалла в каждой зоне представляет собой набор квадратичных парабол, сдвинутых по оси энергии относительно друг друга на величину $\hbar\omega_0$ (рис. 11-19, а). Для $k_z = 0$ спектр энергии изображается совокупностью дискретных уровней (рис. 11-19, б), которые возникают в результате группировки квазинепрерывного спектра (рис. 11-19, в). Уровни энергии электрона кристалла в магнитном поле называются уровнями Ландау.

Магнитное поле приводит не только к квантованию энергии электронов кристалла, но и сдвигает дно зоны проводимости и дно валентной зоны. Действительно, согласно (11-60) минимальная кинетическая энергия электрона в зоне проводимости ($n = 0$) не равна нулю, а составляет $\frac{1}{2} \hbar\omega_0$, т. е. в магнитном поле дно зоны проводимости смещается вверх на величину

$$\frac{\hbar\omega_0}{2} = \frac{\hbar e B}{2m^*}. \quad (11-61)$$

Соответственно минимальная энергия дырок с эффективной массой m_p^* в магнитном поле возрастает на $\hbar eV/2m_p^*$, что выражается в смещении вниз на эту величину потолка валентной зоны. Таким образом, под действием магнитного поля ширина запрещенной зоны полупроводника возрастает на величину ΔE_g , равную:

$$\Delta E_g = \frac{e\hbar B}{2} \left(\frac{1}{m_n^*} + \frac{1}{m_p^*} \right) = \frac{e\hbar B}{2m_r^*}, \quad (11-62)$$

где $\frac{1}{m_r^*} = \frac{1}{m_n^*} + \frac{1}{m_p^*}$ — обратная величина приведенной эффективной массы.

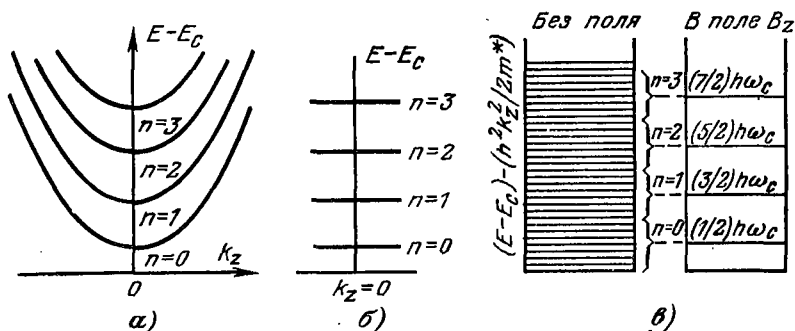


Рис. 11-19. Спектр энергии электронов кристалла в магнитном поле.

Из равенства (11-61) следует, что величина сдвига края соответствующей зоны в магнитном поле обратно пропорциональна эффективной массе носителей заряда в ней. Для антимонида индия с $m_n^* = 0,013 m_0$ сдвиг края зоны проводимости составляет $3,85 \cdot 10^{-7}$ эВ/эрст.

Увеличение ширины запрещенной зоны полупроводника, помещенного в магнитное поле, выразится в сдвиге края полосы фундаментального поглощения в коротковолновую сторону, причем величина сдвига пропорциональна напряженности магнитного поля. Экспериментально, измеряя его величину, на основании (11-62) легко рассчитать приведенную эффективную массу носителей заряда.

Наличие уровней Ландау при помещении полупроводника в магнитное поле создает возможность для переходов электронов между этими уровнями при поглощении квантов света с энергией, кратной $\hbar\omega_0$. Это должно привести к появлению линейчатой структуры на спектральной кривой поглощения. Однако поскольку энергия электрона при движении по направлению вектора напряженности магнитного поля не квантована, то наблюдается некоторое размазывание линейчатого спектра в осциллирующий с последовательностью узких полос с максимумами и минимумами (рис. 11-20, а).

Этот эффект получил название магнитно-осцилляционного эффекта поглощения, осцилляторного магнитно-абсорбционного эффекта или магнетопоглощения.

Магнетопоглощение легче наблюдается в случае прямых переходов, так как тонкая осцилляторная структура при непрямых переходах может оказаться затухевающей. Кроме того, для экспериментального наблюдения магнетопоглощения должно выполняться очевидное условие

$$\hbar\omega_0 \gg kT, \quad (11-63)$$

т. е. расстояние между уровнями энергии в магнитном поле должно быть больше тепловой энергии электронов, иначе квантованный спектр будет для электронов квазинепрерывным.

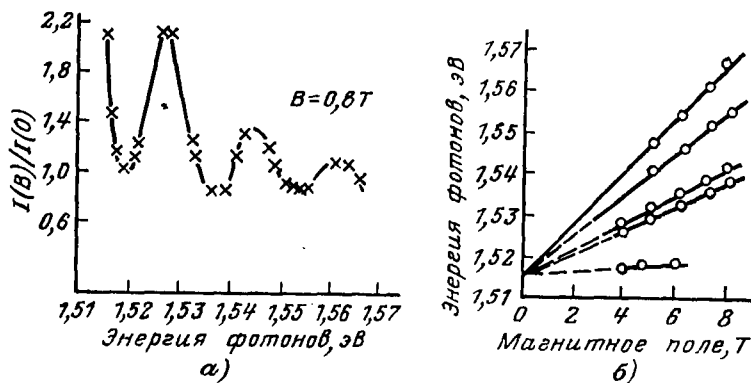


Рис. 11-20. Магнитно-осцилляционный эффект поглощения в арсениде галлия при 77 К.

Из рис. 11-20, а видно, что расстояния между соседними максимумами или минимумами на спектральной зависимости магнетопоглощения составляют величину $\hbar\omega_0$. Поэтому если экспериментально снять спектральную зависимость магнетопоглощения в различных магнитных полях и построить зависимость положения максимумов осциллирующего спектра от величины магнитного поля, то эти зависимости будут представлять собой прямые, сходящиеся к одной точке в нулевом магнитном поле (рис. 11-20, б). Согласно (11-62) изменение ширины запрещенной зоны полупроводника в этой точке равно нулю и ее положение соответствует минимальной энергии, необходимой для осуществления прямого перехода между зонами, т. е. оптической ширины запрещенной зоны. Измерения, проведенные таким методом, дали ширину запрещенной зоны при 300 К в точке $k = 0$ 0,803 эВ для германия, 1,12 эВ для кремния и 1,43 эВ для арсенида галлия. Эти значения были получены для материалов достаточной степени чистоты.

Необходимо отметить, что расщеплением разрешенных энергетических зон на уровни Ландау объясняется как эффект магнетопоглощения, так и эффект циклотронного резонанса. Однако поскольку эффект магнетопоглощения обусловлен переходами между уровнями Ландау, образовавшимися в зоне проводимости и в валентной зоне, то спектр магнетопоглощения лежит внутри полосы фундаментального поглощения. Осциллирующий спектр циклотронного резонанса, обусловленный переходами между уровнями Ландау одной и той же зоны, лежит в спектральной области поглощения на свободных носителях.

11-6. ЭКСИТОННОЕ ПОГЛОЩЕНИЕ

При поглощении света полупроводником возможно такое возбуждение электрона валентной зоны, при котором он не переходит в зону проводимости, а образует с дыркой связанную систему. Такая система получила название **э к с и т о н а**.

Если размеры экситона велики по сравнению с постоянной решетки, то взаимодействие электрона и дырки можно представить как кулоновское взаимодействие двух точечных зарядов, ослабленное средой в ϵ раз. Пусть \mathbf{r}_n и \mathbf{r}_p — радиусы-векторы, определяющие положения электрона и дырки, а m_n^* и m_p^* — их скалярные эффективные массы. Тогда уравнение Шредингера для взаимодействующей пары электрон—дырка имеет вид:

$$\left[-\frac{\hbar^2}{2m_n^*} \nabla_{\mathbf{r}_n}^2 - \frac{\hbar^2}{2m_p^*} \nabla_{\mathbf{r}_p}^2 - \frac{e}{4\pi\epsilon\epsilon_0 (\mathbf{r}_n - \mathbf{r}_p)} \right] \Psi(\mathbf{r}_n, \mathbf{r}_p) = W_{xc} \Psi(\mathbf{r}_n, \mathbf{r}_p), \quad (11-64)$$

где $\nabla_{\mathbf{r}_n}^2$ и $\nabla_{\mathbf{r}_p}^2$ — операторы Лапласа в координатах электрона и дырки; W_{xc} — энергия экситона.

Если ввести радиус-вектор центра тяжести \mathbf{R} , характеризующий положение электронно-дырочной пары, и радиус-вектор \mathbf{r} относительно положения электрона и дырки

$$\mathbf{R} = \frac{m_n^* \mathbf{r}_n + m_p^* \mathbf{r}_p}{m_n^* + m_p^*}; \quad (11-65)$$

$$\mathbf{r} = \mathbf{r}_n - \mathbf{r}_p, \quad (11-66)$$

то уравнение (11-64) в новых координатах будет иметь вид:

$$\left[-\frac{\hbar^2}{2M^*} \nabla_{\mathbf{R}}^2 - \frac{\hbar^2}{2m_r^*} \nabla_{\mathbf{r}}^2 - \frac{e^2}{4\pi\epsilon\epsilon_0 |\mathbf{r}|} \right] \Psi(\mathbf{R}, \mathbf{r}) = W_{xc} \Psi(\mathbf{R}, \mathbf{r}), \quad (11-67)$$

где

$$M^* = m_n^* + m_p^* \quad \text{и} \quad m_r^* = \frac{m_n^* m_p^*}{m_n^* + m_p^*} \quad (11-68)$$

есть эффективная и приведенная масса экситона.

Собственные волновые функции экситонных состояний, удовлетворяющих уравнению (11-67), имеют форму

$$\Psi(\mathbf{R}, \mathbf{r}) = e^{i(\mathbf{k}_{xc}\mathbf{R})} \varphi(\mathbf{r}), \quad (11-69)$$

где $\mathbf{k}_{xc} = \mathbf{k}_n + \mathbf{k}_p$ — волновой вектор экситона, а \mathbf{k}_n и \mathbf{k}_p — волновые векторы электрона и дырки. Плоская волна $e^{i(\mathbf{k}_{xc}\mathbf{R})}$ описывает свободное движение экситона, а волновая функция $\varphi(\mathbf{r})$ описывает состояние внутреннего движения экситона, т. е. движение электрона и дырки относительно центра массы. Подстановка (11-69) в уравнение (11-67) сводит задачу к водородоподобной:

$$\left[-\frac{\hbar^2}{2m_r^*} \nabla_r^2 - \frac{e^2}{4\pi\epsilon\epsilon_0|\mathbf{r}|} \right] \varphi(\mathbf{r}) = \left(W_{xc} - \frac{\hbar^2 k_{xc}^2}{2M^*} \right) \varphi(\mathbf{r}) = E'_{xc} \varphi(\mathbf{r}). \quad (11-70)$$

Собственные значения энергии E'_{xc} , являющиеся энергией связи экситона, образуют водородоподобную серию

$$E'_{xc} = -\frac{e^4 m_r^*}{32\pi^2 \hbar^2 \epsilon^2 \epsilon_0^2} \frac{1}{n^2} = -\frac{E_{xc}}{n^2}, \quad (11-71)$$

где

$$E_{xc} = \frac{e^4 m_r^*}{32\pi^2 \hbar^2 \epsilon^2 \epsilon_0^2}; \quad (11-71a)$$

n — квантовое число, определяющее различные экситонные состояния ($n \geq 1$). Поскольку в экситоне приведенная масса меньше эффективной массы электрона и дырки (11-68), энергия связи экситона будет меньше энергии ионизации донора (2-174).

Следовательно, полная энергия экситона с учетом (11-71) равна:

$$W_{xc} = \frac{\hbar^2 k_{xc}^2}{2M^*} + E'_{xc} = \frac{\hbar^2 k_{xc}^2}{2M^*} - \frac{E_{xc}}{n^2}. \quad (11-72)$$

Примем дно зоны проводимости за начало отсчета энергии и будем считать, что ему соответствует начало сплошного спектра экситона ($n = \infty$). В этом случае энергетическая диаграмма различных экситонных состояний будет подобна представленной на рис. 11-21. Если оптический переход прямой и совершается при $k = 0$, экс-

тон будет характеризоваться водородоподобной серией линий поглощения, удовлетворяющих соотношению

$$h\nu = h\nu_{\infty} - \frac{E_{xc}}{n^2}, \quad (11-73)$$

где $h\nu_{\infty}$ совпадает с шириной запрещенной зоны при $k = 0$. Поэтому формулу (11-73) можно переписать в виде

$$h\nu = E_g - \frac{E_{xc}}{n^2}, \quad (11-74)$$

где E_{xc} — энергия связи экситона, определяемая зависимостью (11-71a).

В полупроводниках с прямыми переходами при $k=0$ экситону ($n=1$) соответствует очень узкий пик на краю поглощения, который уширяется с повышением температуры, как это показано на рис. 11-22.

В полупроводниках с непрямыми переходами также возможно образование экситонов. В этом случае длинноволновый край спектра поглощения экситона ($n=1$) определяется из условия для переходов с поглощением фотона

$$h\nu = E_g - E_p - E_{xc}, \quad (11-75)$$

а для переходов с испусканием фотона

$$h\nu = E_g + E_p - E_{xc}. \quad (11-76)$$

В полупроводниках с непрямыми переходами образование экситона сопровождается возникновением ступеньки на краю собственного поглощения. При этом возможны переходы с участием несколь-

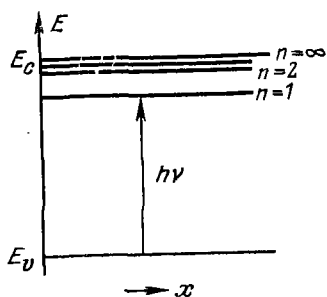


Рис. 11-21. Диаграмма энергетических уровней экситона и его возбужденных состояний. Энергия экситона отсчитывается от края зоны проводимости.

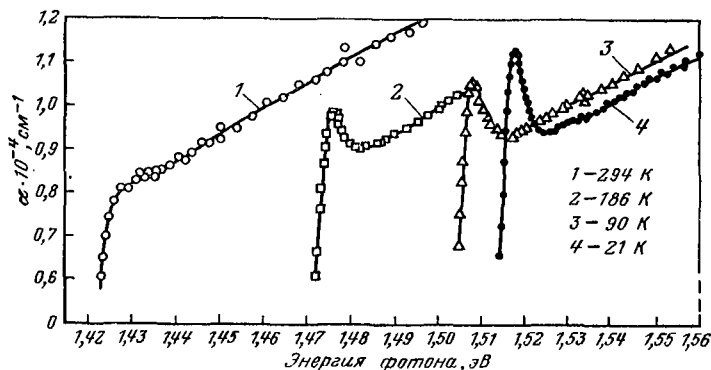


Рис. 11-22. Экситонное поглощение в GaAs.

ких фононов, испускание и поглощение которых возможны в различных комбинациях, поэтому на краю поглощения будет наблюдаться много ступенек, как, например, на рис. 11-23.

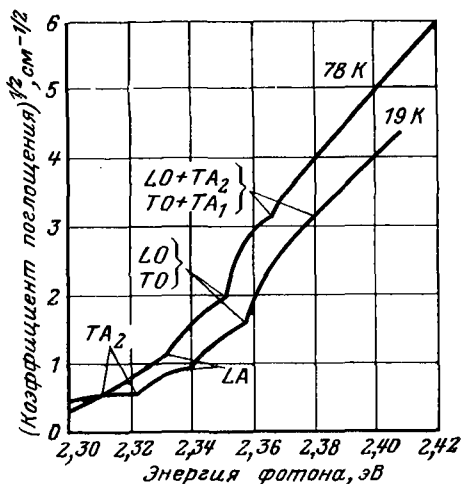
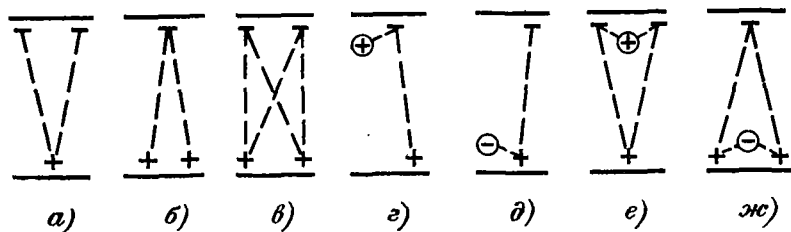


Рис. 11-23. Структура края поглощения GaP. Стрелками указаны пороговые энергии образования экситонов с эмиссией фононов.

Рис. 11-24. Экситонные комплексы.

а, б, в — свободные экситоны; г, д, е, ж — связанные экситоны.



При поглощении света полупроводником возможно и такое возбуждение электронов и дырок, когда кулоновское взаимодействие между ними приводит к объединению их с образованием ионоподобных или молекулородных комплексов. Так, два свободных электрона и одна свободная дырка или две свободные дырки и один свободный электрон, объединяясь, образуют экситонные ионы (рис. 11-24, а и 11-24, б соответственно). Кроме того, два свободных электрона и две свободные дырки могут объединяться в комплекс (рис. 11-24, в), напоминающий молекулу позитрония (рис. 11-25).

Свободная дырка может объединиться с нейтральным донором и образовать положительно заряженный экситонный ион (рис. 11-24, г). В этом случае электрон и дырка, связанные с донором, движутся вокруг донора (рис. 11-26), поэтому такой комплекс называется «связанным экситоном». Электрон, объединяясь

с нейтральным акцептором, также образует связанный экситон (рис. 11-24, д). Образование связанных экситонов может быть и в случае, когда два свободных электрона и одна свободная дырка

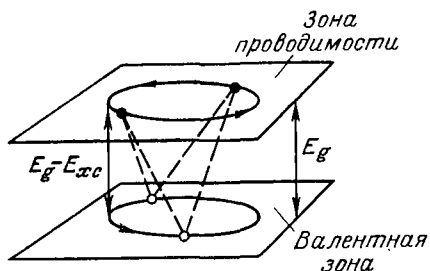


Рис. 11-25. Экситонный комплекс, состоящий из двух электронов, связанных с двумя дырками.

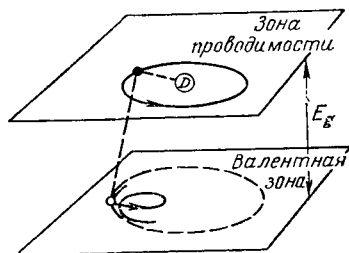


Рис. 11-26. Экситон, связанный с донором D .

связаны с донором (рис. 11-24, е) или когда две дырки и электрон связаны с акцептором (рис. 11-24, ж). Как свободные, так и связанные экситоны были обнаружены в ряде полупроводников.

11-7. ПОГЛОЩЕНИЕ СВОБОДНЫМИ НОСИТЕЛЯМИ ЗАРЯДА

При поглощении фотона электроны могут переходить в пределах той же долины с одного уровня на другой (рис. 11-27). Такие внутрizonные переходы носителей заряда происходят с нарушением правила отбора. В силу закона сохранения импульса они осуществляются либо когда наряду с поглощением фотона происходит поглощение или испускание фонона, либо когда имеется рассеяние на ионизованных примесях. Поэтому поглощение свободными электронами можно рассматривать методами, которые используются для не прямых 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

$$\text{rot } \mathbf{H} = \sigma \mathcal{E} + \epsilon \epsilon_0 \frac{\partial \mathcal{E}}{\partial t}; \quad (11-77)$$

$$\text{rot } \mathcal{E} = -\kappa \kappa_0 \frac{\partial \mathbf{H}}{\partial t}; \quad (11-78)$$

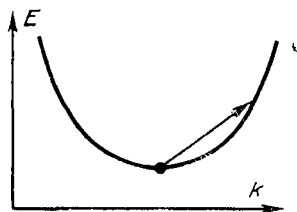


Рис. 11-27. Переход свободного электрона в зоне проводимости.

$$\operatorname{div} \mathbf{H} = 0; \quad (11-79)$$

$$\operatorname{div} \mathcal{E} = 0. \quad (11-80)$$

В этих уравнениях σ — удельная проводимость вещества при данной оптической частоте; κ , κ_0 — магнитная проницаемость, а ε , ε_0 — диэлектрическая постоянная среды и вакуума соответственно.

Из уравнений (11-77) и (11-78) следует:

$$\begin{aligned} \operatorname{rot} \operatorname{rot} \mathcal{E} &= \operatorname{rot} \left[-\kappa \kappa_0 \frac{\partial \mathbf{H}}{\partial t} \right] = -\kappa \kappa_0 \frac{\partial}{\partial t} \operatorname{rot} \mathbf{H} = \\ &= -\kappa \kappa_0 \left(\sigma \frac{\partial \mathcal{E}}{\partial t} + \varepsilon \varepsilon_0 \frac{\partial^2 \mathcal{E}}{\partial t^2} \right). \end{aligned} \quad (11-81)$$

С другой стороны, учитывая (11-80), имеем:

$$\begin{aligned} \operatorname{rot} \operatorname{rot} \mathcal{E} &= [\nabla (\nabla \mathcal{E})] = \nabla (\nabla \mathcal{E} - \nabla^2 \mathcal{E}) = \\ &= \operatorname{grad} \operatorname{div} \mathcal{E} - \nabla^2 \mathcal{E} = -\nabla^2 \mathcal{E}. \end{aligned} \quad (11-82)$$

Следовательно,

$$\nabla^2 \mathcal{E} - \kappa \kappa_0 \frac{\partial \mathcal{E}}{\partial t} - \kappa \kappa_0 \varepsilon \varepsilon_0 \frac{\partial^2 \mathcal{E}}{\partial t^2} = 0. \quad (11-83)$$

Пусть электромагнитная волна, падающая на полупроводник, распространяется вдоль направления x со скоростью v . Для этого случая решением уравнения (11-83) будет

$$\mathcal{E} = \mathcal{E}_0 e^{i\omega \left(t - \frac{x}{v} \right)} \quad (11-84)$$

при условии, если

$$\frac{1}{v^2} = \kappa \kappa_0 \varepsilon \varepsilon_0 - \frac{i \kappa \kappa_0 \sigma}{\omega}. \quad (11-85)$$

Но скорость распространения света в полупроводнике с комплексным показателем преломления

$$\bar{n}^* = \bar{n} - i\chi \quad (11-86)$$

связана со скоростью распространения света в вакууме формулой

$$v = \frac{c}{\bar{n}^*}. \quad (11-87)$$

Здесь χ (мнимая часть показателя преломления) называется коэффициентом экстинкции или показателем поглощения. Следовательно,

$$\bar{n}^{*2} = c^2 \kappa \kappa_0 \left(\varepsilon \varepsilon_0 - \frac{i\sigma}{\omega} \right). \quad (11-88)$$

Если учесть, что для вакуума $\bar{n}^* = 1$, $\kappa = 1$, $\varepsilon = 1$ и $\sigma = 0$, то $c^2 \kappa_0 \varepsilon_0 = 1$ или $c^2 = (\kappa_0 \varepsilon_0)^{-1}$. Кроме того, при определении оптических свойств полупроводников можно положить $\kappa = 1$. В соответствии с этим уравнение (11-88) запишется в виде

$$\bar{n}^{*2} = \varepsilon - \frac{i\sigma}{\varepsilon_0 \omega}. \quad (11-89)$$

На основе равенств (11-86) и (11-89) имеем:

$$\bar{n}^{*2} = \bar{n}^2 - \chi^2 - i2\bar{n}\chi = \varepsilon - \frac{i\sigma}{\varepsilon_0 \omega}, \quad (11-90)$$

откуда следует, что

$$\bar{n}^2 = \chi^2 + \varepsilon; \quad (11-91)$$

$$\chi = \frac{\sigma}{2\bar{n}\varepsilon_0 \omega}. \quad (11-92)$$

Поскольку на основе равенств (11-87) и (11-86)

$$\frac{1}{v} = \frac{\bar{n}}{c} - \frac{i\chi}{c}, \quad (11-93)$$

то, подставляя это выражение в формулу (11-84), получаем:

$$\mathcal{E} = \mathcal{E}_0 e^{i\omega \left(t - \frac{\bar{n}^*}{c} x \right)} = \mathcal{E}_0 e^{-\frac{\omega\chi}{c} x} e^{i\omega \left(t - \frac{\bar{n}}{c} x \right)}. \quad (11-94)$$

В этой формуле множитель $e^{-\frac{\omega\chi}{c} x}$ выражает собой затухание волны. Так как падающая мощность равна

$$P(0) = j\mathcal{E}(0) = \sigma \mathcal{E}^2(0) = \sigma \mathcal{E}_0^2 e^{i2\omega \left(t - \frac{\bar{n}}{c} x \right)},$$

а мощность, оставшаяся после прохождения в полупроводнике расстояния x , есть

$$P(x) = \sigma \mathcal{E}^2(x) = \sigma \mathcal{E}_0^2 e^{-\frac{2\omega\chi}{c} x} e^{i2\omega \left(t - \frac{\bar{n}}{c} x \right)},$$

то прошедшая мощность по отношению к падающей составляет величину

$$\frac{P(x)}{P(0)} = e^{-\frac{2\omega\chi}{c} x}. \quad (11-95)$$

Но согласно формуле (11-4) эту величину можно выразить через коэффициент поглощения

$$\frac{P(x)}{P(0)} = e^{-\alpha x}. \quad (11-96)$$

Из сопоставления формул (11-96) и (11-95), а также с учетом равенства (11-92) получим:

$$\alpha = \frac{2\omega\chi}{c} = \frac{\sigma}{c\bar{n}\epsilon_0}. \quad (11-97)$$

Таким образом, коэффициент поглощения свободными носителями заряда определяется проводимостью вещества.

Поскольку проводимость $\sigma = en\mu_n = \frac{e^2 \langle \tau \rangle}{m_n^*} n$, то коэффициент поглощения запишется через концентрацию свободных носителей заряда в виде

$$\alpha = \frac{\sigma}{c\bar{n}\epsilon_0} = \frac{e^2 \langle \tau \rangle}{c\bar{n}\epsilon_0 m_n^*} n. \quad (11-98)$$

Следовательно, коэффициент поглощения свободными носителями заряда зависит от их концентрации n и эффективной массы m_n^* , среднего времени релаксации $\langle \tau \rangle$ 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 удельной проводимости σ или показателя поглощения χ от длины падающего света λ , 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. $\frac{dx}{dt}$. Поэтому уравнение движения электрона проводимости в электрическом поле падающей волны запишется как

$$m_n^* \frac{d^2x}{dt^2} + m_n^* g \frac{dx}{dt} = e\mathcal{E}e^{i\omega t}. \quad (11-99)$$

Решение этого уравнения будем искать в виде

$$x = Be^{i\omega t}. \quad (11-100)$$

Подставляя (11-100) в выражение (11-99), будем иметь:

$$-m_n^* \omega^2 B + im_n^* g B \omega = e\mathcal{E}, \quad (11-101)$$

откуда найдем:

$$B = \frac{e\mathcal{E}}{m_n^* \omega (-\omega + ig)}. \quad (11-102)$$

Если в единице объема среды имеется n свободных электронов, то поляризуемость ее составляет $P = \frac{enB}{g}$, а диэлектрическая постоянная равна $\epsilon^* = \epsilon_0 + P$. Таким образом,

$$\epsilon^* = \epsilon_0 + \frac{e^2 n}{m_n^* \omega (-\omega + ig)}. \quad (11-103)$$

Но комплексный показатель преломления среды есть $\bar{n}^* = (\epsilon^*/\epsilon_0)^{1/2}$, поэтому

$$(\bar{n} - i\chi)^2 = \frac{\epsilon^*}{\epsilon_0} = 1 + \frac{e^2 n}{\epsilon_0 m_n^* \omega (-\omega + ig)}, \quad (11-104)$$

откуда следует, что

$$\bar{n}^2 - \chi^2 - 1 = - \frac{e^2 n}{\epsilon_0 m_n^* (\omega^2 + g^2)}; \quad (11-105)$$

$$2\bar{n}\chi = \frac{e^2 ng}{\epsilon_0 m_n^* \omega (\omega^2 + g^2)}. \quad (11-106)$$

В случае стационарного поля, для которого из (11-99)

$$m_n^* g \frac{dx}{dt} = e\mathcal{E}, \quad (11-107)$$

а скорость движения электрона проводимости

$$v = \frac{dx}{dt} = \mu_n \mathcal{E}, \quad (11-108)$$

имеем:

$$g = \frac{e\mathcal{E}}{m_n^* v} = \frac{e}{m_n^* \mu_n}. \quad (11-109)$$

Поэтому соотношение (11-106) можно записать в виде

$$2\bar{n}\chi = \frac{en\mu_n}{\epsilon_0 \omega \left[1 + \left(\frac{m_n^* \mu_n \omega}{e} \right)^2 \right]} = \frac{\sigma}{\epsilon_0 \omega \left[1 + \left(\frac{m_n^* \mu_n \omega}{e} \right)^2 \right]}. \quad (11-110)$$

Соответственно

$$\alpha = \frac{2\omega\chi}{c} = \frac{\sigma}{cn\epsilon_0 \left[1 + \left(\frac{m_n^* \mu_n \omega}{e} \right)^2 \right]}. \quad (11-111)$$

Если $\frac{m_n^* \mu_n \omega}{e} \gg 1$ и учитывая, что $\omega = 2\pi\nu = \frac{2\pi c}{\lambda}$, то из равенства (11-111) следует:

$$\alpha = \frac{e^2 n}{4\pi^2 c^3 \epsilon_0 \bar{n} m_n^{*2} \mu_n} \lambda^2 = \frac{e^2 n}{4\pi^2 c^3 \epsilon_0 \bar{n} m_n^* \langle \tau \rangle} \lambda^2. \quad (11-112)$$

Как видно, коэффициент поглощения свободными носителями заряда пропорционален их концентрации, квадрату длины волны падающего света и обратно пропорционален эффективной массе и времени релаксации.

Зависимость α от λ^2 хорошо выполняется для сурьмянистого индия n -типа, которая представлена на рис. 11-28. Кривая a соответствует концентрации электронов $3,5 \cdot 10^{17} \text{ см}^{-3}$, кривая b — концентрации $6,2 \cdot 10^{17} \text{ см}^{-3}$.

Поскольку время релаксации $\langle \tau \rangle$ зависит от природы рассеивающих центров, то можно ожидать, что на зависимость коэффициента поглощения свободными носителями заряда от длины волны падающего света будет оказывать влияние механизм рассеяния.

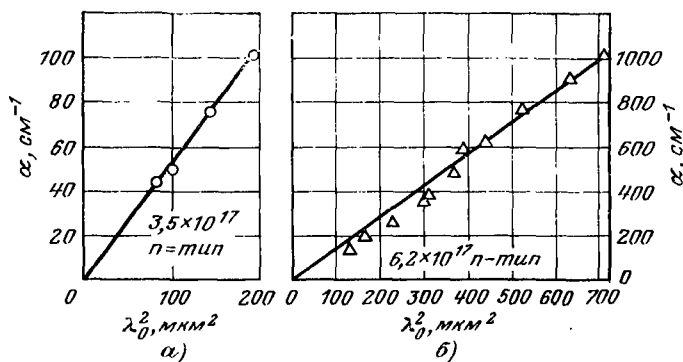


Рис. 11-28. Зависимость коэффициента поглощения InSb n -типа от квадрата длины волны для концентрации электронов $3,5 \cdot 10^{17} \text{ см}^{-3}$ (а) и $6,2 \cdot 10^{17} \text{ см}^{-3}$ (б).

Действительно, рассеяние на акустических фононах приводит к поглощению, меняющемуся как $\lambda^{1,5}$, рассеяние на оптических фононах определяет зависимость $\lambda^{2,5}$, а рассеяние ионизованными примесями дает зависимость $\lambda^{3,5}$. В общем случае, когда реализуются все типы рассеяния, коэффициент поглощения свободными носителями заряда равен сумме трех членов

$$\alpha = A\lambda^{1,5} + B\lambda^{2,5} + C\lambda^{3,5}, \quad (11-113)$$

где A, B, C — константы. В зависимости от концентрации легирующих примесей, а следовательно, от концентрации свободных элект-

тронов будет доминирующим тот или иной механизм рассеяния и $\alpha \sim \lambda^p$. В табл. 11-3 приведены значения p для различных соединений.

Таблица 11-3

Полупроводник	Концентрация электронов, 10^{17} см^{-3}	p	Полупроводник	Концентрация электронов, 10^{17} см^{-3}	p
Ge	0,5—5	2	InP	0,4—4	2,5
GaP	10 $\frac{3}{4}$	1,8	InAs	0,3—8	3
GaAs	1—5	3	InSb	1—3	2
GaSb	0,5	3,5	AlSb	0,4—4	2

У большинства полупроводников валентная зона состоит из трех подзон, разделенных за счет спин-орбитального взаимодействия (такая зонная структура показана на рис. 11-29). Поэтому в полупроводниках p -типа, где вершина валентной зоны занята дырками, возможно оптическое поглощение, обусловленное переходами электронов между отдельными подзонами: а) из зоны легких

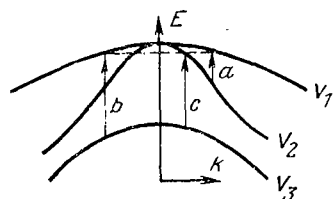
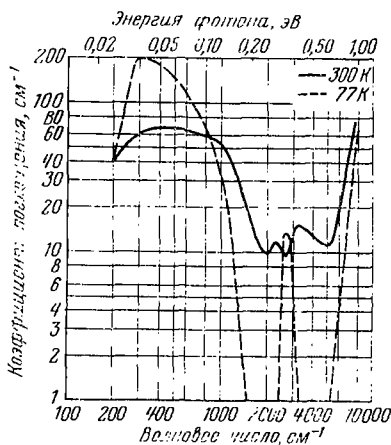


Рис. 11-29. Структура валентных подзон и внутренние переходы.

Рис. 11-30. Поглощение, связанное с переходами между валентными подзонами в германии p -типа.



дырок V_2 в зону тяжелых дырок V_1 ; б) из отщепленной зоны V_3 в зону тяжелых дырок V_1 ; в) из отщепленной зоны V_3 в зону легких дырок V_2 . Для таких прямых переходов соблюдаются правила отбора. Интенсивность и положение пиков поглощения должны меняться с изменением степени легирования или температуры. Если ширина запрещенной зоны полупроводника меньше энергии спин-орбитального расщепления, то переходы $V_3 \rightarrow V_1$ и $V_3 \rightarrow V_2$ будут трудно различимы на фоне собственного поглощения. В качестве иллюстрации на рис. 11-30 приведено поглощение, связанное с переходами электронов между валентными подзонами в германии p -типа.

11-8. ПРИМЕСНОЕ ПОГЛОЩЕНИЕ

Поглощение света, приводящее к ионизации или возбуждению примесных центров в кристалле, называется примесным поглощением. Допустим, что в кристаллической решетке полупроводника имеется донорная или акцепторная примесь (рис. 11-31, а, б) или примесь с глубокими энергетическими уровнями (рис. 11-31, в, г). Тогда при освещении полупроводника светом соответствующей длины волны будет наблюдаться непрерывная полоса поглощения, вызываемая переходами электронов с энергетических уровней примеси в зону проводимости или переходами электронов валентной зоны на примесные уровни. Так как энергия ионизации примесного центра меньше энергии, требуемой для перевода электрона из валентной зоны в зону проводимости, полосы примесного поглощения располагаются за краем фундаментальной абсорбции.

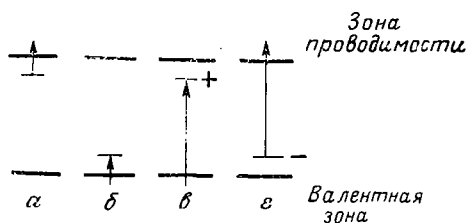


Рис. 11-31. Поглощение, связанное с переходами между примесными уровнями и зонами.

а — донор — зона проводимости; б — валентная зона — акцептор; в — валентная зона — донор; г — акцептор — зона проводимости.

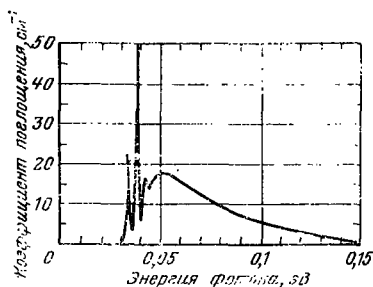


Рис. 11-32. Зависимость коэффициента поглощения кремния, легированного бором, от энергии фотона.

Кроме основных состояний примесные центры могут иметь и энергетические уровни возбуждения, энергия которых для доноров и акцепторов согласно (2-174) определяется выражением

$$E = 13,52 \frac{Z^2}{e^2} \left(\frac{m^*}{m_0} \right) \frac{1}{n^2}, \quad (11-114)$$

где $n = 1, 2 \dots$

Возбуждение электронов примеси, т. е. перевод электронов примеси из основного состояния в возбужденное, приведет к появлению поглощения, в спектре которого будет наблюдаться несколько полос. В качестве примера на рис. 11-32 представлен спектр поглощения кремния, легированного бором с $N_a = 1,8 \cdot 10^{16} \text{ см}^{-3}$, при 4,2 К.

11-9. РЕШЕТОЧНОЕ ПОГЛОЩЕНИЕ

Бинарные соединения $A^{III}B^V$ состоят из атомов различного типа и имеют ионную составляющую химической связи. Поэтому их можно рассматривать как состоящие из электрических диполей. Эти диполи могут взаимодействовать с электромагнитным излучением. Наиболее сильное поглощение электромагнитного излучения имеет место, когда частота излучения равна частоте собственных колебаний диполей. Поглощение света полупроводником, связанное

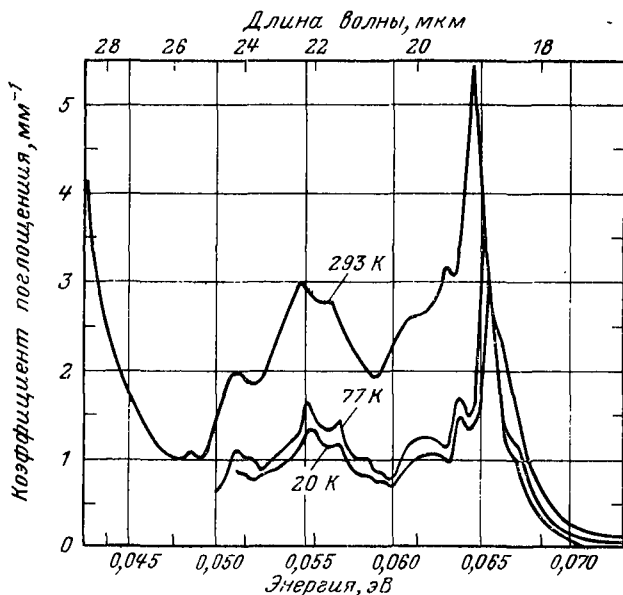


Рис. 11-33. Спектр решеточного поглощения GaAs в области энергии 0,04—0,07 эВ при разных температурах.

с возбуждением колебаний кристаллической решетки, называется решеточным поглощением. Так как решетка кристалла может поглощать энергию электромагнитного поля излучения только при определенных значениях энергии фотона, спектр ее характеризуется рядом пиков поглощения, лежащих в далекой инфракрасной области спектра.

Значения частот излучения, поглощаемого колебаниями решетки, определяются совокупностью оптических и акустических ветвей колебаний. В полупроводниках имеются две поперечные оптические (ТО) ветви колебаний, две поперечные акустические (ТА), одна продольная оптическая (ЛО) и одна продольная акустическая (ЛА).

При поглощении света решеткой должен выполняться закон сохранения квазимпульса, т. е. должны поглощаться только та-

кие фотоны, импульс которых равен квазиимпульсу фонона. Но импульс фотона $P = h/\lambda_{\text{фот}}$ мал по сравнению с квазиимпульсом фонона $q = h/\lambda_{\text{фон}}$, поэтому для выполнения закона сохранения квазиимпульса должны быть испущены два или более фононов. Все это обуславливает сложную структуру спектра поглощения света колебаниями решетки. На рис. 11-33 приведена часть спектра решеточного поглощения электронного арсенида галлия, а идентификация этих пиков поглощения дана в табл. 11-4.

Таблица 11-4

Тип и энергии фононов, участвующих в поглощении в арсениде галлия

Положение максимума, эВ	Тип и энергия фононов, эВ	Положение максимума, эВ	Тип и энергия фононов, эВ
0,0648 {	$TO_1 \rightarrow TO_1$ 0,0324 : 0,0324	0,058 {	$LO \quad LO$ 0,029 : 0,029
0,0631 {	$TO_2 \rightarrow TO_2$ 0,0316 : 0,0316	0,0565 {	$TO_1 \quad LA$ 0,0324 : 0,0241
{	$TO_1 \rightarrow LO$ 0,0324 : 0,0288	0,0548 {	$TO_2 \quad LA$ 0,0316 : 0,0232
0,0612 {	$TO_2 \rightarrow LO$ 0,0316 : 0,0296	0,0510 {	$LO \quad LA$ 0,0288 : 0,0222

В гомополярных полупроводниках, таких как кремний и германий, нет диполей. Но их атомы под воздействием поля световой волны приобретают дипольный момент. Индуцируемый диполь сильно взаимодействует с излучением¹ и в результате возникают фононы. Решеточное поглощение германия лежит в инфракрасной

Таблица 11-5

Фононы, участвующие в поглощении в кремнии

Волновое число, мм^{-1}	Энергия пика, эВ	Тип фононов ¹
144,8	0,1795	3 TO
137,8	0,1708	2 $TO \quad LO$
130,2	0,1614	2 $TO \quad LO$
96,4	0,1195	2 TO
89,6	0,1111	$TO \rightarrow LO$
81,9	0,1015	$TO \rightarrow LA$
74,0	0,0917	$LO \quad LA$
68,9	0,0756	$TO \quad TA$
61,0	0,0702	$LO \quad TA$

¹ Здесь $TO = 0.0598$ эВ; $LO = 0.0513$ эВ; $LA = 0.0414$ эВ; $TA = 0.0158$ эВ.

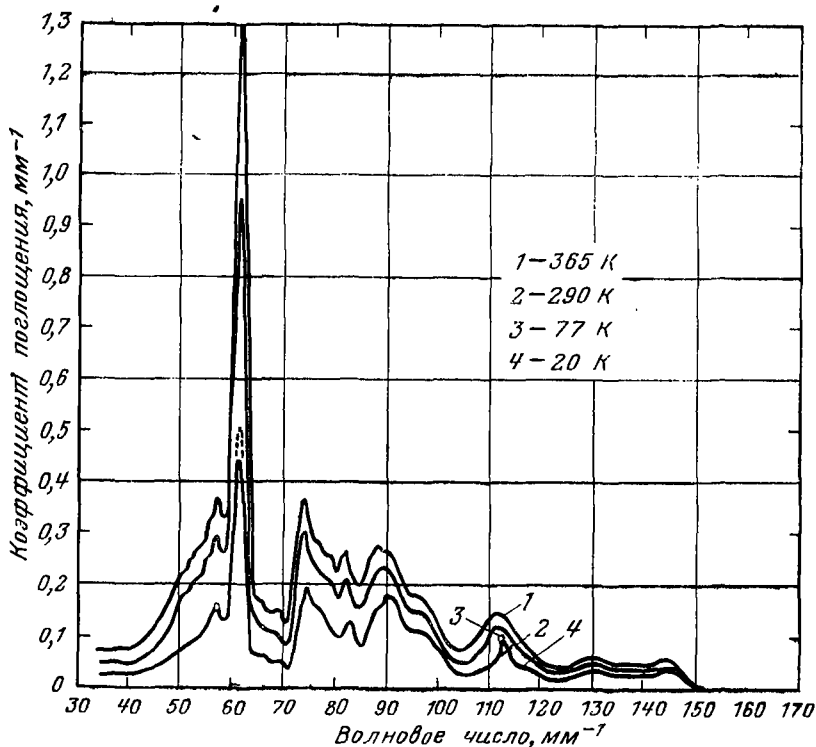


Рис. 11-34. Решеточное поглощение кремния при разных температурах.

области спектра от 14 до 33 мкм, а кремния — в области от 7 до 30 мкм. На рис. 11-34 показан спектр решеточного поглощения кремния, а идентификация фононов приведена в табл. 11-5.

Интенсивность и положение полос поглощения решетки не меняются с изменением в ней концентрации примесей вплоть до величины 10^{18} см⁻³. Однако с понижением температуры, как видно из рис. 11-33 и 11-34, интенсивность полос решеточного поглощения уменьшается, а их положение не изменяется.

Глава двенадцатая

ЛЮМИНЕСЦЕНЦИЯ ПОЛУПРОВОДНИКОВ

12-1. ТИПЫ ЛЮМИНЕСЦЕНЦИИ

У полупроводников, находящихся в возбужденном состоянии, может наблюдаться испускание электромагнитного излучения — лю м и н е с ц е н ц и я, которая является процессом, обратным поглощению. Такое состояние вещества можно осуществлять раз-

личными способами. При фотолюминесценции излучение появляется за счет поглощаемой энергии света. Свечение, возникающее при бомбардировке образца электронами, называют катодолюминесценцией. Люминесценция может наблюдаться при возбуждении вещества рентгеновским (рентгелюминесценция) и γ -излучением (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 излучением 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 и др.

Люминесценция, как любое электромагнитное излучение, характеризуется интенсивностью, спектральным составом, поляризацией, когерентностью и длительностью послесвечения. И только длительность послесвечения отличает люминесценцию от всех других видов неравновесного излучения, которые практически безынерционны; их длительность составляет примерно период световых колебаний (10^{-15} с). В случае люминесценции акты поглощения и излучения света разделены промежуточными процессами, что и приводит к продолжительному существованию свечения после прекращения возбуждения.

У твердых тел различают три вида люминесценции: мономолекулярную, метастабильную и рекомбинационную. Люминесценция, которая происходит во время возбуждения, носит название флюоресценции. Люминесценция, которая продолжается некоторое время после окончания возбуждения, называется фосфоресценцией.

12-2. МОНОМОЛЕКУЛЯРНОЕ СВЕЧЕНИЕ ТВЕРДЫХ ТЕЛ

Мономолекулярное и метастабильное излучение твердых тел имеет место в том случае, когда поглощение и излучение света происходят внутри примесного центра.

Спектры люминесценции твердых тел, т. е. зависимость интенсивности свечения от энергии излученного фотона или длины волны, отличаются от соответствующих спектров изолированных атомов примеси двумя основными чертами. Во-первых, излучение

твердых тел наблюдается обычно в виде широких полос, и, во-вторых, их спектр люминесценции, как правило, смещен в длинноволновом направлении по сравнению со спектром поглощения, т. е. излучаемая энергия меньше, чем поглощаемая. Разницу этих двух энергий называют «сдвигом Франка—Кондона». Эти особенности оптических свойств твердых тел обязаны своим происхождением взаимодействию центра свечения с полем кристаллической решетки. При взаимодействии центра с ближайшим окружением его потенциальная энергия складывается из кулоновского притяжения, обменного взаимодействия и ван-дер-ваальсовского притяжения. В результате этих сил взаимодействия атомные уровни примеси основного U_0 и возбужденного U_1 состояний в кристаллической решетке будут квази-молекулярными, как показано на рис. 12-1.

При такой ситуации возможно поглощение света примесным центром, в результате которого электрон с основного уровня U_0 перейдет на возбужденный U_1 . Самопроизвольный обратный переход электрона с возбужденного уровня центра на основной будет сопровождаться излучением. Если учесть заселенность колебательных уровней электронами в основном и возбужденном состояниях, то совокупность соответствующих фотонов образует полосы поглощения и излучения, которые на рис. 12-1 изображены на оси ординат — шкале энергии. Длительность такой люминесценции твердого тела будет определяться временем жизни электрона в возбужденном состоянии центра. Оно постоянно, не зависит от внешних факторов и, в частности, для дипольного излучения атомов составляет примерно 10^{-8} с.

В том случае, когда люминесценция происходит при воздействии на возбужденный примесный центр внешних факторов, сообщающих ему энергию активации, свечение называют метастабильным. Это излучение запаздывает относительно возбуждения на время, равное длительности пребывания на метастабильном уровне примеси, которое непостоянно и зависит от температуры.

12-3. РЕКОМБИНАЦИОННОЕ ИЗЛУЧЕНИЕ ПОЛУПРОВОДНИКОВ ПРИ ФУНДАМЕНТАЛЬНЫХ ПЕРЕХОДАХ

При возбуждении полупроводника, создающем электронно-дырочные пары, рекомбинационное излучение может возникнуть при непосредственной рекомбинации свободных электронов и свобод-

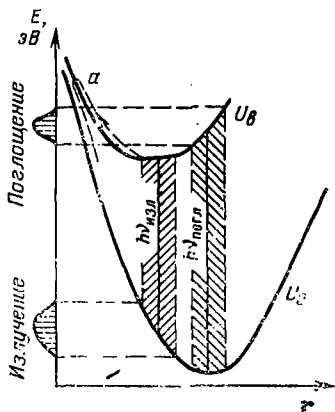


Рис. 12-1. Схема энергетических уровней примесного центра в твердом теле.

ных дырок, при аннигиляции экситона, а также при рекомбинации свободных электронов и свободных дырок через рекомбинационные ловушки, которые выступают здесь в качестве центров свечения. Сначала рассмотрим фундаментальные переходы, при которых изменение энергии равно ширине запрещенной зоны или близко к ней.

1. Прямые переходы «зона проводимости — валентная зона»

В полупроводнике с прямой запрещенной зоной поглощение света сопровождается образованием свободного электрона и свободной дырки, волновые векторы которых одинаковы ($k = k'$).

После процесса поглощения света образовавшиеся свободные носители заряда претерпевают рассеяние, в результате чего за время релаксации электрон опускается на дно зоны проводимости, а дырка поднимается к потолку валентной зоны. Такое состояние между свободными носителями заряда и кристаллической решеткой устанавливается за время порядка $10^{-10} - 10^{-12}$ с. Прямая рекомбинация свободного электрона и свободной дырки с излучением фотона наиболее вероятна, если после процесса релаксации волновые векторы электрона и дырки будут одинаковы (рис. 12-2).

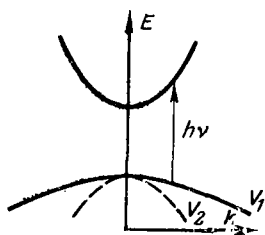


Рис. 12-2. Прямые излучательные переходы.

В этом случае, как и при соответствующих процессах поглощения [формула (11-24)], спектр излучения описывается выражением

$$I(h\nu) = B(h\nu - E_g)^{1/2}, \quad (12-1)$$

где B — постоянная.

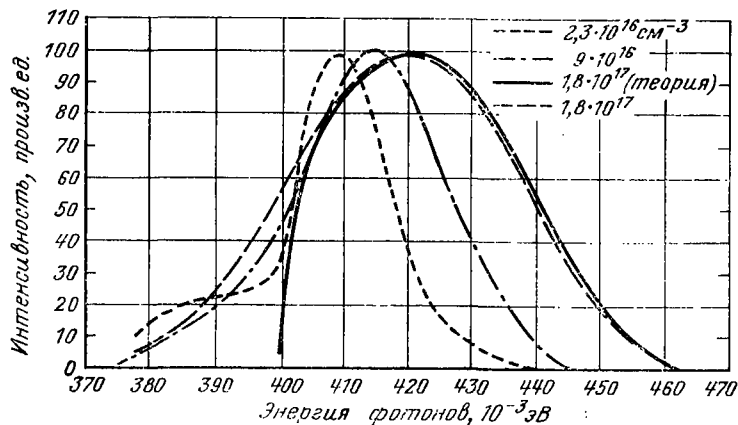


Рис. 12-3. Излучение InSb n -типа, 77 К.

Из этой формулы следует, что излучение при рекомбинации свободных носителей заряда должно иметь порог со стороны низких энергий при $h\nu = E_g$ (рис. 12-3, сплошная кривая). При увеличении температуры, а также при возрастании энергии возбуждения заполняются более высокие состояния в зоне проводимости. Это обусловит излучение при более высоких энергиях фотонов, приводящее к появлению в спектре люминесценции высокоэнергетического «хвоста».

В качестве примера прямой излучательной межзонной рекомбинации на рис. 12-3 представлен спектр фотолюминесценции InAs n -типа с различным уровнем легирования. Сдвиг максимума излучения и его высокоэнергетического края в сторону больших энергий фотонов при возрастании легирования обусловлен смещением уровня Ферми в зону проводимости.

2. Непрямые переходы «зона проводимости — валентная зона»

В полупроводнике с непрямой запрещенной зоной поглощение света происходит с поглощением или излучением фона. В таком полупроводнике при рекомбинации свободного электрона и сво-

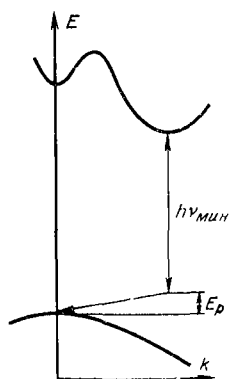


Рис. 12-4. Непрямые излучательные переходы.

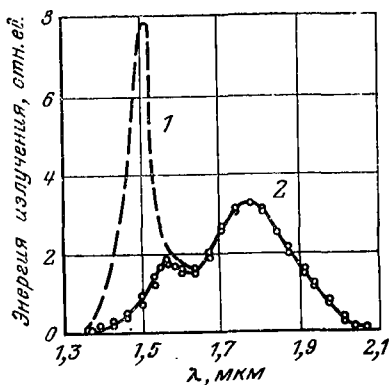


Рис. 12-5. Спектр излучения образца германия толщиной $1,3 \cdot 10^{-3}$ см. Кривая 2 — экспериментальная; кривая 1 получена с учетом реабсорбции (самопоглощения).

бодной дырки должны также принимать участие фононы, которые обеспечивают сохранение квазиимпульса (рис. 12-4). Наиболее вероятным процессом является эмиссия фона. Процесс поглощения фона не существует при излучении фона, поскольку его вероятность в $N_p/(N_p + 1)$ раз меньше вероятности испускания фона. Оптический переход, сопровождаемый испусканием фона, имеет меньшую энергию, чем ширина запрещенной зоны,

равную $h\nu_{\min} = E_g - E_p$. В таком случае по аналогии с коэффициентом поглощения [формула (11-35)] спектр излучения описывается выражением

$$I(h\nu) = B(h\nu - E_g + E_p)^2, \quad (12-2)$$

где E_p — энергия фонона, участвующего в рекомбинации.

Если у полупроводника возможны прямые и непрямые переходы, в спектре его излучения наблюдаются две полосы люминесценции, как, например, в германии (рис. 12-5). Длинноволновая полоса излучения определяется непрямой излучательной рекомбинацией. Коротковолновая полоса люминесценции обусловлена переходом электронов с излучением из прямой долины, которая расположена на 0,15 эВ выше дна самой низкой долины. Кривая 2 на этом рисунке — экспериментальная, а кривая 1 получена из кривой 2 путем введения поправки на поглощение в образце.

3. Сильно легированный полупроводник

У сильно легированного полупроводника уровень Ферми лежит внутри разрешенной зоны (в зоне проводимости для n -типа и в валентной зоне для p -типа). На рис. 12-6 приведена энергетическая диаграмма полупроводника, в котором за счет возбуждения создана концентрация электронов в зоне проводимости.

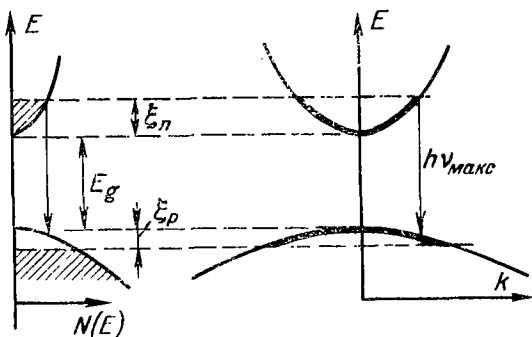


Рис. 12-6. Прямые переходы в случае вырождения.

Для полупроводника с прямой запрещенной зоной для прямых переходов интенсивность излучения с увеличением частоты должна расти как

$$I(h\nu) = B(h\nu - E_g)^{1/2}. \quad (12-3)$$

В случае непрямых переходов, при которых закон сохранения квазиимпульса осуществляется за счет процессов электрон-электронного рассеяния или рассеяния «электрон—примесь», все занятые состояния в зоне проводимости могут давать начало излуча-

тельными переходам во все пустые состояния валентной зоны. В этом случае одной и той же энергии $h\nu$ могут отвечать переходы между различными верхними и нижними состояниями, поэтому спектр излучения описывается выражением

$$I(h\nu) = B(h\nu - E_g)^2. \quad (12-4)$$

4. Экситонная рекомбинация

В достаточно чистом полупроводнике при возбуждении светом, энергия которого близка к ширине запрещенной зоны, возможно возникновение экситона — свободного электрона и свободной дырки, связанных силой кулоновского притяжения.

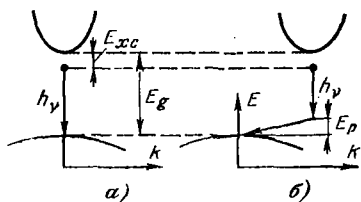


Рис. 12-7. Экситонная рекомбинация.

a — прямая; *б* — непрямая.

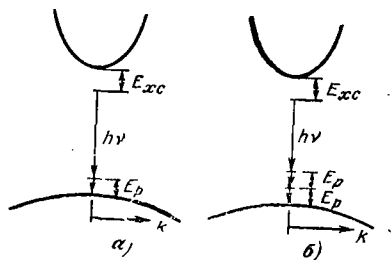


Рис. 12-8. Прямая экситонная рекомбинация.

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

дырки, связанных силой кулоновского притяжения. При рекомбинации таких носителей заряда, т. е. при аннигиляции экситона, будет излучаться узкая спектральная линия. В полупроводнике с прямой запрещенной зоной, имеющего разрешенные переходы (рис. 12-7, *a*), энергия экситонного излучения равна:

$$h\nu = E_g - E_{xc}, \quad (12-5)$$

где E_{xc} — энергия связи экситона.

Экситон может иметь возбужденные состояния, энергия ионизации которых согласно формуле (11-71) в n^2 раз меньше, чем энергия ионизации основного состояния, соответствующего $n = 1$. Поэтому излучение свободного экситона может состоять из серии узких линий.

В полупроводнике с непрямой запрещенной зоной (рис. 12-7, *б*), у которого закон сохранения ква-

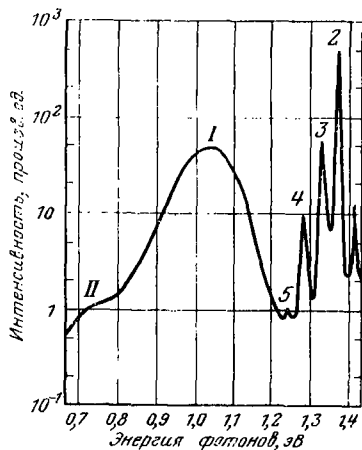


Рис. 12-9. Зависимость фотолуминесценции InP при 6 К от энергии фотонов.

импульса осуществляется за счет испускания оптического фотона с энергией E_p , энергия излученного фотона

$$h\nu = E_g - E_{xc} - E_p. \quad (12-6)$$

У полупроводников с прямыми разрешенными переходами излучение свободного экситона может также происходить с испусканием одного или несколько фононов (рис. 12-8), но вероятность таких переходов уменьшается и соответствующая линия излучения становится слабее.

В присутствии примесей могут образоваться связанные экситоны. Их спектр излучения состоит из узких линий с энергиями меньшими, чем при рекомбинации свободного экситона. На рис. 12-9 в качестве примера приведен спектр фотолюминесценции достаточно чистого фосфида индия. Пик излучения, обозначенный цифрой 1, относят к рекомбинации свободного экситона, а линии 2, 3, 4 и 5 обусловлены рекомбинацией связанного экситона с испусканием соответственно 0, 1, 2 и 3 продольных оптических фононов LO , энергия которого составляет 43 мэВ. Полосы излучения I и II вызваны процессами неэкситонного характера.

12-4. РЕКОМБИНАЦИОННОЕ ИЗЛУЧЕНИЕ ПРИ ПЕРЕХОДАХ МЕЖДУ ЗОНОЙ И ПРИМЕСНЫМИ УРОВНЯМИ

В примесных полупроводниках, в которых в качестве примеси могут выступать не только доноры и акцепторы, но и разного рода дефекты, нарушения кристаллической периодичности на поверхности полупроводника; сверхстехиометрический избыток одного из

компонентов в случае сложного полупроводникового материала, важную роль в процессах люминесценции играют 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

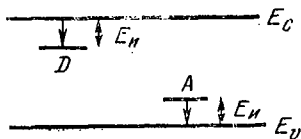


Рис. 12-10. Мелкие переходы.

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.

Переходы при нейтрализации ионизированных доноров и акцепторов (рис. 12-10) в принципе могут являться излучательными, давая свечение в далекой инфракрасной области спектра. Однако, как показали расчеты, вероятность переходов с испусканием фо-

понов значительно больше, чем вероятность излучательных переходов. Например, для германия n -типа при 4 К сечение захвата для переходов с испусканием фотона равно примерно 10^{-12} см², а сечение захвата при переходе с излучением фотона составляет $4 \cdot 10^{-19}$ см². Процесс безызлучательной рекомбинации происходит следующим образом. Сначала электрон зоны проводимости захватывается на возбужденный уровень донора. Затем он совершает каскадные переходы на нижележащие энергетические уровни, испуская фотон при каждом переходе.

1. Переход «зона — уровень примеси»

У полупроводников n - и p -типа, как следует из рис. 12-11, возможны излучательные переходы с донорного уровня в валентную зону или из зоны проводимости на акцепторный уровень. Такие центры отличаются от ловушек лишь несколько большей глубиной залегания уровней и тем, что захват дырки из валентной зоны хотя и менее вероятен по сравнению с захватом электрона из зоны проводимости, но все же имеет заметную величину ($\sigma_n \gg \sigma_p$). Согласно рис. 12-11, *a* сначала происходит захват электрона из зоны проводимости локальным уровнем (переход 2), а затем осуществляется излучательная рекомбинация локализованного электрона и дырки валентной зоны (переход 3).

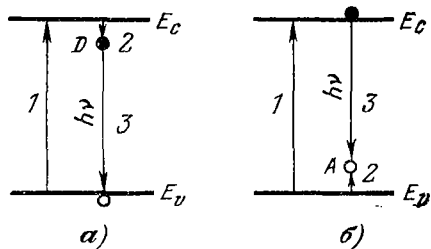


Рис. 12-11. Излучательные переходы между зоной и примесными состояниями.

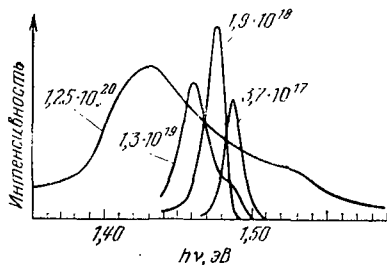


Рис. 12-12. Спектр катодолуминесценции при 4,2 К арсенида галлия, легированного цинком.

В другой модели люминесценции (рис. 12-11, *б*) локализованный уровень расположен немного выше валентной зоны и имеет большую вероятность захвата дырки ($\sigma_p \gg \sigma_n$), но вместе с тем имеет и заметную вероятность захвата электронов зоны проводимости. Люминесцентное излучение возникает при рекомбинации свободного электрона зоны проводимости с захваченной этим уровнем дыркой (переход 3).

При прямых переходах излучается фотон с энергией

$$h\nu = E_g - E_n, \quad (12-7)$$

а при непрямых переходах энергия излученного фотона равна:

$$h\nu = E_g - E_n - E_p, \quad (12-8)$$

где E_n — энергия ионизации примесного атома; E_p — энергия фонона, принимающего участие в рекомбинации. Такого типа излучательная рекомбинация возможна только для полупроводников с относительно низкой концентрацией примесей. Увеличение концентрации доноров (или акцепторов) вызовет размытие энергетического уровня примеси и превращение его в примесную зону, которая будет смыкаться с краями основной зоны. В качестве примера на рис. 12-12 приведен спектр катодолюминесценции арсенида галлия, легированного цинком. Из рисунка видно, что при возрастании концентрации акцепторов имеет место расширение полосы излучения и наблюдается сдвиг ее максимума в сторону низких энергий, что вызывается сужением запрещенной зоны, обусловленным высоким уровнем легирования.

У полупроводников с непрямой запрещенной зоной при малых концентрациях доноров в переходах «донор—валентная зона», а также в переходах «зона проводимости — акцептор» должен принимать участие фонон. На рис. 12-13 приведены спектры излучения для образцов кремния, легированных разными примесями. В рекомбинационных переходах принимает участие поперечный акустический фонон TA с энергией 0,016 эВ и поперечный оптический фонон TO с энергией 0,055 эВ, которые обеспечивают сохранение квазиимпульса носителя заряда. На рисунке длинная вертикальная черта соответствует энергии $h\nu = E_g - E_n - E_p (TO)$, короткая —

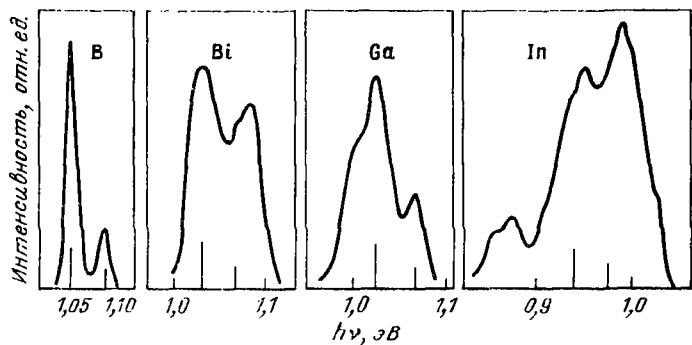


Рис. 12-13. Спектральное распределение примесной излучательной рекомбинации в Si при 20 К.

энергия $h\nu = E_g - E_n - E_p (TA)$. Значения энергий ионизации для примесей (в электронвольтах) составляют: 0,046 (B); 0,069 (Bi); 0,071 (Ga) и 0,16 (In). У образца, легированного индием, в низкоэнергетической части спектра наблюдаются бесфононные переходы.

2. Донорно-акцепторные пары

Если в полупроводнике присутствуют донорная и акцепторная примеси с энергией ионизации соответственно E_d и E_a , то они могут образовывать пары, которые будут действовать как неподвижные молекулы, погруженные в кристалл. Кулоновское взаимодействие между донором и акцептором, находящимися в среде с диэлектрической проницаемостью ϵ , вызывает уменьшение их энергии связи на величину

$$\Delta E = \frac{e^2}{\epsilon r}, \quad (12-9)$$

где r — расстояние между донором и акцептором, образующими пары. Поэтому энергетический зазор между донорным и акцепторным уровнями (рис. 12-14) будет равен:

$$h\nu = E_g - E_d - E_a + \frac{e^2}{\epsilon r}. \quad (12-10)$$

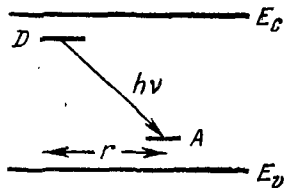


Рис. 12-14. Донорно-акцепторные переходы.

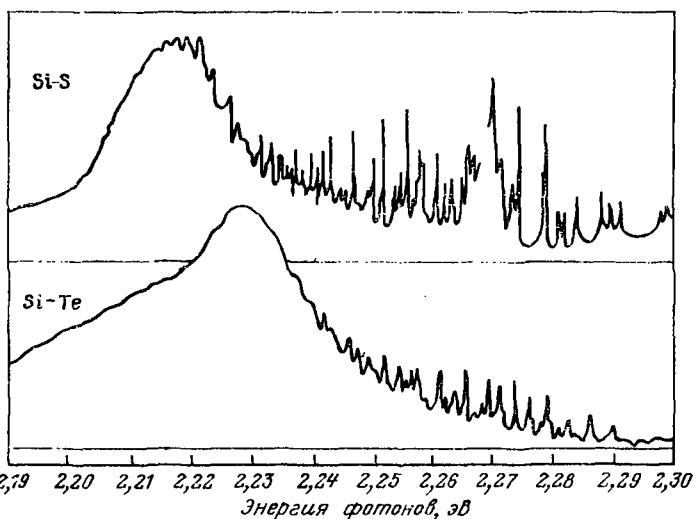


Рис. 12-15. Спектры фотолюминесценции кристаллов GaP при 1,6 К, соответствующие донорно-акцепторной рекомбинации в парах Si—S и Si—Te.

Поскольку примеси могут занимать дискретные положения в решетке (например, примесь замещения занимает узлы), расстояние в донорно-акцепторной паре меняется конечными порциями. Это должно обусловить наличие в спектре тонкой структуры. При боль-

ших расстояниях между донором и акцептором ($r > 40 \text{ \AA}$) линии излучения перекрываются, образуя широкий спектр. Спектральное положение как широкой полосы, так и структуры узких линий должно зависеть от типа примесей. При этом в зависимости от мест, занимаемых примесями, могут образовываться различные типы пар. Если донор и акцептор занимают в решетке одинаковые места, то образуются пары типа I. Если донор и акцептор занимают различные места, то возникают пары типа II.

Излучение донорно-акцепторных пар было обнаружено в ряде полупроводников. В частности, на рис. 12-15 дан спектр фосфида галлия при 1,6 К, имеющего донорно-акцепторные пары Si-S и Si-Te, в которых Si является акцептором. И доноры и акцепторы занимают узлы вместо фосфора.

12-5. РЕЛАКСАЦИЯ ЛЮМИНЕСЦЕНЦИИ ПОЛУПРОВОДНИКОВ

При выключении источника возбуждения излучение полупроводника исчезает не мгновенно. Поскольку интенсивность люминесценции пропорциональна числу актов рекомбинации в секунду

$$I_n \sim - \left(\frac{dn}{dt} \right)_r, \quad (12-11)$$

то изменение ее яркости во времени после снятия возбуждения будет различной в зависимости от вида люминесценции.

В случае мономолекулярного свечения, когда процесс поглощения и излучения происходит в одном и том же центре, уменьшение концентрации возбужденных электронов во времени определяется соотношением вида

$$n = n_0 e^{-t/\tau}, \quad (12-12)$$

где n_0 — концентрация возбужденных электронов в момент $t = 0$; τ — время жизни электрона на возбужденном уровне центра. Следовательно, *интенсивность мономолекулярной люминесценции после прекращения возбуждения*

$$I_n \sim e^{-t/\tau}, \quad (12-13)$$

т. е. убывает по экспоненциальному закону с постоянной времени τ .

Для рекомбинационной люминесценции закон затухания определяется степенью ее возбуждения. Рассмотрим два крайних случая. При слабом возбуждении, когда концентрация избыточных носителей заряда Δn значительно меньше равновесных $\Delta n \ll n_0 + p_0$, интенсивность рекомбинации согласно (7-18) пропорциональна первой степени концентрации, т. е.

$$\frac{d\Delta n}{dt} = - \frac{\Delta n}{\tau}. \quad (12-14)$$

Для этого случая $\Delta n = \Delta n_0 e^{-t/\tau}$, поэтому

$$I_n = I_0 e^{-t/\tau}, \quad (12-15)$$

где Δn_0 и I_0 — концентрация избыточных носителей заряда и интенсивность люминесценции в момент времени $t = 0$, т. е. в момент выключения возбуждающего света. Вид излучения полупроводника, интенсивность которого описывается равенством (12-15), носит название мономолекулярного рекомбинационного свечения. По его экспоненциальной кривой релаксации можно определить время жизни τ неравновесных носителей заряда.

В случае сильного возбуждения, когда $\Delta n \gg n_0 + p_0$, интенсивность рекомбинации пропорциональна квадрату концентрации избыточных носителей заряда [равенство (7-20)], т. е.

$$\frac{d\Delta n}{dt} = -\gamma_r \Delta n^2. \quad (12-16)$$

Поскольку концентрация избыточных носителей убывает во времени как

$$\Delta n = \frac{\gamma_r \Delta n_0}{1 + \gamma_r \Delta n_0 t}, \quad (12-17)$$

то интенсивность люминесценции описывается соотношением вида

$$I_n = \frac{\gamma_r^2 \Delta n_0^2}{(1 + \gamma_r \Delta n_0 t)^2}, \quad (12-18)$$

а люминесценция называется бимолекулярным рекомбинационным свечением.

12-6. ТЕМПЕРАТУРНОЕ ТУШЕНИЕ ЛЮМИНЕСЦЕНЦИИ ПОЛУПРОВОДНИКОВ

У полупроводников интенсивность излучения уменьшается при повышении температуры, т. е. наблюдается температурное тушение люминесценции. Если через P_r обозначить вероятность излучательных переходов, которая предполагается не зависящей от температуры, а через P_{nr} — вероятность безызлучательных переходов, то квантовый выход люминесценции η описывается выражением

$$\eta = \frac{P_r}{P_r + P_{nr}}. \quad (12-19)$$

Вероятность P_{nr} зависит от температуры

В этом выражении E^* — энергия активации; P_{nr0} — коэффициент, не зависящий от температуры. Следовательно, температурная зависимость квантового выхода излучения имеет вид:

$$\eta = \frac{1}{1 + Ce^{-E^*/kT}}, \quad (12-21)$$

где $C = P_{nr0}/P_r$ — константа.

Тушение мономолекулярного свечения твердых тел, когда поглощение и излучение света происходят в одном центре, можно объяснить следующим образом. При поглощении кванта света электрон центра, переходя на возбужденный уровень, может при высокой температуре попасть на колебательный энергетический уровень, соответствующий ближайшему расположению потенциальных кривых U_0 и U_v (точка a на рис. 12-1). Затем он безызлучательным способом перейдет в основное состояние. Этот переход на рис. 12-1 показан пунктирной стрелкой. Для такого механизма температурного тушения выход люминесценции будет описываться формулой (12-21). В этом случае E^* есть энергия активации, необходимая для перевода электрона в возбужденном состоянии в точку ближайшего расположения потенциальных кривых U_0 и U_v .

Температурное тушение люминесценции полупроводника, обусловленной рекомбинацией свободных носителей заряда через локализованный центр, также будет описываться соотношением (12-21). Например, центр свечения расположен на некотором расстоянии от потолка валентной зоны (рис. 12-11, б). Такой центр, поглощая квант света, может возбуждаться (см. рис. 11-31, з). Затем 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. Например, безызлучательные переходы при участии меди в германии имеют термическую энергию активации $E^* = 0,14$ эВ.

В случае излучательной рекомбинации свободных электронов и свободных дырок температурное тушение люминесценции можно описать формулой

$$\eta = \frac{1}{1 + CT^g}, \quad (12-22)$$

где C и g — константы. При выводе этого соотношения использовалась теория прямой безызлучательной рекомбинации электрона зоны проводимости и дырки валентной зоны.

12-7. СПОНТАННОЕ И ВЫНУЖДЕННОЕ ИЗЛУЧЕНИЕ АТОМА

Прежде чем переходить к описанию условий возникновения лазерного излучения в полупроводниках, рассмотрим процесс поглощения и излучения света в атоме. Из всей совокупности кван-

товых состояний, в которых может находиться атом, для простоты рассмотрения выберем только два состояния с энергией E_1 и E_2 (рис. 12-16).

Пусть на такой атом падает излучение, имеющее спектральную плотность энергии u_ν . Невозбужденный атом, поглощая энергию излучения $h\nu$, перейдет в возбужденное состояние (рис. 12-16, а). При этом переходы с поглощением совершаются тем чаще, чем больше величина u_ν . Если N_1 — число атомов в невозбужденном состоянии, а $u_\nu B_{12}$ — вероятность перехода с поглощением, то число переходов с поглощением в единицу времени в единице объема равно:

$$dN_{12} = B_{12}u_\nu N_1. \quad (12-23)$$

Атом, находясь в возбужденном состоянии в результате поглощения кванта света, через некоторый промежуток времени из возбужденного состояния возвратится в нормальное, отдавая избыток энергии $E_2 - E_1$ в виде излучения $h\nu$. Переход с излучением может произойти самопроизвольно, независимо от воздействия внешнего поля (рис. 12-16, б). Такой переход называется спонтанным, а свечение носит название спонтанного излучения.

Спонтанные переходы являются случайными, их вероятность обозначим через A_{21} . Если N_2 — число атомов, находящихся в возбужденном состоянии, то число спонтанных переходов в единицу времени

$$dN'_{21} = A_{21}N_2. \quad (12-24)$$

Следовательно, уменьшение количества возбужденных атомов в результате спонтанных переходов за время dt составит величину

$$-dN_2 = A_{21}N_2 dt. \quad (12-25)$$

Интегрируя это уравнение, получаем закон убывания числа возбужденных атомов со временем при спонтанных переходах:

$$N_2 = N_{20}e^{-A_{21}t}, \quad (12-26)$$

где N_{20} — число атомов, находящихся в возбужденном состоянии в момент времени $t = 0$.

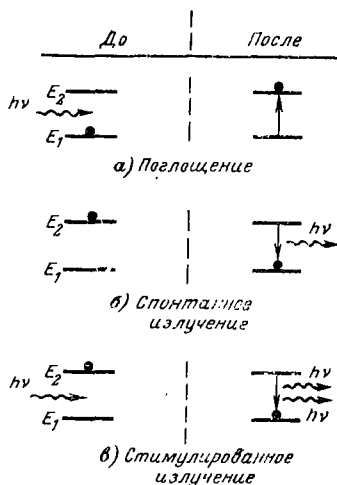


Рис. 12-16. Три основных типа переходов между двумя энергетическими уровнями E_1 и E_2 . Черные точки обозначают состояние атома. Начальному состоянию соответствует рисунок слева; конечное состояние после протекания соответствующего процесса изображено слева.

Так как за каждый спонтанный переход излучается энергия $E_2 - E_1 = h\nu$, то в результате dN'_{21} переходов в единицу времени выделяется энергия в количестве

$$h\nu dN'_{21} = h\nu A_{21} N_2. \quad (12-27)$$

Тогда интенсивность спонтанного свечения (энергия, испускаемая в единицу времени) равна:

$$I = h\nu A_{21} N_{20} e^{-A_{21}t} = I_0 e^{-A_{21}t}, \quad (12-28)$$

где

$$I_0 = h\nu A_{21} N_{20}. \quad (12-29)$$

Так как $A_{21} N_2 dt$ есть число атомов, совершивших за время dt переход $2 \rightarrow 1$, то оно определяет количество атомов, «проживших» в возбужденном состоянии в течение времени t . Поэтому продолжительность пребывания в возбужденном состоянии всех атомов, испытавших спонтанные переходы в течение времени от 0 до ∞ , составит величину

$$T = \int_0^{\infty} t A_{21} N_2 dt = \int_0^{\infty} N_{20} A_{21} t e^{-A_{21}t} dt = \frac{N_{20}}{A_{21}}. \quad (12-30)$$

Отсюда следует, что средняя продолжительность жизни атома в возбужденном состоянии

$$\tau = \frac{T}{N_{20}} = \frac{1}{A_{21}}. \quad (12-31)$$

В соответствии с этим формула (12-28) запишется в виде

$$I = I_0 e^{-t/\tau}. \quad (12-32)$$

Таким образом, после прекращения возбуждения интенсивность спонтанного излучения во времени уменьшается по экспоненциальному закону с постоянной времени τ .

Поскольку спонтанные переходы являются случайными, то спонтанное излучение имеет статистический характер, оно некогерентно, т. е. генерируются фотоны, распространяющиеся в произвольных направлениях и имеющие произвольные фазы, хотя их частота фиксирована. Это обусловлено тем, что акты спонтанного излучения в силу случайного характера происходят независимо друг от друга в разные моменты времени.

Однако атом, находящийся в возбужденном состоянии (на уровне E_2), под действием поля излучения может возвратиться в нормальное состояние (на уровень E_1). Такой переход называется вынужденным, а излучение носит название вынужденного, индуцированного или стимулированного (рис. 12-16, в). Индуцированные переходы будут совершаться тем чаще, чем больше плотность энергии возбуждающего света u_ν . Поэтому вероятность этих переходов в единицу времени

будет $u_\nu B_{21}$. И если N_2 — число атомов в возбужденном состоянии, то число индуцированных переходов в единицу времени в единице объема составит:

$$dN_{21}'' = u_\nu B_{21} N_2. \quad (12-33)$$

Следовательно, внешнее излучение с частотой ν вызывает не только поглощение света, но и добавочное, индуцированное, излучение.

Таким образом, среднее число переходов с излучением в единицу времени в единице объема определяется спонтанными и вынужденными переходами и равно:

$$dN_{21} = dN_{21}' + dN_{21}'' = (A_{21} + u_\nu B_{21}) N_2. \quad (12-34)$$

В стационарном состоянии на основании принципа детального равновесия число прямых и обратных процессов в единицу времени одинаково. Следовательно,

$$dN_{12} = dN_{21} \quad (12-35)$$

или

$$(A_{21} + u_\nu B_{21}) N_2 = u_\nu B_{12} N_1. \quad (12-36)$$

Отсюда найдем, что

$$\frac{N_2}{N_1} = \frac{u_\nu B_{12}}{A_{21} + u_\nu B_{21}}. \quad (12-37)$$

Поскольку равновесие имеет статистический характер, то отношение N_2/N_1 можно выразить при помощи распределения Больцмана, по которому для системы, имеющей состояния с энергией $E_1, E_2, \dots, E_m, \dots$, вероятность пребывания ее в энергетическом состоянии E_m равна:

$$W_m = C g_m e^{-E_m/kT}, \quad (12-38)$$

где g_m — статистический вес, учитывающий вырождение состояния m . Поэтому можно записать:

$$\frac{N_2}{N_1} = \frac{g_2 e^{-E_2/kT}}{g_1 e^{-E_1/kT}} = \frac{g_2}{g_1} e^{-(E_2 - E_1)/kT}. \quad (12-39)$$

Из сравнения равенств (12-37) и (12-39) следует, что

$$\frac{u_\nu B_{12}}{A_{21} + u_\nu B_{21}} = \frac{g_2}{g_1} e^{-(E_2 - E_1)/kT}, \quad (12-40)$$

откуда с учетом $E_2 - E_1 = h\nu$ получим:

$$u_\nu = \frac{g_2 A_{21}}{g_1 B_{12} e^{h\nu/kT} - g_2 B_{21}}. \quad (12-41)$$

Для определения коэффициентов A и B используем следующие предельные условия. При $T \rightarrow \infty$ плотность энергии u_ν должна

стремиться к бесконечности, а это значит, что знаменатель в выражении (12-41) стремится к нулю. Из этого следует, что

$$g_1 B_{12} = g_2 B_{21}. \quad (12-42)$$

С учетом этого формула (12-41) запишется в виде

$$u_\nu = \frac{g_2 A_{21}}{g_1 B_{12}} \frac{1}{e^{h\nu/kT} - 1}. \quad (12-43)$$

При $\frac{h\nu}{kT} \ll 1$ в первом приближении имеем $e^{h\nu/kT} = 1 + \frac{h\nu}{kT}$. Следовательно,

$$u_\nu = \frac{g_2 A_{21}}{g_1 B_{12}} \frac{kT}{h\nu}. \quad (12-44)$$

С другой стороны, при условии $\frac{h\nu}{kT} \ll 1$ справедлива формула Рэлея—Джинса*, по которой

$$u_\nu = \frac{8\pi\nu^2}{c^3} kT. \quad (12-45)$$

Из сопоставления формул (12-44) и (12-45) имеем:

$$\frac{g_2 A_{21}}{g_1 B_{12}} = \frac{8\pi h\nu^3}{c^3}, \quad (12-46)$$

откуда

$$A_{21} = \frac{8\pi h\nu^3}{c^3} \frac{g_1}{g_2} B_{12} = \frac{8\pi h\nu^3}{c^3} B_{21}. \quad (12-47)$$

Таким образом, при вынужденном переходе генерируется фотон, имеющий ту же частоту, направление и фазу, что и фотон, индуцирующий рекомбинацию.

Рассмотрим теперь, при каких условиях может наблюдаться индуцированное излучение. Пусть имеем параллельный пучок монохроматического света интенсивности I . При прохождении слоя вещества толщиной Δx интенсивность света уменьшится на величину $-\Delta I = \alpha I \Delta x$. Это уменьшение интенсивности света определяется тем, что в атомах происходят переходы с поглощением энергии в количестве $u_\nu B_{12} h\nu N_1 \Delta x$ и индуцированные переходы с излучением энергии, равной $u_\nu B_{21} h\nu N_2 \Delta x$. В результате можно записать:

$$-\Delta I = \alpha I \Delta x = (B_{12} N_1 - B_{21} N_2) u_\nu h\nu \Delta x. \quad (12-48)$$

Из формулы (12-47) 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.

* Шпольский Э. В. Атомная физика. М., «Наука», 1974, т. 1, § 86.

Если учесть, что плотность излучения u_ν в параллельном пучке света сечением 1 см^2 , распространяющегося со скоростью c , равна:

$$u_\nu = \frac{I}{c}, \quad (12-49)$$

то коэффициент поглощения α в соответствии с равенствами (12-48), (12-49) и (12-42) приобретает вид:

$$\alpha = \frac{h\nu}{c} (B_{12}N_1 - B_{21}N_2) = \frac{h\nu}{c} \left(1 - \frac{g_1}{g_2} \frac{N_2}{N_1}\right) B_{12}N_1. \quad (12-50)$$

Учитывая (12-39), имеем:

$$\alpha = B_{12}N_1 \left(1 - e^{-(E_2 - E_1) / kT}\right) \frac{h\nu}{c}. \quad (12-51)$$

Так как $E_2 > E_1$, то экспоненциальный член в (12-51) всегда меньше единицы, и величина, стоящая в скобках, положительна. Это означает, что в случае термодинамического равновесия при любой конечной температуре суммарный коэффициент поглощения всегда положителен, т. е. $\alpha > 0$. В этом случае имеет место обычное положительное поглощение света, а индуцированные переходы лишь уменьшают его. Причиной этому является то, что в равновесном состоянии число невозбужденных атомов всегда больше числа возбужденных атомов, т. е. $N_1 > N_2$.

Из распределения Больцмана (12-39) следует, что

$$T = - \frac{E_2 - E_1}{k \ln \frac{g_1 N_2}{g_2 N_1}}. \quad (12-52)$$

Поскольку $E_2 > E_1$, а в равновесии $g_1 N_2 < g_2 N_1$, то $T > 0$. Однако если выполняется неравенство

$$g_1 N_2 > g_2 N_1 \quad \text{или} \quad \frac{N_2}{N_1} > \frac{g_2}{g_1}, \quad (12-53)$$

то $T < 0$. Таким образом, для того чтобы создать среду с отрицательным поглощением, необходимо осуществить неравновесное состояние, при котором число возбужденных атомов было бы больше числа атомов, находящихся в нормальном, невозбужденном состоянии. Заселенность энергетических уровней атомов, удовлетворяющих неравенству (12-53), носит название *инверсионности*.

12-8. СТИМУЛИРОВАННОЕ ИЗЛУЧЕНИЕ ТВЕРДЫХ ТЕЛ

Одним из способов получения среды с отрицательным поглощением может явиться создание косвенным путем избытка атомов на более высоком уровне энергии по отношению к более низкому.

Такой случай можно осуществить на атомной или молекулярной системе, которая может находиться в трех энергетических состояниях: в нормальном с энергией E_0 и двух возбужденных (1 и 2) с энергиями E_1 и E_2 , между которыми возможны спонтанные переходы с вероятностями A_{20} , A_{21} и A_{10} (рис. 12-17). Будем считать, что эти уровни не имеют вырождения.

В такой системе с помощью поглощаемых квантов света производится заселение возбужденного уровня 2 . Селективное заселение уровня получило название оптической накачки. Уровень 1 заселяется только в результате спонтанного перехода $2 \rightarrow 1$, и число таких переходов в единицу времени равно $A_{21}N_2$.

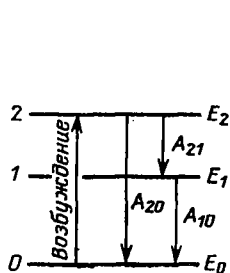


Рис. 12-17. Трехуровневая схема переходов.

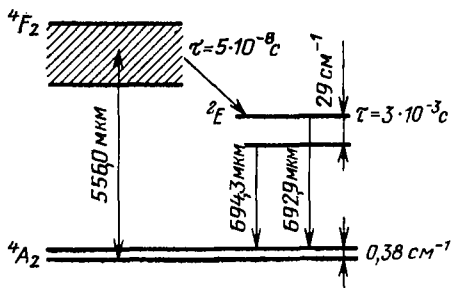


Рис. 12-18. Схема уровней в кристалле рубина.

Уменьшение населенности уровня 1 происходит только за счет переходов $1 \rightarrow 0$, и число таких переходов в единицу времени будет $A_{10}N_1$. В стационарных условиях количество заселяющих переходов равно числу обратных переходов:

$$A_{21}N_2 = A_{10}N_1. \quad (12-54)$$

Для получения индуцированного излучения необходима инверсная населенность уровней, т. е. населенность уровня 2 должна быть больше, чем уровня 1 . Для рассматриваемой системы инверсная населенность уровней возникнет в том случае, если

$$A_{21} < A_{10}. \quad (12-55)$$

Из этого соотношения следует, что уровень 2 за счет перехода $2 \rightarrow 1$ должен опустошаться медленнее, чем уровень 1 за счет переходов $1 \rightarrow 0$. При таком соотношении вероятностей в состоянии 2 будет накапливаться больше атомов, чем в состоянии 1 . Однако для осуществления такого процесса накопления атомов необходимо также, чтобы уровень 2 достаточно медленно опустошался за счет переходов $2 \rightarrow 0$, т. е. вероятность перехода из состояния 2 в состояние 0 должна быть мала.

Генераторы стимулированного излучения 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), содержащий в качестве примеси замещения трехвалентные ионы хрома Cr^{3+} . У трехвалентного иона хрома, энергетическая схема которого изображена на рис. 12-18, имеется энергетическая полоса 4F_2 , а нормальное 4A_2 и возбужденное 2E состояния представляют собой узкие дублеты.

Переходу ${}^4A_2 \rightarrow {}^4F_2$ соответствует полоса поглощения с $\lambda = 556$ мкм. Для состояния ${}^4F_2 \rightarrow {}^2E$ среднее время жизни $\tau = 5 \cdot 10^{-8}$ с, поэтому вероятность перехода ${}^4F_2 \rightarrow {}^2E$ велика, а для перехода ${}^2E \rightarrow {}^4A_2$ мала, так как среднее время жизни для состояния 2E примерно $5 \cdot 10^{-3}$ с.

При помощи оптической накачки производится инверсная населенность метастабильного уровня 2E по отношению к уровню 4A_2 . Переходы ${}^2E \rightarrow {}^4A_2$ сопровождаются двумя линиями излучения с $\lambda = 694,3$ мкм и $\lambda = 692,9$ мкм. В спонтанном свечении эти линии слабые, так как вероятность соответствующих переходов мала.

Для создания излучения за счет вынужденных переходов кристалл рубина используется в виде стерженька длиной \mathcal{L} в несколько сантиметров со строго параллельными и оптически отполированными торцами. Торцы стержня покрываются тонким слоем серебра. При этом один из них делается полупрозрачным.

Возникшее спонтанное излучение за счет перехода ${}^2E \rightarrow {}^4A_2$ в среде с отрицательным коэффициентом поглощения усиливается индуцированными переходами. Его интенсивность при прохождении длины стержня \mathcal{L} возрастает по закону

$$I = I_0 e^{\alpha \mathcal{L}}. \quad (12-56)$$

Увеличение интенсивности излучения будет происходить при многократном прохождении световой волны вдоль кристалла в результате отражения излучения от торцов кристалла. Многократно отражаться будут лишь те лучи, которые падают на торцевые поверхности кристалла перпендикулярно. В силу этого излучение, вызванное индуцированными переходами, сосредоточено в пределах малого телесного угла $\Delta\Omega$. В результате возникает узконаправленный пучок излучения большой интенсивности, так как яркость его возрастает в отношении $4\pi/\Delta\Omega$. Спонтанное же излучение распределяется равномерно во всех направлениях в пределах телесного угла 4π , поэтому оно уходит из кристалла через боковые поверхности.

Индукцированное излучение является когерентным. Оно происходит в том же направлении и имеет ту же фазу, что и первичное. Поэтому лазер можно рассматривать как «резонатор», внутри которого возникнут стоячие волны. Стоячие волны возникают в том

случае, если на длине стержня \mathcal{L} укладывается целое число полуволн излучения

$$m \frac{\lambda}{2n} = \mathcal{L}, \quad (12-57)$$

где m — целое число; λ/n — длина волны излучения в рубине.

Оптическая накачка в рубиновом лазере может осуществляться с помощью лампы-вспышки (ксеноновой). В этом случае накопленная энергия в возбужденном состоянии 2E во время импульсной накачки высвечивается в виде вспышки. При работе рубинового лазера в непрерывном режиме требуются большая мощность оптической накачки и охлаждение кристалла. Возможно и уменьшение времени лазерной вспышки. Это достигается при помощи зеркала, которое отсутствует во время накачки, а затем быстро вводится в систему. В результате возникает лавинообразный процесс индуцированного излучения.

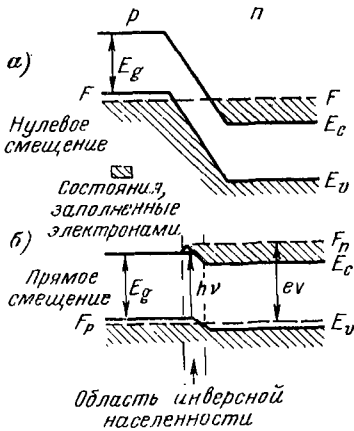


Рис. 12-19. Сильно легированный p - n переход в равновесном состоянии (а) и при смещении в пропускном направлении (б).

Создать среду с отрицательным поглощением также возможно, если искусственно понизить населенность более низкого уровня по отношению к более высокому. Такой механизм возбуждения индуцированного свечения можно осуществить в полупроводниках, у которых экстремумы зон имеют одинаковые значения k , например InSb , InAs , InP , GaAs . У них имеет место преимущественно излучательная рекомбинация. В этом случае при прохождении тока через p - n переход в пропускном направлении рекомбинация инжектированных неосновных носителей заряда сопровождается свечением. Это свечение является спонтанным.

На излучающих диодах можно получить и стимулированное излучение, но для этого необходимо создать инверсную населенность уровней, при которой верхние уровни должны быть населены больше чем наполовину по отношению к нижним. Следовательно, в случае p - n перехода носители заряда должны находиться в вырожденном состоянии. Диаграмма энергетических зон инжекционного лазера, которая представлена на рис. 12-19, подобна диаграмме туннельного диода. Но степень легирования областей лазера ниже, чем туннельного диода, поэтому на вольт-амперной характеристике лазерного диода не имеется области отрицательного сопротивления или эта область пренебрежимо мала. В условиях равновесия квазиуровни Ферми лежат в зоне проводимости и валентной зоне

(рис. 12-19, а). У сильно легированного p - n перехода достаточно высокое смещение в пропускном направлении может создать область инверсной населенности (рис. 12-19, б), которая обычно очень тонкая (около 1 мкм). В этом случае электроны могут перемещаться через барьер n -области и переходить на свободные состояния валентной зоны с испусканием фотонов, энергия которых примерно равна E_g . Переходы могут также происходить между краем зоны проводимости и акцепторным уровнем (или зоной) вблизи потолка валентной зоны. Кроме того, дырки могут переходить в n -область и рекомбинировать с электронами с испусканием фотонов. При этом вид преобладающего процесса определяется относительной концентрацией примеси, значениями подвижности и времени жизни неравновесных носителей заряда.

При малых токах, текущих через p - n переход, возникает спонтанное излучение, распространяющееся по всем направлениям. Возрастание тока увеличивает скорость рекомбинации, что соответственно увеличивает плотность фотонов. Они индуцируют дальнейшую рекомбинацию. Поскольку на максимум спектрального распределения спонтанного излучения приходится больше фотонов, на этой частоте вызывается больше переходов, чем на краях спектра. Это приводит к значительному сужению спектра излучения и, следовательно, усилению интенсивности излучения в его максимуме. Если интенсивность пика излучения растет сверхлинейно с уровнем возбуждения, то процесс называется «сверхсвечением». Возникающие при сверхсвечении фотоны, так же как и при спонтанном излучении, распространяются во всех направлениях с произвольными фазами. Для создания лазера необходимо, чтобы усиление было по крайней мере равно потерям, а излучение — когерентным. Когерентность достигается помещением источника в резонатор, обеспечивающий селективное усиление электромагнитных колебаний определенной частоты и фазы, которые образуют стоячие волны в резонаторе. Степень когерентности определяется качеством резонатора.

Типичный характер зависимости интенсивности излучения лазера от уровня возбуждения представлен на рис. 12-20. Как видно, ниже некоторого порогового уровня возбуждения излучение в основном имеет спонтанный характер (кривая 1). При возрастании тока усиление также возрастает. Но когда достигнут порог генерации, т. е. когда световая волна проскакивает резонатор без затухания, интенсивность вынужденного излучения резко увеличивается (кривая 2) и из окружающих плоскостей p - n перехода будет выходить когерентное излучение.

На рис. 12-21 изображен спектр излучения типичного лазерного диода. При значениях тока ниже пороговых спектр излучения определяется спонтанными переходами и относительно широк. При возрастании тока выше пороговых значений спектр излучения значительно сужается. Зависимость спектрального излучения этого лазера от температуры представлена на рис. 12-22. Как видно, при

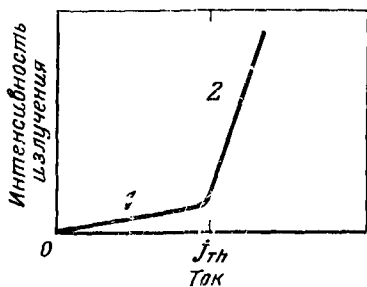


Рис. 12-20. Излучаемая мощность полупроводникового лазера в зависимости от уровня возбуждения.

1 -- спонтанное излучение; 2 — вынужденное излучение.

Рис. 12-21. Спектральное распределение излучения GaAs — лазера при токах ниже и выше пороговых.

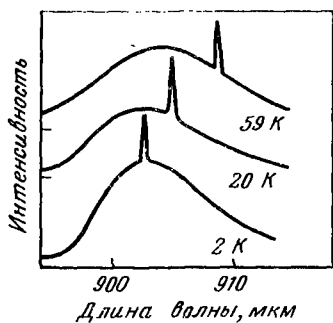
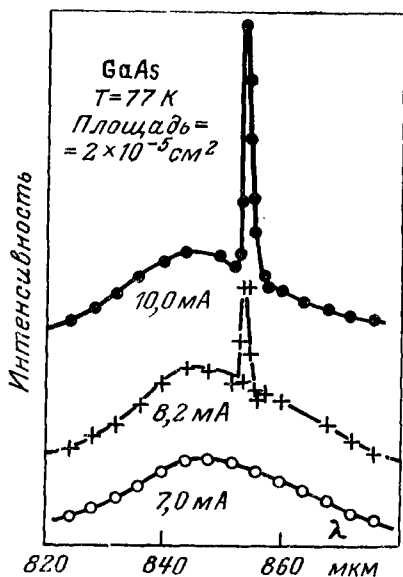


Рис. 12-22. Зависимость спектрального излучения GaAs лазерного диода от температуры.

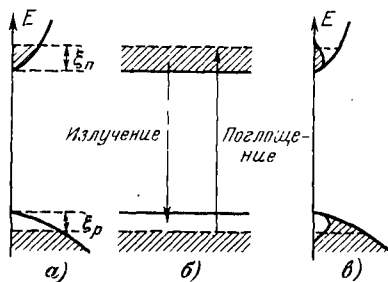


Рис. 12-23. Заполнение зон до вырождения, предотвращающее межзонное поглощение на частоте излучения.

возрастании температуры линия вынужденного излучения сдвигается в область более длинных волн, что согласуется с температурным изменением ширины запрещенной зоны.

Инверсная населенность в полупроводниках может быть осуществлена также фотовозбуждением и возбуждением с помощью пучка электронов. При оптическом возбуждении инверсия осуществляется фотонами, обладающими энергией, превышающей ширину запрещенной зоны. Пусть при абсолютном нуле за счет фотовозбуждения состояния зоны проводимости полупроводника с пря-

мыми оптическими переходами заполнены до квазиуровня Ферми ξ_n и введены дырки в валентную зону до квазиуровня Ферми ξ_p , как это изображено на рис. 12-23, а. При такой инверсной населенности ни одно состояние, участвующее в процессе излучения, не может участвовать в поглощении (рис. 12-23, б). При конечных температурах распределение носителей заряда будет такое, как показано на рис. 12-23, в. Хотя для всей системы тепловое равновесие отсутствует, но носители заряда в данной энергетической зоне находятся в тепловом равновесии друг с другом. Фотоны, имеющие энергию $h\nu$ [$E_g < h\nu < (F_n - F_p)$], вызывают переходы электронов на нижние уровни, что приводит в свою очередь к возникновению стимулированного излучения.

Получение вырожденной электронно-дырочной населенности в полупроводниках можно достичь за счет перехода электронов из валентной зоны в зону проводимости при бомбардировке кристалла пучком электронов. Преимуществом оптического возбуждения и возбуждения с помощью электронного пучка является то, что в этих случаях возбуждается большой объем материала и не требуется создание p - n перехода. Это особенно важно для материалов, в которых трудно получить p - n переход.

Глава тринадцатая

ФОТОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

13-1. ВНУТРЕННИЙ ФОТОЭФФЕКТ

Изменение электрического сопротивления полупроводника под действием излучения называется внутренним фотоэлектрическим эффектом или фоторезистивным эффектом. Добавочная проводимость, обусловленная носителями заряда, созданными излучением, носит название фотопроводимости.

При внутреннем фотоэффекте первичным актом является поглощение фотона. Поэтому процесс образования свободных носителей заряда под воздействием излучения будет происходить по-разному в зависимости от особенностей процесса поглощения света. Если оптическое возбуждение электронов происходит из валентной зоны в зону проводимости, то имеет место с о б с т в е н н а я ф о т о п р о в о д и м о с т ь, обусловленная электронами и дырками (переход 1, рис. 13-1). Для полупроводников с прямыми долинами при вертикальных переходах энергия фотона $h\nu$ должна быть не меньше ширины запрещенной зоны, т. е.

$$h\nu \geq E_g. \quad (13-1)$$

В случае не прямых переходов, когда сохранение квазиимпульса обеспечивается за счет эмиссии фонона, нижняя граница спектрального распределения фотопроводимости будет лежать при

$$h\nu = E_g + E_p. \quad (13-2)$$

Для сильно легированного полупроводника n -типа, когда уровень Ферми расположен выше края зоны проводимости на величину ξ_n , нижняя граница фотопроводимости будет соответствовать

$$h\nu = E_g + \xi_n. \quad (13-3)$$

В сильно легированном полупроводнике p -типа уровень Ферми лежит на величину ξ_p ниже края валентной зоны, поэтому

$$h\nu = E_g + \xi_p. \quad (13-4)$$

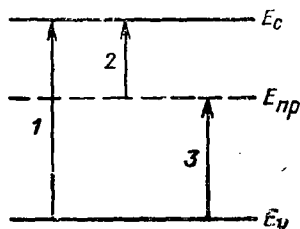


Рис. 13-1. Схема возможных оптических переходов, обуславливающих фотопроводимость.

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 сплошную зону. Поэтому спектральное распределение фоточувствительности в зависимости от энергии фотонов или длины волны света должно простираться далеко в коротковолновую область. При

больших энергиях фотонов поглощение в фундаментальной области ведет к увеличению фотопроводимости за счет роста коэффициента поглощения. Вместе с тем такое поглощение увеличивает концентрацию носителей заряда вблизи поверхности полупроводника, которые имеют меньшее время жизни, чем носители заряда в объеме. Эти процессы приведут к тому, что спектральная кривая фототока должна проходить через максимум. Очевидно, чем больше скорость поверхностной рекомбинации, тем сильнее будет выражен максимум фототока. Это действительно имеет место, как это видно из рис. 13-2, на котором представлены кривые спектрального распределения фототока для некоторых полупроводников.

При наличии в запрещенной зоне полупроводника локальных уровней примеси оптическое поглощение может вызвать переходы электронов между уровнями примеси и зонами (переходы 2 и 3, рис. 13-1). Фотопроводимость, обусловленная такими переходами, называется примесной фотопроводимостью. Поскольку энергия ионизации примеси $E_{пр}$ меньше ширины запрещенной зоны E_g , то длинноволновая граница примесного фототока сдвинута в длинноволновую сторону по отношению к собственной фотопроводимости. Но спектральные кривые примесного внутреннего фотоэффекта $\sigma_{ф}(\lambda)$ достаточно хорошо совпадают с кривыми

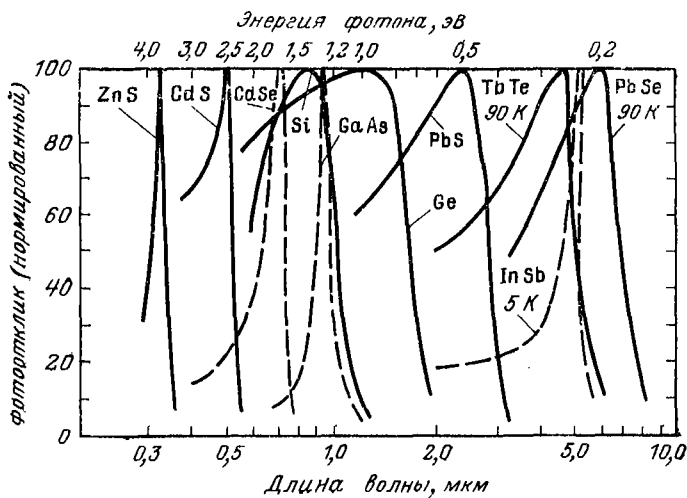
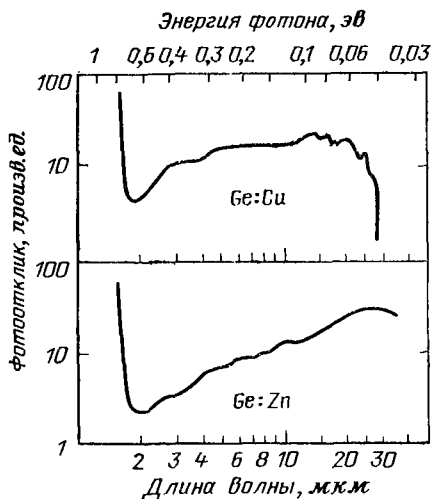


Рис. 13-2. Спектральное распределение фототока в области собственного поглощения.



← Рис. 13-3. Спектральное распределение фототока германия, легированного медью и цинком.

примесного поглощения α (λ). В качестве примера на рис. 13-3 приведен спектр фототока германия, легированного медью и цинком.

При экситонном поглощении света имеет место создание связанной пары электрон—дырка, которая является электрически нейтральным образованием. Поэтому поглощение света, связанное с образованием экситонов, первоначально не ведет к возникновению свободных носителей заряда. Однако в реальных кристаллических структурах экситоны имеют значительно большую вероятность диссоциировать безызлучательно (с образованием свободных электронов и дырок), чем рекомбинировать с испусканием кванта

света. Таким образом, образование экситонов в конечном итоге ведет к возникновению свободных носителей заряда, а следовательно, фототока. Экситонное поглощение, характеризующееся узкими полосами поглощения, определяет и узкие полосы фототока. При этом спектр фототока в области экситонного поглощения будет зависеть как от состояния поверхности, так и от соотношения диффузионной длины экситонов $L_{\text{экс}}$ и биполярной диффузионной длины L . Так, например, при обедненной поверхности, когда время жизни неосновных носителей заряда в приповерхностной области полупроводника τ_s будет меньше времени жизни τ , характерного для его объема, максимумы полос экситонного поглощения будут соответствовать минимумам на кривой фототока в случае, если $L_{\text{экс}} \ll L$ (кривая 1, рис. 13-4, а), и будут соответствовать максимумам, если $L_{\text{экс}} \gg L$ (кривая 1, рис. 13-4, б). Напротив, если $\tau_s \gg \tau$, то следует ожидать противоположных соотношений между

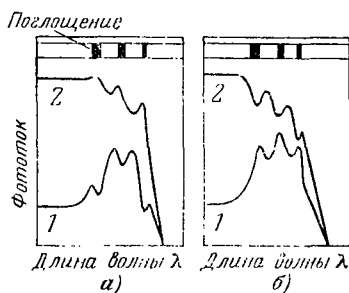
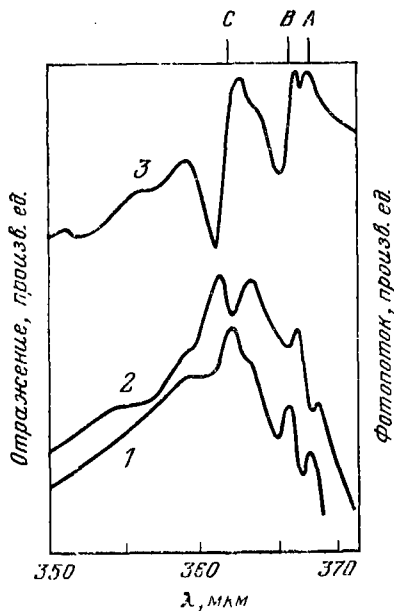


Рис. 13-4. Возможное расположение максимумов полос экситонного поглощения и фототока.

Рис. 13-5. Спектральное распределение отражения (кривая 3) и фототока монокристалла ZnO до (кривая 1) и после травления (кривая 2).



положением максимумов полос экситонного поглощения и фототока в зависимости от соотношения $L_{\text{экс}}$ и L (кривые 2 на рис. 13-4, а, б).

Состояние поверхности полупроводника можно легко изменить путем различных воздействий на нее (механическое, химическое и т. д.). Таким образом можно изменить характер наблюдаемого спектра фототока, обусловленного экситонным поглощением. Для примера на рис. 13-5 представлены спектры фототока монокристалла ZnO. Положение максимумов экситонных полос поглощения А, В и С (показано стрелками) определено из спектра отражения (кри-

вая 3). Видно, что максимумам полос экситонного поглощения соответствуют минимумы на кривой фототока (кривая 1). Травление поверхности такого кристалла приводит к случаю, когда максимумам поглощения соответствуют максимумы фототока (кривая 2). Очевидно, травление поверхности кристалла ZnO способствовало повышению диффузионной длины экситонов над биполярной диффузионной длиной.

Поглощение света свободными носителями заряда и колебаниями решетки непосредственно не могут привести к изменению концентрации носителей заряда. Однако возрастание концентрации носителей заряда в этих случаях может происходить в результате вторичных эффектов, когда поглощение света значительно увеличивает кинетическую энергию свободных носителей заряда или увеличивает концентрацию фононов, которые затем отдают свою энергию на возбуждение носителей заряда.

13-2. ФОТОПРОВОДИМОСТЬ

Избыточные электроны Δn и дырки Δp , образованные в результате взаимодействия вещества с фотонами, после процесса ионизации могут иметь энергии значительно большие, чем средняя энергия равновесных носителей заряда. Однако в результате взаимодействия с фононами и дефектами кристаллической решетки энергия неравновесных носителей заряда становится равной средней тепловой энергии равновесных носителей заряда. Этот процесс происходит за время порядка 10^{-10} — 10^{-12} с, которое равно времени релаксации носителей заряда по энергии. Как правило, время жизни неравновесных носителей заряда значительно превосходит эту величину, составляя 10^{-2} — 10^{-7} с, и, следовательно, почти в течение всего этого времени кинетическая энергия неравновесных носителей заряда соответствует средней тепловой энергии равновесных носителей заряда. Поэтому можно считать, что распределение по энергиям неравновесных носителей заряда в зонах является таким же, как и равновесных. Значит и подвижности неравновесных носителей заряда 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 заряда по энергиям.

Таким образом, при воздействии света полная проводимость полупроводника определяется равновесными носителями заряда n_0 , p_0 и носителями заряда, Δn и Δp , генерируемыми светом, поэтому равна:

$$\sigma = e [(n_0 + \Delta n) \mu_n + (p_0 + \Delta p) \mu_p]. \quad (13-5)$$

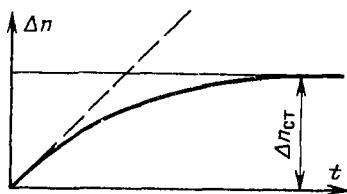
Так как темновая проводимость $\sigma_T = \sigma_0 = e (n_0 \mu_n + p_0 \mu_p)$, то фотопроводимость полупроводника, обусловленная непосредственным действием излучения, есть

$$\sigma_\Phi = \sigma - \sigma_0 = e (\Delta n \mu_n + \Delta p \mu_p). \quad (13-6)$$

Естественно, что концентрации избыточных носителей заряда Δn и Δp зависят от интенсивности и длины волны света. Пусть на слой вещества толщиной dx , имеющего коэффициент поглощения α , падает свет интенсивности I . Тогда количество световой энергии, поглощаемой в единицу времени в единице объема этого вещества есть

$$\frac{dI}{dx} = \alpha I. \quad (13-7)$$

Следовательно, при поглощении квантов света энергии $h\nu$ в единице объема полупроводника в единицу времени для области собственного поглощения образуются избыточные электроны и дырки в количестве



$$G = \frac{\beta \alpha I}{h\nu}. \quad (13-8)$$

Рис. 13-6. Изменение во времени концентрации носителей заряда, возбуждаемых светом.

Здесь коэффициент пропорциональности β , который называется квантовым выходом фотоионизации, определяет число пар носителей заряда (или число носителей заряда при примесной фотопроводимости), образуемых одним поглощенным фотоном, если интенсивность света измерять числом квантов в секунду на единицу поверхности.

Однако сразу после начала освещения фотопроводимость полупроводника не достигает максимального значения, ибо по мере увеличения концентрации неравновесных носителей заряда нарастает и процесс рекомбинации. Поскольку скорость генерации неравновесных носителей остается постоянной при неизменной интенсивности света, то через какой-то промежуток времени интенсивность рекомбинации достигнет интенсивности генерации и установится стационарное состояние, характеризующееся постоянным значением концентрации фотоносителей заряда $\Delta n_{ст}$ и $\Delta p_{ст}$ (рис. 13-6).

Стационарные концентрации избыточных носителей заряда можно определить, если воспользоваться уравнением непрерывности, в котором генерационный член записан в виде (13-8) в предположении однородной генерации. Поэтому

$$\Delta n_{ст} = G_n \tau_n = \frac{\beta \alpha I \tau_n}{h\nu}; \quad (13-9)$$

$$\Delta p_{ст} = G_p \tau_p = \frac{\beta \alpha I \tau_p}{h\nu}. \quad (13-10)$$

Они обуславливают стационарную фотопроводимость

$$\sigma_{ф. ст} = e\beta\alpha(\mu_n \tau_n + \mu_p \tau_p) \frac{I}{h\nu}. \quad (13-11)$$

Отношение фотопроводимости σ_{ϕ} к интенсивности света I , которое определяет удельную фоточувствительность полупроводника

$$S_{\phi} = \frac{\sigma_{\phi}}{I}, \quad (13-12)$$

используется для сравнения величины фотопроводимости различных полупроводниковых материалов.

Если один из членов в скобках соотношения (13-11) значительно больше другого (за счет разницы в величинах подвижности или

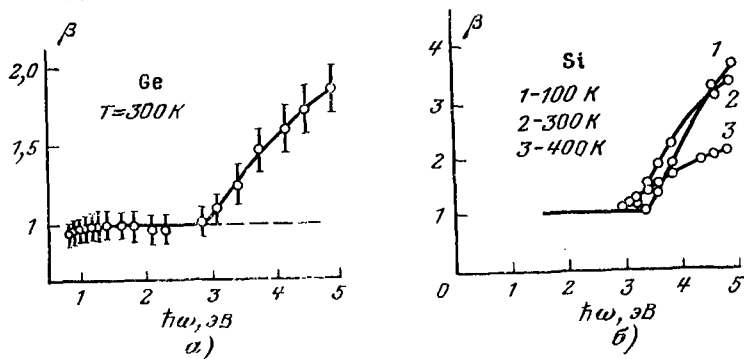


Рис. 13-7. Зависимость квантового выхода (число возбужденных электронов на один поглощенный фотон) от энергии фотона в германии (а) и кремнии (б).

времени жизни электронов и дырок), то фотопроводимость определяется носителями заряда одного знака и называется монополярной. В этом случае

$$\sigma_{\phi, \text{ст}} = e\beta\alpha\tau\mu \frac{I}{h\nu}. \quad (13-13)$$

Выражение для стационарного значения плотности фототока будет иметь вид:

$$j_{\phi} = \sigma_{\phi, \text{ст}}\mathcal{E} = e\beta\alpha(\mu_n\tau_n + \mu_p\tau_p) \frac{I\mathcal{E}}{h\nu}. \quad (13-14)$$

Но $\mu_n\mathcal{E} = v_n$, а $\mu_p\mathcal{E} = v_p$ и если через l обозначить размер полупроводника в направлении поля, то время дрейфа $t_n = l/v_n$ и $t_p = l/v_p$. С учетом этого соотношение (13-14) запишется как

$$j_{\phi} = e\beta\alpha \left(\frac{\tau_n}{t_n} + \frac{\tau_p}{t_p} \right) \frac{Il}{h\nu}. \quad (13-15)$$

Если величины, входящие в (13-15), известны, то, измеряя j_{ϕ} , можно определить квантовый выход β . На рис. 13-7, а приведена спектральная характеристика квантового выхода для германия и кремния. Видно, что при 300 К вплоть до 2,7 эВ для германия

и вплоть до 3 эВ для кремния квантовый выход равен единице. При дальнейшем росте энергии фотона квантовый выход резко увеличивается. Это происходит потому, что поглощение фотона столь большой энергии сопровождается образованием носителей заряда, кинетическая энергия которых значительно больше ширины запрещенной зоны. «Горячие» носители заряда обладают энергией, достаточной для образования вторичных электронно-дырочных пар путем ударной ионизации. Если $m_n^* = m_p^*$, то граница роста квантового выхода должна лежать при $h\nu \approx 3E_g$, а при $m_p^* \gg m_n^*$ она сдвигается в область до $h\nu \approx 2E_g$. Поскольку ширина запрещенной зоны уменьшается при повышении температуры, то граница роста квантового выхода, как видно из рис. 13-7, б, смещается в сторону меньших энергий. Следовательно, значение квантового выхода, большее единицы, определяется не рождением одним фотоном нескольких пар носителей заряда, а обусловлено ударной ионизацией.

13-3. РЕЛАКСАЦИЯ ФОТОПРОВОДИМОСТИ

Пусть полупроводник освещается импульсом света прямоугольной формы, как это изображено на рис. 13-8, а. Стационарное значение фотопроводимости достигается не мгновенно, а лишь через некоторое время после начала освещения (рис. 13-8, б). При выключении света неравновесная проводимость исчезнет также через некоторое время после прекращения освещения (рис. 13-8, б). *Кривые нарастания и спада неравновесной проводимости называются кривыми релаксации фотопроводимости.* нас будут интересовать законы нарастания и спада фотопроводимости в зависимости от уровня освещения. Рассмотрим случаи малого и большого уровня возбуждения.

Изменение концентрации избыточных носителей заряда в единицу времени $\left(\frac{d\Delta n}{dt}\right)$ есть разность между скоростями генерации и рекомбинации носителей заряда. Поскольку скорость генерации $G = \frac{\beta\alpha I}{h\nu}$, то

$$\frac{d\Delta n}{dt} = \frac{\beta\alpha I}{h\nu} - R. \quad (13-16)$$

Второй член правой части уравнения (13-16) учитывает уменьшение концентрации неосновных носителей заряда в результате процесса рекомбинации.

При малом уровне освещения, для которого $\Delta n \ll n_0 + p_0$, согласно выражению (7-18) $R = \Delta n/\tau$, поэтому

$$\frac{d\Delta n(t)}{dt} = \frac{\beta\alpha I}{h\nu} - \frac{\Delta n(t)}{\tau}. \quad (13-17)$$

Разделяя переменные и интегрируя с учетом начального условия $\Delta n = 0$ при $t = 0$, получаем выражения, характеризующие изменения концентрации избыточных электронов во времени: для процесса нарастания

$$\Delta n = \beta \alpha \tau \frac{I}{h\nu} (1 - e^{-t/\tau}) = \Delta n_{\text{ст}} (1 - e^{-t/\tau}); \quad (13-18)$$

аналогично для процесса спада

$$\Delta n = \beta \alpha \tau \frac{I}{h\nu} e^{-t/\tau} = \Delta n_{\text{ст}} e^{-t/\tau}. \quad (13-19)$$

Отсюда следует, что *релаксация фотопроводимости при малой освещенности определяется экспоненциальным законом с постоянной времени, соответствующей времени жизни неравновесных носителей заряда.*

При большом уровне возбуждения, когда $\Delta n \gg n_0 + p_0$, согласно (7-20) $R = \gamma (\Delta n)^2$ и уравнение (13-17) примет вид:

$$\frac{d \Delta n}{dt} = \beta \alpha \frac{I}{h\nu} - \gamma (\Delta n)^2. \quad (13-20)$$

Для решения этого уравнения также используем начальное условие: $\Delta n = 0$ при $t = 0$. В результате получим, что неравновесная концентрация электронов при квадратичной рекомбинации описывается выражением для нарастания

$$\Delta n = \sqrt{\frac{\beta \alpha I}{\gamma h\nu}} \operatorname{th} t \sqrt{\frac{\beta \alpha \gamma I}{h\nu}}. \quad (13-21)$$

Аналогично для процесса спада концентрации носителей заряда при выключении света

$$\Delta n = \sqrt{\frac{\beta \alpha I}{\gamma h\nu}} \frac{1}{1 + \sqrt{\frac{\beta \alpha \gamma I}{h\nu}} t}. \quad (13-22)$$

Следовательно, *при большом уровне освещения нарастание фотопроводимости описывается гиперболической тангенсойдой, а спад — гиперболическим законом.*

Величина фотопроводимости зависит от интенсивности света. При линейном законе рекомбинации стационарная фотопроводимости

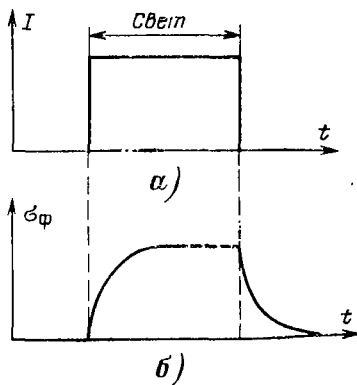


Рис. 13-8. Релаксация фотопроводимости (б) при возбуждении ее импульсом света прямоугольной формы (а).

мость согласно (13-19) пропорциональна интенсивности света Γ . е.

$$\sigma_{\Phi} \sim \frac{I}{h\nu} \quad \text{и} \quad j_{\Phi} \sim \frac{I}{h\nu}. \quad (13-23)$$

При квадратичной рекомбинации из (13-22) следует, что

$$\sigma_{\Phi} \sim \sqrt{\frac{I}{h\nu}} \quad \text{и} \quad j_{\Phi} \sim \sqrt{\frac{I}{h\nu}}. \quad (13-24)$$

В общем случае можно считать, что

$$\sigma_{\Phi} \sim \frac{I^b}{h\nu}. \quad (13-25)$$

При $b = 1$ фоторезистивный эффект называется линейным, при $b < 1$ — сублинейным. Следовательно, люкс-амперная характеристика полупроводника, определяющая зависимость стационарного фототока от интенсивности света, должна иметь два участка: при слабой освещенности — линейный, а при большой освещенности — сублинейный (рис. 13-9).

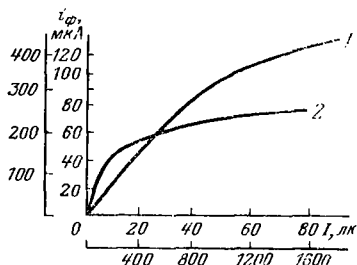


Рис. 13-9. Зависимость стационарного фототока от интенсивности освещения для сернистого висмута.

1 — слабое освещение; 2 — сильное освещение.

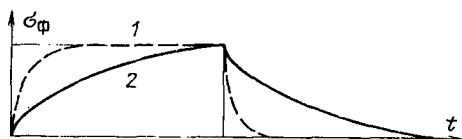


Рис. 13-10. Кривые релаксации фотопроводимости при отсутствии (1) и при наличии (2) центров захвата.

При некоторых условиях возможно соотношение $b > 1$, тогда фоторезистивный эффект называется сверхлинейным.

Если в полупроводнике имеются центры захвата, то они оказывают влияние на кинетику фотопроводимости. Свободные электроны будут не только рекомбинировать с дырками, но будут также захватываться ловушками, что уменьшит скорость нарастания стационарной концентрации носителей заряда. При выключении возбуждающего света опустошение ловушек будет затягивать спад концентрации носителей заряда. В результате этого будет иметь место замедление процесса нарастания и спада фотопроводимости (рис. 13-10).

13-4. ФОТОПРОВОДИМОСТЬ ПРИ НАЛИЧИИ ПОВЕРХНОСТНОЙ РЕКОМБИНАЦИИ И ДИФФУЗИИ НОСИТЕЛЕЙ ЗАРЯДА

В предыдущем анализе фотопроводимости предполагалась однородная генерация носителей заряда по всему образцу. Кроме того, не была учтена поверхностная рекомбинация, которая приводит к уменьшению концентрации неравновесных носителей заряда вблизи поверхности полупроводника.

Рассмотрим теперь случай, когда падающее излучение поглощается неравномерно по толщине образца и имеет место диффузия носителей заряда.

При отсутствии уровней ловушек захвата для вычисления стационарной концентрации избыточных носителей заряда нужно найти решение уравнения непрерывности для неосновных носителей заряда (например, для дырок):

$$-\frac{\Delta p}{\tau_p} + D \frac{d^2 \Delta p}{dx^2} + G e^{-\alpha x} = 0. \quad (13-26)$$

В этом уравнении первое слагаемое учитывает рекомбинацию носителей заряда; второе слагаемое — диффузию носителей заряда, которая возникает как результат неоднородной генерации носителей заряда и влияния поверхностной рекомбинации; третье слагаемое учитывает генерацию носителей заряда по закону:

$$G e^{-\alpha x} = \frac{\beta(1 - R_s) I_0}{h\nu} e^{-\alpha x}. \quad (13-27)$$

Здесь β — квантовый выход; $h\nu$ — энергия падающего кванта света; R_s — коэффициент отражения поверхности полупроводника.

Если излучение интенсивности I_0 падает на поверхность образца с координатой $x = 0$, то граничное условие на этой поверхности можно записать в виде

$$D \left. \frac{d \Delta p}{dx} \right|_{x=0} = s \Delta p|_{x=0}, \quad (13-28)$$

где s — скорость поверхностной рекомбинации на освещенной поверхности. Условия на второй поверхности образца толщиной d можно не учитывать, если считать образец достаточно толстым: $d \gg L$ и $d \gg 1/\alpha$.

Общее решение уравнения (13-26) имеет вид:

$$\Delta p(x) = A e^{-x/L} - \frac{\tau G e^{-\alpha x}}{L^2 \alpha^2 - 1}. \quad (13-29)$$

Используя граничное условие (13-28), будем иметь:

$$\Delta p(x) = \frac{G \tau}{\alpha^2 L^2 - 1} \left[\frac{\alpha L^2 + s \tau}{L + s \tau} e^{-x/L} - e^{-\alpha x} \right]. \quad (13-30)$$

Полное количество генерированных дырок, отнесенных к единице поверхности образца толщиной d , получим, если проинтегрируем $\Delta p(x)$ в пределах от 0 до d и учтем равенство (13-27):

$$\Delta P = \int_0^d \Delta p(x) dx = \frac{\beta(1-R_s)I_0\tau\alpha}{hv(\alpha^2L^2-1)} \left[\frac{\alpha L^2 + s\tau}{L + s\tau} L(1 - e^{-dL}) - \frac{1}{\alpha}(1 - e^{-\alpha d}) \right]. \quad (13-31)$$

Поскольку $\Delta P = \Delta N$, а фотопроводимость пропорциональна ΔP (13-6), то выражение (13-31) позволяет проанализировать влияющие скорости поверхностной рекомбинации на фотопроводимость, а также выяснить некоторые особенности спектральной зависимости фотопроводимости вблизи границы собственного поглощения полупроводника. Для этого кривую собственного поглощения условно разделим на два участка.

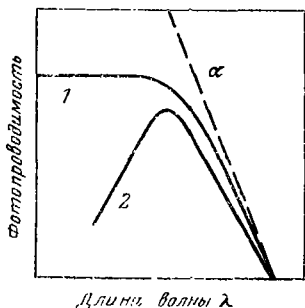


Рис. 13-11. Зависимость фототока от длины поглощенного света при отсутствии (кривая 1) и наличии (кривая 2) поверхностной рекомбинации.

1. Область малого коэффициента поглощения, когда $\alpha d \ll 1$ и $\alpha L \ll 1$. В этом случае $1 - e^{-\alpha d} \approx \alpha d$, а $e^{-dL} \approx 0$, так как $d \gg L$. Поэтому уравнение (13-31) примет вид:

$$\Delta P = \frac{\beta(1-R_s)I_0\tau\alpha}{hv} \left(d - \frac{\alpha L^2 + s\tau}{L + s\tau} L \right) \approx \frac{\beta(1-R_s)I_0\tau}{hv} \alpha \left(d - \frac{s\tau L}{L + s\tau} \right). \quad (13-32)$$

Из этого выражения следует, что в области малого коэффициента поглощения фотопроводимость линейно растет с увеличением α . При этом, если $s\tau \ll d$, то поверхностная рекомбинация слабо влияет на фотопроводимость, если же $s\tau \gg d$, то величина фотопроводимости будет меньше (кривая 2, рис. 13-11), чем при $s = 0$ (кривая 1, рис. 13-11).

2. Область средних значений коэффициента поглощения, когда $\alpha d > 1$, а $\alpha L < 1$. В этом случае

$$\Delta P = \frac{\beta(1-R_s)I_0\tau\alpha}{hv} \left[\frac{1}{\alpha} - \frac{\alpha L^2 + s\tau}{L + s\tau} \right] \approx \frac{\beta(1-R_s)I_0\tau}{hv} \left[1 - \frac{\alpha s\tau}{L + s\tau} \right]. \quad (13-33)$$

Видно, что с дальнейшим ростом коэффициента поглощения α фотопроводимость будет убывать (кривая 2, рис. 13-11). В случае отсутствия поверхностной рекомбинации ($s = 0$) фотопроводимость стремится к насыщению (кривая 1, рис. 13-11).

Следовательно, наличие поверхностной рекомбинации приводит к появлению максимума на кривой спектральной зависимости фотопроводимости (см. рис. 13-2).

13-5. ЭФФЕКТ ДЕМБЕРА

Если полупроводник осветить сильно поглощаемым светом, то вблизи поверхности, где происходит поглощение света, возникнут электронно-дырочные пары. Эти носители заряда диффундируют из освещенной области в глубь полупроводника (рис. 13-12). Поскольку электроны обладают большей подвижностью, чем дырки, то они продвигнутся дальше в глубь кристалла, чем дырки. Эта разница в диффузии носителей заряда приведет к тому, что поверхность полупроводника окажется положительно заряженной по отношению к его объему. Возникшее электрическое поле направлено так, что оно ускоряет дырки, имеющие меньшую подвижность, и замедляет подвижные электроны, поэтому суммарный ток равен нулю. *Объемную э. д. с., возникающую при неравномерном освещении полупроводника, называют э. д. с. Дембера.*

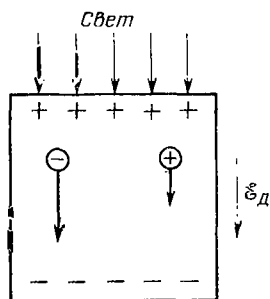


Рис. 13-12. Возникновение э. д. с. Дембера.

Вычисление э. д. с. Дембера проведем в предположении, что в полупроводнике мало ловушек и нет поверхностной рекомбинации. Для этого необходимо решить уравнения

$$j_n = e(n_0 + \Delta n) \mu_n \mathcal{E}_D + eD_n \frac{d\Delta n}{dx}; \quad (13-34)$$

$$j_p = e(p_0 + \Delta p) \mu_p \mathcal{E}_D - eD_p \frac{d\Delta p}{dx}, \quad (13-35)$$

где \mathcal{E}_D — поле Дембера и $\Delta n = \Delta p$ — концентрация избыточных носителей заряда. В стационарных условиях полная плотность тока равна нулю. Так как $eD_p = \mu_p kT$ и $eD_n = \mu_n kT$, то

$$j = j_n + j_p = e[(n_0 + \Delta n) \mu_n + (p_0 + \Delta p) \mu_p] \mathcal{E}_D + e \left(D_n \frac{dn}{dx} - D_p \frac{dp}{dx} \right) \quad (13-36)$$

и для поля Дембера получим выражение

$$\mathcal{E}_D = - \frac{kT}{e} \frac{(\mu_n - \mu_p) \frac{d\Delta n}{dx}}{n\mu_n + p\mu_p}, \quad (13-37)$$

где $n = n_0 + \Delta n$ и $p = p_0 + \Delta p$.

Проинтегрируем выражение (13-37) по x от поверхности ($x = 0$) до такой точки, где концентрация неравновесных носителей обращается в нуль. Это расстояние должно быть порядка нескольких длин диффузии ($x = 3L$). Так как $\sigma = e(n\mu_n + p\mu_p)$, а $\Delta\sigma = e(\mu_n + \mu_p)\Delta n$, то

$$\mathcal{E}_D = -\frac{kT}{e} \frac{b-1}{b+1} \frac{1}{\sigma} \frac{d\Delta\sigma}{dx}, \quad (13-38)$$

где $b = \mu_n/\mu_p$.
Следовательно,

$$V_D = -\frac{kT}{e} \frac{b-1}{b+1} \int_{x=0}^{x=3L} \frac{d\Delta\sigma}{\sigma} = \frac{kT}{e} \ln \frac{\sigma(0)}{\sigma(3L)}. \quad (13-39)$$

На расстоянии $x = 3L$ все неравновесные носители прорекомбинируют, а поэтому $\sigma(3L) = e(n_0\mu_n + p_0\mu_p)$. У поверхности полупроводника $\sigma(0) = e(n_0 + \Delta n)\mu_n + e(p_0 + \Delta p)\mu_p = \sigma(3L) + e(\mu_n + \mu_p)\Delta n$. Поэтому

$$V_D = \frac{kT}{e} \frac{b-1}{b+1} \ln \left[1 + \frac{(b+1)\Delta n}{bn_0 + p_0} \right]. \quad (13-40)$$

Из (13-43) следует, что *э. д. с. Дембера тем больше, чем сильнее различаются подвижности электронов и дырок*. Электродвижущая сила Дембера обычно очень мала, она немного больше kT/e . При $\mu_n = \mu_p$ $V_D = 0$.

13-6. ФОТОЭЛЕКТРОМАГНИТНЫЙ ЭФФЕКТ

Если неравномерно освещенный полупроводник поместить в магнитное поле, перпендикулярное направлению диффузии носителей заряда, то в нем возникает поперечная э. д. с. (рис. 13-13). Это явление называется *фотоэлектромагнитным эффектом* (ФЭМ эффектом).

На диффундирующие электроны и дырки, созданные светом, в магнитном поле действует сила Лоренца, отклоняющая их в направлении, перпендикулярном к направлению их движения и направлению магнитного поля. При этом потоки электронов и дырок отклоняются в разные стороны. Вследствие этого у противоположных сторон образца, как это представлено на рис. 13-14, скапливаются заряды противоположного знака и возникает электрическое поле, а следовательно, возникает напряжение V_y ФЭМ эффекта. Накопление зарядов и возрастание э. д. с. будут продолжаться до тех пор, пока ток проводимости, обусловленный возникшим электрическим полем, не компенсирует магнитно-диффузионный ток.

Определение V_y и тока короткого замыкания проведем для полупроводника, у которого скорость поверхностной рекомбинации

мала, а поглощение света происходит в приповерхностном слое образца с образованием электронно-дырочных пар.

В случае слабого магнитного поля ($\mu B \ll 1$) углы Холла для электронов φ_n и для дырок φ_p согласно (6-88) запишутся в виде

$$\varphi_n = -\mu_n^{(H)} B; \quad (13-41)$$

$$\varphi_p = \mu_p^{(H)} B, \quad (13-42)$$

где $\mu_n^{(H)}$ и $\mu_p^{(H)}$ — холловские подвижности электронов и дырок. Они связаны с дрейфовыми подвижностями соотношениями $\mu_n^{(H)} = r\mu_n$ и $\mu_p^{(H)} = r\mu_p$, где r — Холл-фактор, зависящий от механизма рассеяния (§ 6-5).

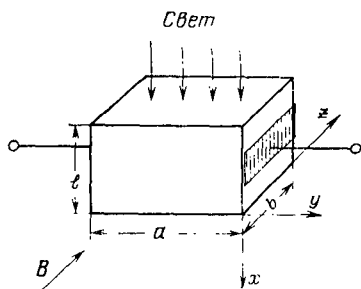


Рис. 13-13. Образец для измерения напряжения фотоэлектроманнитного эффекта.

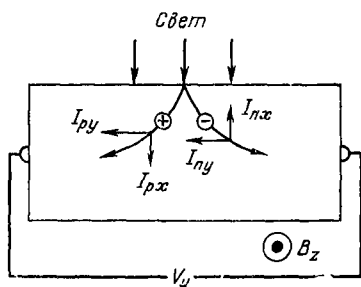


Рис. 13-14. Возникновение напряжения V_y при фотоэлектроманнитном эффекте.

Если j_{nx} и j_{px} есть диффузионные токи в направлении оси x , то магнитно-диффузионные токи согласно рис. 13-14 равны:

$$j_{ny} = j_{nx} \operatorname{tg} \varphi_n = -r\mu_n B j_{nx}; \quad (13-43)$$

$$j_{py} = j_{px} \operatorname{tg} \varphi_p = r\mu_p B j_{px}. \quad (13-44)$$

Полный магнитно-диффузионный ток составит:

$$j_y = j_{ny} + j_{py} = rB (\mu_p j_{px} - \mu_n j_{nx}). \quad (13-45)$$

Если в полупроводнике нет ловушек, то $\Delta n = \Delta p$. С учетом этого диффузионные токи запишутся в виде

$$j_{px} = -j_{nx} = -eD \frac{d\Delta n}{dx}, \quad (13-46)$$

где D — коэффициент биполярной диффузии.

Подставляя (13-46) в (13-45), получаем:

$$j_y = -erDB (\mu_n + \mu_p) \frac{d\Delta n}{dx}. \quad (13-47)$$

Для того чтобы найти ток короткого замыкания во внешней цепи $j_{к.з.}$, 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

$$j_{к.з.} = \int_0^l j_y b dx = -erBD (\mu_n + \mu_p) b \int_0^l d \Delta n = \\ = erDB (\mu_n + \mu_p) b [\Delta n(0) - \Delta n(l)]. \quad (13-48)$$

В стационарном состоянии полный ток через все сечение образца в плоскости xz равен нулю, т. е. ток короткого замыкания равен полному току проводимости, который определяется напряжением ФЭМ эффекта $V_{фэм}$ и полной проводимостью образца σ :

$$j = V_{фэм} \sigma = j_{к.з.} \quad (13-49)$$

Поскольку полная проводимость образца

$$\sigma = \frac{bl}{a} \sigma_0 + \frac{b}{a} \int_0^l \Delta \sigma dx = \frac{bl}{a} \sigma_0 + \frac{eb}{a} (\mu_n + \mu_p) \int_0^l \Delta n dx, \quad (13-50)$$

где a — размер образца по оси y ; σ_0 — проводимость в темноте, то полная величина фото-э. д. с. составляет:

$$V_{фэм} = \frac{j_{к.з.}}{\sigma} = \frac{eraDB (\mu_n + \mu_p) [\Delta n(0) - \Delta n(l)]}{l\sigma_0 + e (\mu_n + \mu_p) \int_0^l \Delta n dx}. \quad (13-51)$$

Следовательно, напряжение фотоэлектромагнитного эффекта пропорционально индукции магнитного поля и зависит от концентрации неравновесных носителей заряда.

Если генерация носителей заряда происходит вблизи освещенной поверхности, то их концентрация из-за диффузии вдоль оси x изменяется по закону

$$\Delta n = \Delta n(0) e^{-x/L}. \quad (13-52)$$

В том случае, если толщина образца l значительно превосходит диффузионную длину L , т. е. $l \gg L$, то $\Delta n(l) = 0$, а

$$\int_0^l \Delta n dx = \Delta n(0) \int_0^l e^{-x/L} dx = \Delta n(0) L. \quad (13-53)$$

Следовательно,

$$V_{фэм} = \frac{earDB (\mu_n + \mu_p) \Delta n(0)}{l\sigma_0 + e (\mu_n + \mu_p) \Delta n(0) L}. \quad (13-54)$$

При высоком уровне возбуждения, когда $\Delta n \gg n_0 + p_0$, в равенстве (13-54) можно пренебречь членом $l\sigma_0$, а поэтому

$$V_{\text{фэм}} = raDB \frac{1}{L}. \quad (13-55)$$

При низком уровне возбуждения, когда $\Delta n \ll n_0 + p_0$, в знаменателе выражения (13-54) можно пренебречь вторым слагаемым. В результате получим:

$$V_{\text{фэм}} = eraDB (\mu_n + \mu_p) \frac{\Delta n(0)}{l\sigma_0}. \quad (13-56)$$

В случае, если скорость поверхностной рекомбинации пренебрежимо мала, то полное число неравновесных носителей заряда в образце, приходящихся на единицу освещенной поверхности, определяется квантовым выходом β , интенсивностью возбуждающего света I и временем жизни носителей заряда τ . Следовательно, учитывая (13-53), можно написать:

$$\int_0^l \Delta n dx = \beta\tau \frac{I}{hv} = \Delta n(0)L, \quad (13-57)$$

откуда

$$\Delta n(0) = \frac{\beta\tau I}{hvL} = \frac{\beta LI}{hvD}. \quad (13-58)$$

Подставляя (13-58) в (13-56), получаем:

$$V_{\text{фэм}} = eraB (\mu_n + \mu_p) \beta L \frac{I}{hvl\sigma_0}. \quad (13-59)$$

Из анализа формул (13-59) и (13-55) следует, что напряжение ФЭМ эффекта при малом уровне возбуждения пропорционально интенсивности света, а с ростом уровня освещения стремится к насыщению (рис. 13-15).

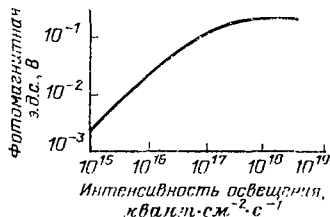


Рис. 13-15. Зависимость фотомагнитной э. д. с. от интенсивности освещения для германия.

13-7. ФОТОВОЛЬТАИЧЕСКИЙ ЭФФЕКТ В p - n ПЕРЕХОДАХ

На p - n переходе, как известно (§ 9-8), существует контактная разность потенциалов. Этот потенциальный барьер обусловлен электрическим полем, которое появляется в результате диффузии основных носителей заряда через p - n переход. При термодинамическом равновесии положение уровня Ферми во всей системе постоянно и энергетическая схема p - n перехода имеет вид, изобра-

женный на рис. 13-16, а. В этом случае токи обусловлены свободными носителями заряда, генерируемыми за счет теплового возбуждения, и в равновесии суммарный ток равен нулю. При принятом обозначении токов, как это сделано на рис. 13-16, а, условие равновесия запишется в виде

$$-j_n^{(n)} + j_n^{(p)} - j_p^{(p)} + j_p^{(n)} = 0. \quad (13-60)$$

В этом равенстве каждая пара токов электронов и дырок равна нулю

$$-j_n^{(n)} + j_n^{(p)} = 0; \quad -j_p^{(p)} + j_p^{(n)} = 0, \quad (13-61)$$

так как количества переходящих носителей заряда в прямом и обратном направлении при термодинамическом равновесии равны.

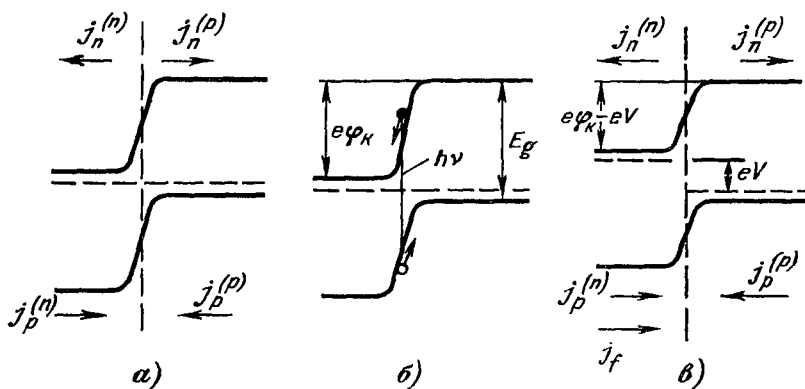


Рис. 13-16. Энергетическая схема p - n перехода и токи носителей заряда при термодинамическом равновесии (а), при освещении (б, в).

Но потоки неосновных носителей заряда — электронов из p -области и дырок из n -области есть не что иное, как электронная и дырочная составляющие тока насыщения в вольт-амперной характеристике диода (9-154). Полный ток насыщения

$$j_s = j_{ns} + j_{ps} = \frac{eD_n n_p}{L_n} + \frac{eD_p p_n}{L_p}. \quad (13-62)$$

Рассмотрим теперь p - n переход, на который падают фотоны с энергией большей, чем ширина запрещенной зоны (рис. 13-16, б). В результате поглощения фотона возникает электронно-дырочная пара. Под действием внутреннего поля в p - n переходе созданные светом носители заряда движутся в противоположных направлениях: дырки в p -область, а электроны — в n -область (см. рис. 13-16, б). Эти перешедшие через p - n переход носители заряда создадут добавочный ток, который обозначим j_f . Так как перешедшие в p -область избыточные дырки уменьшают отрицательный объемный заряд, то энергетические уровни в p -области снижаются

и в результате этого происходит понижение потенциального барьера. Следовательно, разделение зарядов приводит к возникновению разности потенциалов V , приложенной к p - n переходу в прямом направлении (рис. 13-16, θ). Электроны из n -области и дырки из p -области, преодолевая пониженный потенциальный барьер $e(\varphi_k - V)$, будут инжектированы в другую область, где они превращаются в неосновные носители заряда и рекомбинируют. При этом токи, обусловленные инжектированными носителями заряда, направлены из p - в n -область (рис. 13-16, θ). Стационарное состояние установится тогда, когда число создаваемых светом электронно-дырочных пар сравняется с числом носителей заряда, уходящих через пониженный потенциальный барьер p - n перехода. Если p - n переход соединен с внешней цепью, то можно измерить фото-э. д. с., которая носит название вентильной э. д. с., или, если сопротивление нагрузки мало, фототок. Следовательно, освещенный n - p переход действует как фотоэлемент.

Для определения вентильной фото-э. д. с. $V_{\text{вент}}$ запишем уравнение для тока j , текущего через p - n переход:

$$j = j_f - j_n^{(n)} + j_n^{(p)} - j_p^{(p)} + j_p^{(n)}. \quad (13-63)$$

Здесь токи неосновных носителей заряда при освещении $j_n^{(p)}$ и $j_p^{(n)}$, как следует из рис. 13-16, θ , равны своим значениям в равновесии:

$$j_n^{(p)} = j_{sn}; \quad j_p^{(n)} = j_{sp}. \quad (13-64)$$

Токи основных носителей заряда при освещении $j_n^{(n)}$ и $j_p^{(p)}$ в результате понижения потенциального барьера на величину eV увеличиваются и становятся равными:

$$j_n^{(n)} = j_{sn} e^{eV/kT}; \quad (13-65)$$

$$j_p^{(p)} = j_{sp} e^{eV/kT}. \quad (13-66)$$

Объединяя выражения (13-64), (13-65) и (13-66), получаем:

$$j_f - j_{sn} (e^{eV/kT} - 1) - j_{sp} (e^{eV/kT} - 1) = j \quad (13-67)$$

или

$$j_f - j = j_s (e^{eV/kT} - 1), \quad (13-68)$$

откуда будем иметь:

$$V = \frac{kT}{e} \ln \left(1 + \frac{j_f - j}{j_s} \right). \quad (13-69)$$

Уравнение (13-69) является уравнением фотодиода для любого режима.

Для определения вентильной фото-э. д. с. $V_{\text{вент}}$, которая соответствует напряжению на зажимах разомкнутой цепи, необходимо положить $j = 0$. Тогда

$$V_{\text{вент}} = \frac{kT}{e} \ln \left(1 + \frac{j_f}{j_s} \right). \quad (13-70)$$

Величина j_f определяется числом избыточных носителей заряда, созданных светом и дошедших до $p-n$ перехода. Если через $I/h\nu$ обозначим число фотонов, падающих каждую секунду на единицу поверхности, β — квантовый выход, т. е. количество электронно-дырочных пар, возникающих на один фотон, а через γ — долю непрорекомбинировавших пар носителей заряда, пришедших к $p-n$ переходу, то

$$j_f = e\gamma\beta \frac{I}{h\nu}. \quad (13-71)$$

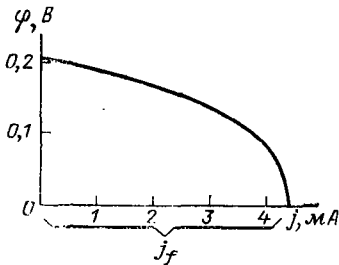


Рис. 13-17. Вольт-амперная характеристика германиевого фотодиода в вентильном режиме.

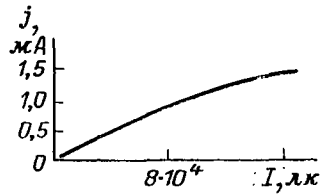


Рис. 13-18. Люкс-амперная характеристика германиевого фотодиода.

Это выражение справедливо для случая, когда весь падающий на полупроводник свет поглощается. С учетом (13-71) выражение (13-70) примет вид:

$$V_{\text{вент}} = \frac{kT}{e} \ln \left(1 + \frac{e\beta\gamma I}{h\nu j_s} \right). \quad (13-72)$$

При большом уровне освещения, когда $\frac{j_f}{j_s} \gg 1$, имеем:

$$V_{\text{вент}} = \frac{kT}{e} \ln \frac{e\beta\gamma I}{h\nu j_s}. \quad (13-73)$$

При малом уровне возбуждения, когда $\frac{j_f}{j_s} \ll 1$, используя разложение в ряд, получаем:

$$V_{\text{вент}} = kT \frac{\beta\gamma I}{h\nu j_s}, \quad (13-74)$$

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. 13-17 и 13-18,

на которых приведены вольт-амперная и люкс-амперная характеристики для германиевых фотодиодов в вентильном режиме. Из рис. 13-17 видно, что малым токам соответствует предельное для данной освещенности значение фото-э. д. с.

Из формулы (13-73) следует, что при увеличении интенсивности света фото-э. д. с. возрастает до тех пор, пока не сравняются j_f и j_s , т. е. пока не исчезнет потенциальный барьер, препятствующий переходу носителей заряда. Высота барьера представляет собой максимально достижимое значение $V_{\text{вент}}$. Поэтому вентильная фото-э. д. с. зависит от степени легирования, а следовательно, от положения уровня Ферми. Практически этот предел соответствует ширине запрещенной зоны.

У некоторых полупроводников, изготовленных в виде тонких слоев, наблюдается фотовольтаический эффект, который характеризуется напряжением, величина которого во много раз превышает ширину запрещенной зоны. Были получены фото-э. д. с., достигающие 500 В. На рис. 13-19 представлены кривые зависимости фото-э. д. с. от интенсивности света для различных полупроводников. Этот эффект наблюдается в том случае, если изолирующая подложка, на которой происходит рост пленки полупроводника, наклонена по отношению к направлению движения пара испаряемого вещества. Угол наклона подложки должен лежать в пределах от 30° до 60° , а температура предварительного нагрева подложки составляет $50-100^\circ\text{C}$. Состав и давление остаточных газов в вакуумной камере имеют большое значение, а иногда после осаждения пленки полупроводника на подложку требуется дополнительная ее обработка. Существенны также выбор материала подложки и ее чистота. При этом адсорбция и десорбция кислорода влияют на величину фотовольтаического эффекта и могут привести к изменению его знака.

Для объяснения этого эффекта привлекаются модели, основанные на представлении о суммировании элементарных фото-э. д. с., генерируемых цепочкой малых фотоэлементов. Эти фото-э. д. с. возникают либо в результате эффекта Дембера в микроячейках, либо из-за фотовольтаического эффекта в $p-n$ переходах или на границах зерен. Однако ни одна из этих моделей не описывает все наблюдаемые факты.

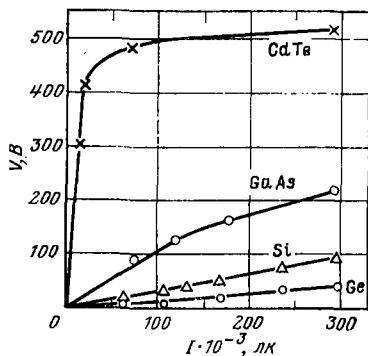


Рис. 13-19. Зависимость напряжения от интенсивности света для пленок, изготовленных из различных полупроводников.

13-8. ФОТОВОЛЬТАИЧЕСКИЙ ЭФФЕКТ НА БАРЬЕРЕ ШОТТКИ

На контакте металла с полупроводником имеет место перераспределение зарядов, приводящее к возникновению в полупроводнике обедненного слоя, сопровождающегося искривлением зон

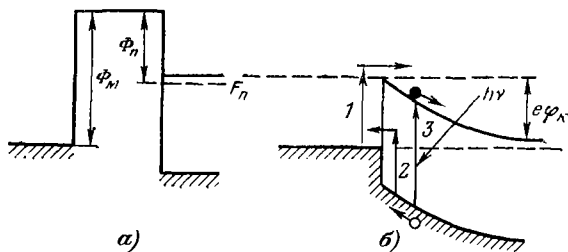


Рис. 13-20. Образование барьера Шоттки в полупроводнике, находящемся в контакте с металлом, и возможные процессы внутренней фотоэмиссии.

полупроводника. Это искривление зон вблизи поверхности раздела называется барьером Шоттки. Если работа выхода из металла Φ_M больше работы выхода из полупроводника n -типа Φ_n (рис. 13-20, а), образуется барьер Шоттки для электронов (рис. 13-20, б). Если $\Phi_M < \Phi_n$, то образуется барьер Шоттки для дырок. Высота этого барьера равна:

$$e\varphi_K = \Phi_M - \Phi_n. \quad (13-75)$$

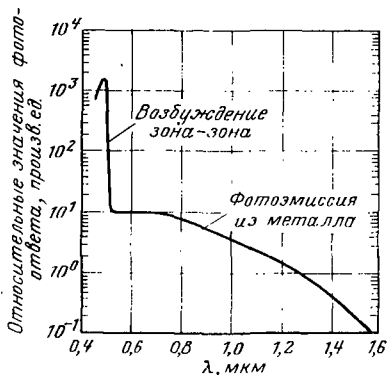


Рис. 13-21. Спектральная зависимость фотоответа контакта металл—полупроводник.

Пусть на структуру, изображенную на рис. 13-20, б, падают фотоны с энергией $h\nu > e\varphi_K$. В результате оптического возбуждения электроны в металле приобретают энергию, достаточную для преодоления барьера (процесс 1). Электроны, входя в полупроводник, заряжают его отрицательно, т. е. на барьере возникает фото-э. д. с. Электроны, возбужденные из валентной зоны полупроводника в металл, будут в металле занимать состояния выше уровня Ферми

(процесс 2). Но этот переход совершается при помощи туннелирования, его вероятность намного меньше, чем для процесса 1. Кроме того, за счет процесса 2 не будет возникать фото-э. д. с., поскольку дырки не будут переходить в металл.

Если $h\nu > E_g$, то электронно-дырочные пары, генерируемые в пределах барьера Шоттки (процесс 3), разделяются контактным полем и создают фото-э. д. с. между металлическим электродом и объемом полупроводника.

На рис. 13-21 приведена спектральная зависимость фототока контакта металл—полупроводник. Видно, что на характеристике имеется две области: область большого фототока, обусловленного генерацией пар носителей заряда при межзонных переходах, и область меньших значений фототока, простирающаяся в длинноволновую сторону и связанная с миграцией горячих электронов из металла в полупроводник с переходом через барьер.

Фотогальванические диоды с барьером Шоттки могут быть выполнены в виде фотоэлементов большой площади. Они могут быть использованы для регистрации фототока, вызванного α -частицами. Кроме того, они могут применяться для регистрации электронов, возбужденных рентгеновскими лучами, γ -лучами и тепловыми нейтронами. Тепловые нейтроны представляют собой неионизирующие частицы, но, например, в InP они вызывают превращение In^{115} в радиоактивный изотоп In^{116} , который превращается в Sn^{116} в результате β -распада с эмиссией электронов высокой энергии, которые и генерируют электронно-дырочные пары.

13.9. ВНЕШНИЙ ФОТОЭФФЕКТ

На рис. 13-22 представлена энергетическая схема полупроводника разной степени легирования. Здесь $E_{\text{вак}}$ есть энергия, которой обладает электрон, вышедший из полупроводника и имеющий в вакууме практически нулевую кинетическую энергию.

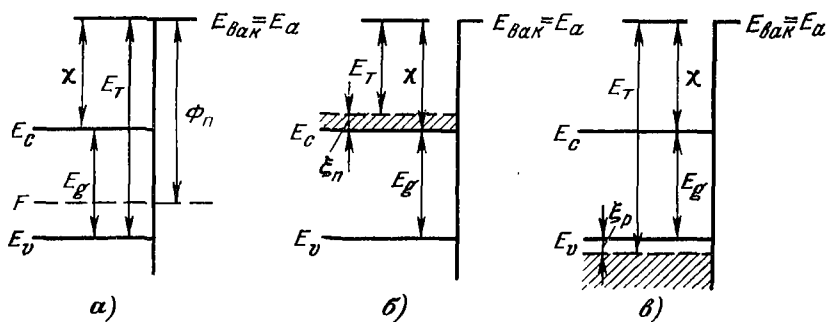


Рис. 13-22. Зависимость порогового фотоэффекта от легирования.

Энергия χ , отделяющая край зоны проводимости от уровня вакуума, есть энергия электронного сродства. Величина Φ , равная разности энергий, соответствующих уровню Ферми и уровню вакуума, есть работа выхода.

Рассмотрим теперь взаимодействие между фотоном и полупроводником, в результате которого происходит эмиссия электрона из полупроводника. Процесс эмиссии электронов из полупроводника под действием излучения называется внешним фотоэффектом. Внеш-

ний фотоэффект представляет собой последовательность трех процессов (рис. 13-23): 1) электрон валентной зоны полупроводника переходит в высокое энергетическое состояние зоны проводимости в результате взаимодействия с фотоном; 2) возбужденный электрон в результате рассеяния теряет часть энергии и переходит на нижний уровень зоны проводимости; 3) электрон выходит с нижнего уровня зоны проводимости полупроводника в вакуум с энергией, равной разности его полной энергии и $E_{\text{вак}}$. Порог внешнего фотоэффекта E_T есть наименьшая энергия фотона, которая достаточна, чтобы удалить электрон из полупроводника.

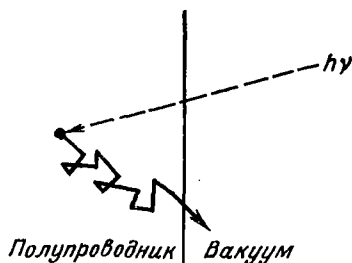


Рис. 13-23. Возбуждение, рассеяние и выход фотоэлектронов из полупроводника.

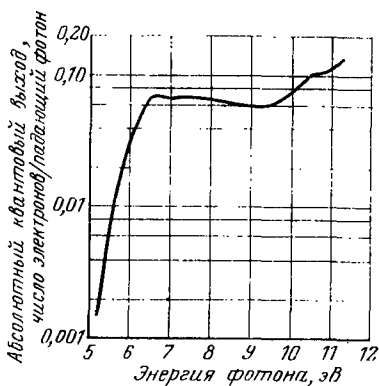


Рис. 13-24. Спектральное распределение квантового выхода электронов из CdTe.

Порог фотоэффекта равен примерно 5 эВ.

У собственного полупроводника или невырожденного полупроводника, когда уровень Ферми находится в запрещенной зоне (рис. 13-22, а), эмиссия электронов под действием света происходит из валентной зоны. Поэтому

$$E_T = \chi + E_g. \quad (13-76)$$

Для непрямых переходов, когда сохранение квазиимпульса обеспечивается за счет эмиссии фонона с энергией E_p ,

$$E_T = \chi + E_g + E_p. \quad (13-77)$$

Для сильно легированного полупроводника n -типа, у которого уровень Ферми лежит выше края зоны проводимости на величину ξ_n (рис. 13-22, б), имеем:

$$E_T = \chi - \xi_n. \quad (13-78)$$

В сильно легированном полупроводнике p -типа уровень Ферми расположен на величину ξ_p ниже края валентной зоны (рис. 13-22, в), поэтому

$$E_T = \chi + E_g + \xi_p. \quad (13-79)$$

Работа выхода Φ , порог фотоэффекта E_T и порог для прямых переходов E_{Td} для некоторых полупроводников приведены в табл. 13-1. В ней даны также значения ширины запрещенной зоны E_g , энергия электронного сродства χ и указаны плоскости, с которых происходил фотоэффект.

Таблица 13-1

Работа выхода и порог внешнего фотоэффекта для некоторых полупроводников

Полупроводник и плоскость	E_g , эВ	Φ , эВ	E_T , эВ	E_{Td} , эВ	$\chi = E_T - E_g$, эВ
AlSb (110)	1,5	4,86	5,15	5,75	3,65
GaAs (110)	1,40	4,71	5,47	5,75	4,07
GaSb (110)	0,70	4,76	4,76	5,24	4,06
InAs (110)	0,41	4,90	5,31	5,58	4,90
InSb (110)	0,18	4,75	4,77	5,26	4,59
InP (110)	1,3	4,45	5,68	5,94	4,38
Si (111)	1,09	4,83	5,10	5,45	4,01
Ge (111)	0,67	4,80	4,80	5,22	4,13

Типичная кривая зависимости квантового выхода фотоэлектронов от энергии приведена на рис. 13-24. Вначале по мере увеличения энергии фотонов, превышающей пороговое значение, число эмиттируемых электронов возрастает и квантовый выход резко нарастает. Затем он переходит в плато, на котором имеется структура, отражающая свойства зонной структуры полупроводника. Начальный участок резкого нарастания кривой квантового выхода $Y(h\nu)$ описывается зависимостью вида

$$Y(h\nu) = A(h\nu - E_T)^r, \quad (13-80)$$

где A и r — константы, r определяется механизмом рассеяния. В случае прямого перехода электрона из валентной зоны без рассеяния $r = 1$, а при наличии рассеяния $r = 2$. Для непрямого перехода электрона из валентной зоны без рассеяния и с рассеянием $r = 5/2$.

ПРИЛОЖЕНИЯ

I. Физические константы

Заряд электрона $e = 4,8 \cdot 10^{-10}$ ед. CGSE $= 1,602 \cdot 10^{-19}$ Кл

Масса свободного электрона $m_0 = 9,11 \cdot 10^{-28}$ г

Скорость света в вакууме $c = 2,998 \cdot 10^{10}$ см/с

Боровский радиус $a_0 = 5,29 \cdot 10^{-9}$ см

Постоянная Планка $h = 6,62 \cdot 10^{-34}$ Дж·с $= 4,5 \cdot 10^{-15}$ эВ·с

Постоянная Дирака $\hbar = h/2\pi = 1,054 \cdot 10^{-34}$ Дж·с

Постоянная Больцмана $k = 1,380 \cdot 10^{-23}$ Дж/К $= 8,62 \cdot 10^{-5}$ эВ/К

Тепловая энергия

$kT = 25,9$ мэВ при комнатной температуре;

$kT = 6,7$ мэВ при температуре жидкого азота;

$kT = 0,36$ мэВ при температуре жидкого гелия

Энергия, соответствующая 1эВ, 1эВ $= 1,602 \cdot 10^{-19}$ Дж

Длина волны в вакууме излучения с энергией фотона 1эВ λ_0 (1 эВ)
 $= 1,239 \cdot 10^{-4}$ см

Волновое число, соответствующее 1 эВ ν_0 (1 эВ) $= 8,06 \cdot 10^4$ см⁻¹

II. Свойства Ge, Si и GaAs (при 300 К)

Свойства	Ge	Si	GaAs
Количество атомов/см ³	$4,42 \times 10^{22}$	$5,0 \times 10^{22}$	$2,21 \times 10^{22}$
Атомная масса	72,6	28,08	144,63
Поле пробоя, В/см	$\sim 10^5$	$\sim 3 \times 10^5$	$\sim 4 \times 10^5$
Кристаллическая структура	Алмаза	Алмаза	Цинковой обманки
Плотность, г/см ³	5,3267	2,328	5,32
Диэлектрическая проницаемость	16	11,8	10,9
Эффективная плотность состояний в зоне проводимости N_c , см ⁻³	$1,04 \times 10^{19}$	$2,8 \times 10^{19}$	$4,7 \times 10^{17}$
Эффективная плотность состояний в валентной зоне N_v , см ⁻³	$6,1 \times 10^{18}$	$1,02 \times 10^{19}$	$7,0 \times 10^{18}$
Эффективные массы m^*/m_0 : электронов	m_l 1,58 m_t 0,082	m_l 0,97 m_t 0,19	0,068
дырок	m_{pl}^* 0,04 m_{pt}^* 0,3	m_{pl}^* 0,16 m_{pt}^* 0,5	m_{pl}^* 0,12 m_{pt}^* 0,5
Сродство к электрону, эВ	4,0	4,05	4,07
Ширина запрещенной зоны при 300 К, эВ	<u>0,66</u>	1,11	1,43
Собственная концентрация, см ⁻³	$2,5 \times 10^{13}$	$1,6 \times 10^{10}$	$1,1 \times 10^7$
Постоянная решетки, Å	5,65748	5,43086	5,6534
Температурный коэффициент линейного расширения $\Delta L/\Delta T$, °C ⁻¹	$5,8 \times 10^{-6}$	$2,6 \times 10^{-6}$	$5,9 \times 10^{-6}$
Точка плавления, °C	937	1420	1238
Время жизни неосновных носителей, с	10^{-3}	$2,5 \times 10^{-3}$	10^{-8}
Подвижность дрейфовая, см ² /(В·с): электронов μ_n дырок μ_p	3900 1900	1500 600	8500 400
Энергия рамановских фононов E_p , эВ	0,037	0,063	0,035
Удельная теплота, Дж/(г·°C)	0,31	0,7	0,35
Теплопроводность (при 300 К), Вт/(см·°C)	0,61	1,45	0,46
Коэффициент тепловой диффузии, см ² /с	0,36	0,9	0,44
Давление паров, Па	0,133 при 1270°C $1,33 \cdot 10^{-6}$ при 800°C	0,133 при 1600°C $1,33 \cdot 10^{-6}$ при 930°C	$1,01 \cdot 10^5$ при 1050°C $1,01 \cdot 10^7$ при 1220°C
Работа выхода, эВ	4,4	4,8	4,7

III. Свойства полупроводников

Полупроводники	Ширина запрещенной зоны		Тип абсолютного минимума зоны проводимости	$\left(\frac{dE_g}{dT}\right) \cdot 10^4$ (300 К), эВ/К	$\left(\frac{dE_g}{dT}\right) \cdot 10^4$ эВ/К	m_n^*/m_0	Эффективная масса		Показатель преломления n	Статическая диэлектрическая проницаемость ϵ	Постоянная решетки a , Å	Подвижность	
	$E_g^{эВ}$ (0 К)	E_g (300 К), эВ					m_p^*/m_0	m^*/m_0				μ_n см ² /В·с	μ_p см ² /В·с
IV	Si	1,166	1,11	Непр. 100	-2,3	-1,5	$m_n^* 0,98$	$m_p^* 0,5$	3,44	11,7	5,43	1350	480
		0,74	0,67	Непр. 111	-3,7	5,0	$m_n^* 0,19$	$m_p^* 0,16$	4,00	16,3	5,66	3900	1900
IV - IV SiC	$\alpha - \text{Sn}$	-0,2		Пр. 000		5,0	$m_n^* 0,082$	$m_p^* 0,04$			6,489	2000	1000
		3,0 (6H)	2,8-3,2	Непр.	-3,3				2,69 с		$a 3,0817$ $c 15,1123$	400	
VI	Se	1,95	1,74	Пр. 0001	-14	-20		0,12				1	
		0,334	0,32	Пр. 0001	-0,3	-19	0,038 ±	0,26 0,10 ±	2,68 ± с	10,2		1100	
BP AlP AlAs	BP AlP AlAs	2,5	2	Непр.	-3,5		0,13		2,6	6,9	4,538	80	
		2,24	2,43	Непр. 100			0,5		3,0	9,8	5,462	1000	
		2,24	2,16	Непр.				$m_n^* 1,06$ $m_p^* 0,49$		12	5,66		

III-V	AlSb	1,6	1,6	Нср. 100	-4	-1,6	0,11	0,39	3,4	11	6,135	50	400
	GaN	3,6	3,5	Пр. 0000	-3,9	3,7	0,2		2,4		α 3,18 β 5,16 5,450	150	
	GaP	2,1	2,25	Нср. 100	-5,4	-1,7	0,13	0,8	3,37	10		120	120
	GaAs	1,520	1,43	Пр. 000	-5,0	11	0,07	$m_{pT}^* 0,5$ $m_{pA}^* 0,12$	3,4	12	5,653	8 600	400
	GaSb	0,81	0,69	Пр. 000	-4,1	12	0,045	0,39	3,9	15	6,095	4 000	650
II-VI	InP	3,12	1,28	Пр. 000	-4,6	4,6	0,07	0,40	3,37	12,1	5,8687	4 000	650
	InAs	0,12	0,36	Пр. 000	-3,3	5	0,028	0,33	3,42	12,5	6,058	30 000	240
	InSb	0,235	0,17	Пр. 000	-2,9	15	0,0133	$m_{pT}^* 0,6$ $m_{pA}^* 0,012$	3,75	18	6,4787	76 000	5000 (78 K)
	ZnO		3,2	Пр. 0000	-9,5	0,6	0,32	0,27	2,2	7,9	α 3,2495 β 5,2065 γ 3,811 δ 6,257	180	
	ZnS		3,8	Пр. 0000	-3,8	9	0,28	0,5	2,4	8,3	5,106 5,067 6,101	100	7
II-VI	ZnSe	2,80	3,6	Пр. 000	-5,3	5,7	0,39	0,7	2,4	8,3		210	
	ZnTe	2,39	2,58	Пр. 000	-7,2	6	0,17	2,5	2,89	8,1		500	
	CdS	2,38	2,53	Пр. 0000	-5	6	0,15	0,35	3,56	9,7		600	
	CdSe	1,85	1,71	Пр. 0000	-4,6	3,3	0,20	0,45	2,5	8,9	α 4,136 β 6,713	210	
	CdTe	1,60	1,50	Пр. 000	-4,1	1,5	0,11	2,5	2,75	10,6	α 4,299 β 7,010 6,477	500	
IV-VI	HgS	-0,24	2,5	Пр. 000	-5,6		0,045	~0,3	3,7	25	6,085	5 500	100 (20 K)
	HgSe	-0,28	-0,15	Пр. 000	-4		0,029			20	6,12	22 000	
	HgTe		0,14	Пр. 000	-4								
	PbS	0,29	0,37	Пр. 111	-4	-7	0,1	0,1	3,7	170	5,936	550	600
	PbSe	0,15	0,26	Пр. 111	-4	-8	$m_f 0,07$ $m_{pT}^* 0,06$	$m_{pA}^* 0,03$	3,8	250	6,124	1020	930
IV-VI	PbTe	0,19	0,29	Пр. 111	-4	-9	$m_f 0,21$ $m_{pT}^* 0,3$	$m_{pA}^* 0,02$		412	6,460	1620	750
	SnTe	0,3	0,18	Пр. 111			$m_f 0,02$				6,325		

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

А

- Адиабатическое приближение 31
- Аккумуляция носителей заряда 262
- Акустические ветви колебаний решетки 89
- Акцептор, определение 25
- Акцепторный полупроводник 25, 134
- Ангармонизм колебаний 104
- Ангармонический осциллятор 105
- Антизапирающий слой 277

Б

- Барьер Шоттки 304
- Барьерная емкость 293
- Биоломинесценция 360
- Бимолекулярное рекомбинационное свечение 371
- Биполярная диффузионная подвижность 265
- дрейфовая подвижность 264
- световая генерация носителей заряда 224
- Биполярный коэффициент диффузии 264

В

- Вероятность переходов 147
- поглощения фонона 232, 327
- рассеяния 155
- релаксации 161
- Вертикальные переходы 326
- Виды рекомбинации 228
- Влияние давления 338
- поля магнитного 340
- — электрического 339
- температуры 337
- Внешний фотоэффект 405
- Внешняя контактная разность потенциалов 275
- Внутренний фотоэффект 383
- Внутренняя контактная разность потенциалов 274
- Водородоподобные центры 76
- Волюновое число 35, 83
- Волюнов вектор решетки 83
- — фонона 83
- — электрона 35
- пакет 56, 57

- Вольт-амперная характеристика контакта металл — полупроводник 283, 286, 288
- — — $p-n$ перехода 295, 298
- Время жизни дырки 230, 241
- — мгновенное 226, 236
- — неравновесных носителей заряда 225, 230, 235, 242
- — — температурная зависимость 243
- — — фотона 233, 327
- — — электрона 230, 235, 242
- — — электроно-дырочной пары 225, 235, 240
- — релаксации 150, 167, 169
- — максвелловское 226
- — — при рассеянии на ионах примеси 176, 178
- — — решеточном рассеянии 176
- Выпрямление на контакте металла — полупроводник 270
- — — $p-n$ переходе 293
- Вырожденный примесный полупроводник 120
- собственный полупроводник 127

Г

- Гармонический осциллятор 87, 93
- Генерация носителей заряда 19, 221
- — биполярная 224
- — монополярная 226

Д

- Демаркационный уровень 248
- Дефекты 155, 229
- — линейные 156
- — точечные 156
- Диодная теория выпрямления 283
- Дислокации 156, 162
- Диффузия носителей заряда 249, 254, 257
- Диффузионная длина 260
- — скорость 260
- — теория выпрямления 286
- Диффузионное рассеяние 312
- Диффузионный ток 251
- Диэлектрическая проницаемость 76
- Диэлектрическое время релаксации 227
- Длина диффузионная 260
- — дрейфа 261

Длина затягивания 259
— свободного пробега носителей заряда 20, 155, 161, 167, 169
— — — фотона 233, 327
— экранирования 256, 270

Долины 69
Домен 213

Донорно-акцепторные пары 369
— переходы 370

Донорный полупроводник 25, 129

Доноры, определение 25

Дрейфовая длина 261

— скорость носителей заряда 21, 28, 57

Дрейфовый ток 253

Дырки 19, 23, 72

— легкие 73

— тяжелые 12

Е

Емкость контакта металла — полупроводник 280
— p - n перехода 293

З

Закон Ома 206, 253

— сохранения квазиниульса 325

— — энергии 325

Запирающий слой 277

Зинеровский пробой 298

Зона Бриллюэна, первая 46, 47, 69

— валентная 22

— запрещенная 22

— примесной проводимости 140

— проводимости 22

Зонная структура антимонида индия 73

— — арсенида галлия 70

— — германия 70

— — кремния 70

И

Изгиб зон 270

Избыточная концентрация носителей заряда 222, 258

Изоэнергетические поверхности 54, 63

— — сферические 65

— — эллипсоидальные 63, 64

Импульс фотона 232

— — электрона 59

Инверсная населенность 377

Инверсный слой 268, 364

Индукция магнитная 181

Инжекция 261, 294

Интеграл столкновения 149

Ионизация примесей 131

К

Катодолюминесценция 360

Квазиниульс 45

Квазиуровень Ферми 223, 281

Квантовые генераторы 378

Квантовый выход излучения 371

— — фотоионизации 388

Кинетическая энергия решетке 86

Кинетическое уравнение Больцмана 144

Ковалентные кристаллы 18

Колебания атомов решетки 81, 82

— — струны 79

Компоненты тензора 61, 63

Контакт металл — металл 273

— металл — полупроводник 276

— — электронного и дырочного полупроводников 288

Контактная разность потенциалов 198, 274, 275, 276

Концентрация дырок 115, 119, 121

— носителей заряда 105, 106, 114

— — — вырождение 122

— — — зависимость от температуры 126, 133

— — электронов 114, 118, 121

Коэффициент диффузии 252, 262

— — захвата 237

— — ионизации 237

— — отражения 323

— — Пельтье 201

— — поглощения 233, 323

— — пропускания 324

— — рекомбинации 221, 234

— — теплового расширения 104

— — теплопроводности 203

— — Томсона 201

— — Холла 360

— — экстинкции 350

Л

Лавинный пробой 298

Лазеры 378

Ловушки захвата 237, 246

— — рекомбинационные 229, 246

Люкс-амперная характеристика 392

Люминесценция 359

Люминесценция, гашение 371

— — метастабильная 360

— — мономолекулярная 360

— — рекомбинационная 360

М

Магнитная проницаемость 250

Максвелловское время релаксации 227

МДП-структура 312

Модуль Юнга 80, 83

Моды колебаний 85

Н

Наклон зон 55, 218

Невырожденный примесный полупроводник 12, 118

Невырожденный собственный полупроводник 124
Непрямые переходы 330
Неравновесная функция распределения 145, 170
Неравновесные носители заряда 13, 222
Нормальные координаты решетки 85, 98

О

Область сильной ионизации примеси 131
— слабой ионизации примеси 138
Обменный интеграл 38
Образование хвоста плотности состояния 143
Обращенный слой 268
Одноэлектронное приближение 32
Оператор Гамильтона 30
Оптическая накачка 378
Оптические ветви колебаний решетки 89

П

Переходы вертикальные 326
— внутрizonные 355
— между хвостами зон 337
— межзонные 326
— не прямые 330
— прямые 326
Периодический потенциал решетки 34
Плотность состояний 105, 107
— потока энергии 203
— тока 26, 250
— — дырочного 175
— — электронного 174
p-n переход 288
— — физический 268, 278
Поверхностная проводимость 310
— рекомбинация 317
Поверхностный потенциал 303
Поверхностные состояния 301
— — быстрые 316
— — медленные 316
— уровни 301
— явления 301
Подвижность носителей заряда 28, 175, 176, 178, 179, 189
— — при эффекте поля 313
— Холлова 189
Поглощение примесное 325, 326
— решеточное 325, 357
— света 323
Поглощение свободными носителями заряда 325, 349
— собственное 325
— — при прямых переходах 325
— — — непрямых переходах 330
— экситонное 345
Показатель поглощения 350

Показатель преломления 232, 327
— — комплексный 350
Поле Холла 183
Полупроводник 14, 15, 167
— акцепторный 25
— вырожденный 117, 120
— донорный 24
— компенсированный 136
— — частично 137
— невырожденный 117, 118
— примесный 24, 116
— сильно легированный 120
— собственный 19, 22, 124
— — вырожденный 127
— — невырожденный 124
Поляризация 352
Поляризуемость 353
Постоянная Больцмана 110
— Планка 35
Потенциальная энергия решетки 86
Правила отбора 326
Приведенная масса 327
Приведенный квазиуровень Ферми 223
— уровень Ферми 114
Примесные зоны 140
Принцип детального равновесия 221
— макроскопической обратности 148
— Паули 44
Проводимость 11, 173
— примесная 24, 180
Процессы в *p-n* переходе при обратном смещении 294
— — — — прямом смещении 293
— генерации 249
— переноса 146, 153
— рассеяния 149

Р

Работа выхода электрона 270
— — — из акцепторного полупроводника 273
— — — собственного полупроводника 272
— — — — электронного полупроводника 273
Равновесная концентрация носителей заряда 121
Равновесное состояние 150
Равновесные носители заряда 13, 221
Радиоломинесценция 360
Разогрев электронно-дырочного газа 206
Рассеяние тили 154
— диффузионное 312
— междолинное 211
— на акустических фононах 354
— — атомах примеси 161
— — дислокациях 162

Ш

- Ширина запрещенной зоны 22, 125, 196, 216, 219, 244, 327
- — — зависимость от давления 338
- — — — — температуры 337

Э

- Э. д. с. Дембера 395
- термоэлектродвижущая 196
- Холла 193
- Экситоны 345
- непрямые 347
- прямые 346
- свободные 348
- связанные 348
- Экситонное излучение 365
- поглощение 325, 345
- Экситонные комплексы 348
- Эксклюзия носителей заряда 262
- Экстракция носителей заряда 262, 295
- Электролюминесценция 360
- Электропроводность примесного полупроводника 18, 24, 26
- собственного полупроводника 18
- Электростатическая ионизация 206, 218
- Элементы тензора 61, 63
- Эллипсоидальные поверхности равной энергии 63, 108, 109

- Энергия активации 119, 125
- гармонического осциллятора 87
- ионизации примеси 77, 142
- связи экситона 346
- Ферми 110
- фонона 332
- электронного сродства 271
- Энергетическая структура p - n перехода 290
- щель 22
- Эффект Ганна 206, 211
- Дембера 395
- Зеебека 195
- Зинера 206, 218
- магнетопоглощения 344
- магниторезистивный 190
- Пельтье 201
- поля 309
- Томсона 201
- фононного увлечения 199
- фотоэлектромагнитный 396
- Франца—Келдыша 340
- Холла 181, 185
- Эффективная масса 60, 74
- — дырки 53
- — легкой 73
- — тяжелой 72
- — плотности состояний 108, 109
- — поперечная 64
- — продольная 64
- Эффективное сечение захвата 241
- — проводимости 160
- — рассеяния 154

