

Лазарь Соломонович Стильбанс
ФИЗИКА ПОЛУПРОВОДНИКОВ

Книга представляет собой систематическое рассмотрение основных разделов физики полупроводников: качественного и количественного описания строения полупроводниковых кристаллов, энергетического спектра и статистики электронов и фононов, теории явлений переноса, оптических и фотоэлектрических свойств и контактных явлений.

В первой главе эти вопросы рассмотрены в качественной форме, а в последующих дается количественный анализ, но при этом везде делается упор на физическую сущность явлений; необходимые для понимания этого материала сведения из теоретической физики (квантовой механики, статистики и термодинамики) приводятся в тексте попутно с основным материалом. Вторая глава посвящена описанию основных свойств кристаллов: симметрии, характера химической связи, дефектов, тепловых колебаний и теплоемкости. В третьей главе дается представление об электронной теории кристаллов (предпосылки введения адиабатического и одноэлектронного приближения, методы анализа и особенности зонной структуры полупроводников). Статистике электронов в полупроводниках посвящена четвертая глава, в которой также приведены некоторые положения термодинамики.

В пятой, шестой и седьмой главах излагаются основы теории явлений переноса (анализ кинетического уравнения, электро- и теплопроводности полупроводников, термоэлектрических, гальвано- и термомагнитных явлений). Восьмая глава посвящена теориям выпрямления на контакте металл — полупроводник и p - n переходе, и девятая — оптическим явлениям (поглощению света, фотопроводимости, фотовольтаическим эффектам и стимулированному излучению).

Книга рассчитана на широкий круг читателей — инженеров, научных работников и студентов старших курсов технических вузов.

СОДЕРЖАНИЕ

| | |
|---------------------------------------------------|-----|
| Предисловие | 3 |
| Глава первая. ОСНОВНЫЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ | 6 |
| 1. 1. Некоторые сведения о строении атома | 6 |
| 1. 2. Энергия и движение электрона в твердом теле | 10 |
| 1. 3. Электропроводность полупроводников | 36 |
| 1. 4. Теплопроводность полупроводников | 43 |
| 1. 5. Контактные явления | 55 |
| 1. 6. Термоэлектрические явления | 75 |
| 1. 7. Гальваномагнитные и термомагнитные явления | 83 |
| 1. 8. Фотопроводимость | 100 |
| Глава вторая. СТРОЕНИЕ КРИСТАЛЛОВ | 113 |
| 2. 1. Некоторые вопросы квантовой теории | 113 |
| 2. 2. Геометрия кристаллической решетки | 147 |

| | |
|-----------------------------------------------------------------------------------|-----|
| 2. 3. Дефекты в кристаллах | 163 |
| 2. 4. Тепловые колебания кристаллов | 174 |
| 2. 5. Теплоемкость | 184 |
| Глава третья. ОСНОВНЫЕ ПРЕДСТАВЛЕНИЯ ЭЛЕКТРОННОЙ ТЕОРИИ КРИСТАЛЛОВ | 190 |
| 3. 1. Адиабатическое приближение | 190 |
| 3. 2. Одноэлектронное приближение | 194 |
| 3. 3. Приближение почти свободных электронов | 198 |
| 3. 4. Приближение сильно связанных электронов | 207 |
| 3. 5. Основные особенности структуры энергетических зон полупроводников | 209 |
| Глава четвертая. СТАТИСТИКА ЭЛЕКТРОНОВ В ПОЛУПРОВОДНИКАХ | 217 |
| 4. 1. Некоторые понятия статистики и термодинамики | 217 |
| 4. 2. Распределение Ферми | 224 |
| 4. 3. Статистика невырожденного электронного газа в полупроводниках | 226 |
| 4. 4. Энергия электронов в зоне проводимости, вырождение | 235 |
| Глава пятая. НЕКОТОРЫЕ ВОПРОСЫ ТЕОРИИ ЯВЛЕНИЙ ПЕРЕНОСА | 244 |
| 5. 1. Элементарный расчет электропроводности и подвижности | 245 |
| 5. 2. Кинетическое уравнение (учет энергетической зависимости времени релаксации) | 260 |
| 5. 3. Феноменологический анализ явлений переноса | 270 |
| 5. 4. Вычисление времени релаксации | 271 |
| 5. 5. Явления в сильных электрических полях | 278 |
| Глава шестая. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ И ТЕПЛОПРОВОДНОСТЬ | 292 |
| 6. 1. Термоэлектродвижущая сила | 294 |
| 6. 2. Вывод коэффициента термо- э.д.с. из кинетического уравнения | 296 |
| 6. 3. Увлечение электронов фононами | 299 |
| 6. 4. Зависимость термо- э.д.с. от температуры и концентрации носителей | 304 |
| 6. 5. Электронная теплопроводность | 311 |
| 6. 6. Теплопроводность кристаллической решетки | 317 |
| 6. 7. Фотонная теплопроводность | 329 |
| Глава седьмая. ГАЛЬВАНОМАГНИТНЫЕ И ТЕРМОМАГНИТНЫЕ ЯВЛЕНИЯ | 331 |
| 7. 1. Общие сведения | 331 |
| 7. 2. Эффект Холла и изменение сопротивления в магнитном поле | 341 |
| 7. 3. Эффект Эттингсгаузена | 350 |
| 7. 4. Гальваномагнитные явления в сильных магнитных полях | 351 |
| 7. 5. Термомагнитные явления | 355 |
| Глава восьмая. КОНТАКТНЫЕ ЯВЛЕНИЯ | 362 |
| 8. 1. Особенности контактных явлений | 362 |
| 8. 2. Контакт полупроводника и металла | 366 |

| | |
|--------------------------------------------------------------------------------------------------------------|-----|
| 8. 3. Диффузионная теория выпрямления Мотта (химический барьерный слой на границе металла с полупроводником) | 368 |
| 8. 4. Диодная теория Бете | 373 |
| 8. 5. Теория физического запирающего слоя (теория истощения Шоттки) | 375 |
| 8. 6. Теория p-n перехода | 378 |
| Глава девятая. ОПТИЧЕСКИЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ | 400 |
| 9. 1. Поглощение света | 400 |
| 9. 2. Фотопроводимость | 409 |
| 9. 3. Фотовольтаические эффекты | 421 |
| 9. 4. Циклотронный резонанс | 426 |
| 9. 5. Стимулированное излучение | 430 |
| Литература | 441 |
| Предметный указатель | 443 |

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

| | |
|-----------------------------------------------------|---------------------------------------------------------|
| Адиабатическое приближение | 190—194, 196 |
| Альфа-подход (в теории термоэлектричества) | 302 |
| Ангармоничность (колебаний) | 48, 54, 55, 186, 319 |
| Антизапорный слой | 64, 65 |
| Атомные кристаллы | 170 |
| Базисный вектор (решетки) | 152, 155, 157, 160, 198, 210 |
| Бозе — Эйнштейна распределение | 322, 324, 328 |
| Больцмана распределение | 184, |
| Бравэ решетки | 156, 157 |
| Бриллюэна зоны | 204, 205, 211, 213, 216 |
| Валентная зона | 13, 14, 16, 18, 36, 109, 203, 434, 435 |
| Вероятность нахождения частицы | 122, 134 — перехода 130, 131, 138, 141 |
| «Вертикальные» переходы — см. Переходы | |
| Видемана и Франца закон | 45, 256, 311, 315, 316 |
| Волна электрона | 121, 122, 128, 158, 249 |
| Волновая функция | 122, 123, 130, 190, 192, 196—200, 274, 275, 285 |
| Волновое число | 47, 160, 259, 285 |
| Волновой вектор | 47, 53, 117, 122, 159-161, 176, 179, 181, 202, 320, 322 |
| Волны упругие | 46, 176, 317, |
| Вольтамперная характеристика контакта двух металлов | 61 |
| — — — полупроводника и металла в диодной теории | 66—68, 373—376 |
| — — — — — в диффузионной теории | 66, 68, 69, 368—373, 376—378 |
| — — p-i-n перехода | 396 |
| — — p-n перехода | 72—75, 385— 396 |
| — — фотоэлемента | 425 |
| Время жизни | 72, 106, 392, 409, 416, 420 |
| Время релаксации | 41, 218—220, 256, 258—261, 267—278, 355, 404, 415 |

— —, энергетическая зависимость 256, 260, 261, 271, 343, 357
— — фононов 300, 301, 317, 323—325
— — —, дисперсионные зависимости 325—326
Время свободного пробега 37, 100, 247, 249, 257, 259—261, 279, 429
Выпрямление на контакте двух металлов 60, 61
— — — — полупроводников 61
— — — — полупроводника и металла 66—70
Вырождение (см. также «масса эффективная», «невырожденное», «электронный газ», «дырочный газ», «состояния», «полупроводники») 31, 78, 235, 282, 339, 413
—, кратность 224
—, критерий 242
— ориентационное 235
— полное 237, 241, 243, 342
—, связь с эффективной массой 240
— сильное 44, 242, 243
— снятие 31, 137, 138, 201, 235
—, степень 30, 238, 261, 312, 316
— частичное 31, 35, 239, 242, 433
Гальвано-магнитные явления 83, 85—98, 99, 244, 331, 334, 341, 351
— —, коэффициенты 270
— — поперечные, продольные 92, 99, 351
Гамильтона оператор (гамильтониан) 131—134, 139, 142, 190, 207
Генерация носителей 101, 104, 378, 379, 418
Групповая скорость 117, 179, 180
— — распространения тепловых колебаний 49
— — электронной волны 202, 236
Дебаевский радиус экранирования 288, 363
Дебая температура 41, 51, 52, 54, 55, 89, 96, 97, 187, 255, 276, 277, 283, 327
— функция 188
Дембера эффект 422
Дефекты решетки 34, 36, 37, 39, 48, 163—174, 319, 324
Диффузии коэффициент 69, 72, 364, 422
— — для ионов 172, 173
Диффузионная длина 72, 365, 381—383, 392, 394, 423
Диффузионная теория — см. выпрямление
Диффузия биполярная 82, 349, 357
— ионов 171
— носителей 68, 72, 74, 78, 363, 397, 421, 422
Диффузия электронно-дырочных пар 45
Длина свободного пробега 37, 39, 40, 42, 44, 48, 49, 54, 66, 68, 81, 94, 96, 97, 247, 249, 257, 261, 275—276, 279, 317, 324, 337, 343, 368, 373, 374

Емкость p - n перехода 280, 301, 364, 366, 376, 389—392
Жидкость фермиевская 239
—электронная 30, 238
Запорный слой 63, 64, 66, 72, 290, 363, 368—378, 423
Заселенность инверсная 433, 434, 436, 437, 440
Захват носителей 104, 105
Зеркального изображения силы 371, 372
Зона (энергетическая) 11—15, 56, 100, 209—216
— запрещенная 16, 18, 20, 32, 46, 108, 166, 202, 235, 286, 289, 306, 399, 416, 434, 435, 440
Импульс фонона 53, 54, 118, 254, 258
— фотона 108, 109, 405, 408—409
— электрона (см. также квазиимпульс) 27, 121, 190, 254, 259, 405—406
Импульса закон сохранения 54, 104, 108, 110, 111, 251, 253, 405, 435
Инжекция 74, 75, 362, 380—385, 394—395, 440
Интеграл перекрытия 207, 208
— столкновений 267, 269, 323
Ионизация двухступенчатая 410
— термоэлектронная 285—287, 373
— ударная 285, 289—291, 362, 365, 384, 392
— электростатическая 285, 362, 375, 384
Квазиимпульс фонона 253
— электрона 27, 214, 241, 253, 257
Квазистатический процесс 220
Квазиуровень Ферми 382—384, 417
Квазичастицы 219, 253
Квантовый выход 106, 112
Кванты энергии волны 114, 116
— — излучения 120
— — осциллятора 51, 53
Кикоина — Носкова фотомагнитный эффект 422—423
Кинетическое уравнение (Больцмана) 260—270, 340, 341
— — для фононов 318, 325, 328
Ковалентные кристаллы 89, 144, 147, 176, 196, 197
Когерентного излучения генераторы 142
Колебания атомов решетки, акустические 89, 181—183, 188, 317, 318
— — — оптические 41, 89, 112, 181—183, 189, 255, 276, 283, 314, 316, 317
— — — поляризация 176
— — — поперечные, продольные 176, 317, 318
— — — тепловые 14, 15, 29, 36, 37, 39, 40, 146, 148, 166, 174—184, 237, 238, 274—278, 409
Контактная разность потенциалов 58, 59, 63, 78, 367, 368, 381, 395

Контактные явления 55, 63, 362—399
Концентрация носителей тока 19, 22, 24, 30, 35, 44, 68, 77, 88, 100, 226, 229—234,
241, 278, 304—306, 418
— — , влияние поля 284— 291
— — — неравновесная 102— 103, 380, 382, 384
— электронов в металлах и полуметаллах 19, 30, 31, 43, 242
Край полосы поглощения 108
Кристаллическая решетка 36, 147-163, 171, 208, 211, 281
Лавинный эффект 363, 365
Лазер 104, 165, 436, 437, 440
Ландау уровни 340
Ловушки (см. также прилипание) 416—417, 420
Лоренца сила 84, 85, 264, 331, 333, 423
— число 44, 257, 3J6
Магнитная сила 84, 85, 86, 90, 91, 95, 97, 332, 334, 335
Магнитное поле сильное 94, 95, 338, 339, 341, 351—355
— — —, классический и квантовый критерии 340, 351
— — слабое 93, 94, 95, 337, 338, 341, 347
Маджи — Риги — Ледюка эффект 100
Максвелла — Больцмана распределение 35, 227, 239, 328, 347
— закон распределения для скоростей 15
Масса эффективная 24, 27, 35, 42, 120, 212, 213, 214, 277, 331, 392, 427
Мелкие уровни 120
Металлическая связь 144, 147
Металлы 13, 16, 18, 58—62, 78, 176, 237, 242, 298
—, схема валентной зоны 16
— щелочно-земельные 17, 238
— щелочные 16, 34, 199, 237
Миллеровские индексы 163
Модуль сжимаемости 49
Невырожденное состояние электронного газа 30, 31, 35, 44, 77, 81, 227, 277
Неопределенностей соотношения Гейзенберга 124, 162
Неопределенности принцип 123—128, 175, 249
Непрерывности уравнение 378—379, 389
Нернста — Эттинггаузена эффекты 98—100, 356—358, 423
Обратная решетка 158, 203, 205, 210
— —, вектор 161, 200, 203, 321, 322
— —, — базисный 162
Обратное пространство 159, 161, 162, 198, 210
— —, основная (или приведенная) область 160, 161
— —, элементарная ячейка 160
Обращенный диод 364, 397

Объемный заряд 62, 64
Одноэлектронное приближение 194—198, 209
Паули принцип 13, 56, 126, 175, 196, 199
Пельтье явление 63, 75, 76, 79, 292, 293, 300, 303
— коэффициент 76, 81, 82, 303
— теплота 80
Переноса явления 244—292, 413
— —, феноменологический анализ 270—271
Переход *p-i-n* 393—396
— электронно-дырочный 70, 71, 280, 362, 363, 366, 378—399, 423, 437, 440
— — на границе полупроводник — металл 65, 66
— — толстый, тонкий 72, 392—396
Переходы (носителей) — см. также «рекомбинация»
— безызлучательные 104, 194
— «вертикальные» 256, 405
— межзонные (прямые) 107, 108, 404—407, 435
— «непрямые» 406, 407, 435, 436
— при столкновениях горизонтальные, вертикальные 313—314
Пи-подход (в теории термоэлектричества) 294, 302
Писаренко формула 82
Плотность состояний (электронных) 28, 32, 33, 42, 227— 229, 308, 339, 382, 417, 418
Поглощение (излучения, света) 107, 400—409
— отрицательное 431, 432
— примесное 107, 108, 110
— свободными носителями 106, 112, 402—404, 435
— собственное (фундаментальное) 107, 109, 409
— —, красная граница 410
—, спектр 107, 408
— спектральное распределение 409
— экситонное 111, 407—410
Подвижность ионов 173, 174
Подвижность носителей тока 19, 36, 245, 256, 257, 260, 403, 410
— — —, температурная зависимость 19, 256, 261
— — —, экспериментальное определение 88, 97, 356
— —, зависимость от электрического поля 279—284
Полупроводники 13, 18
— вырожденные (см. также вырождение) 35, 89, 227, 357, 407
— ковалентные 41, 89
— невырожденные 40, 41, 92, 237, 418
— примесные 20, 31

- — дырочные 23, 24, 36, 232
- — электронные 21, 31, 35, 232
- с атомной решеткой 96, 98
- с ионной решеткой 41, 89, 96, 97
- собственные 20, 32
- Полярон 192—194, 249
- Потенциал ионизации (ионизационный) 11, 55, 56, 145—147, 166
- решетки периодический— 195—198, 200—206, 210, 214, 412
- химический (см. также уровень Ферми) 222, 223
- — приведенный 226, 241, 304
- —, уровень 33, 35, 57, 58, 71, 78, 80, 166, 227, 241, 294, 309, 381, 396, 431
- электростатический 157, 222
- Поток тепла 46, 98, 245, 318, 321, 326, 355
- энергии 245, 270, 329
- Преломления показатель (коэффициент) 112, 400, 401
- Приближение почти свободных электронов 198—206
- сильно связанных электронов 206—209
- Прилипание (см. также ловушки) 414
- , уровень 105
- Прицельное расстояние 248, 258
- Проводимость ионная 171—175
- Процессы нормальные («N» при фонон-фононных взаимодействиях) 321, 323, 326
- переброса (umclapp) 321—324, 326
- Работа выхода 55—58, 62, 64, 70, 78, 373
- Распределение фононов — см. Бозе
- — неравновесное 322, 323, 325, 329
- электронов по волновым векторам 266
- — — — —, неравновесная добавка 269—270, 296
- — — — —, — функция 266, 270, 315
- — — скоростям 262, 263, 344
- — — — —, изменение под действием градиента температуры 313, 314
- — — — —, — — — электрического поля 262, 263, 312, 313, 344
- — — — —, неравновесное 263—265, 312, 313, 345
- Рассеяние фононов 319
- — на фононах 319, 321, 324
- электронов 39
- — межзонное 310
- — на дефектах решетки 39, 245, 247, 314
- — на тепловых колебаниях решетки 40, 54, 81, 89, 261, 263, 274—278, 281, 282, 316
- — примесное 40, 81, 89, 97, 248, 251, 258, 261, 263, 271—274, 277, 281, 282, 394

Рекомбинация 18, 45, 72, 74, 101, 104, 363, 378, 392, 397, 413—421, 426
— безызлучательная 104
—, вероятность 102, 106
— излучательная 104
— квадратичная 414
— линейная 414
— поверхностная 72, 409, 424, 426,
— прямая (межзонная) 101, 104, 105
— ударная 106
— через центры 101, 104, 105, 416
— экситона 112
Риги — Ледюка эффект 98, 355—356
Симметрия кристалла 198
— трансляционная 111, 157, 158, 192, 193, 198, 209—211
Скорость дрейфа электронов 37, 38, 69
— теплового движения электронов 37, 39, 40, 78
— — — — средняя 43, 68, 254, 279
Скрещенные электрическое и магнитное поля 334, 335
Слой объемного заряда 63, 71, 383
— — — на контакте металл — полупроводник 64
— — —, толщина 380
Столкновение 248—251
Столкновения, вероятность 219, 246
— межэлектронные, неупругие, упругие, почти упругие 251, 252, 254—258, 268,
269, 283, 284, 314—316
Столкновения фононов с дефектами, с фононами 54, 219, 320
— — — —, 2-3-и 4-фононные 320
—, частота 219, 246
— электронов сфононами 54, 251, 253, 254, 258
Тепловое возбуждение электрона 15, 20
— — —, частота следования 15, 18
Теплоемкость решетки 43, 49, 50, 52, 184—189, 317
— фотонная 329
— электронного газа 44, 311
Теплопроводность 244, 256, 270, 271
— решеточная 14, 43, 46—49, 219, 317—329, 351
— фотонная 311, 329—330
— электронная 43—44, 46, 257, 264, 311—316, 351, 356
— —, отношение к электропроводности 44, 45
Терромагнитные коэффициенты 257, 270
Терромагнитные явления 83, 85, 98—100, 244, 331, 334, 355—357

Термоэлектрические явления 75, 76, 78, 82, 100, 244, 264 292—311, 315, 351, 422—423

Томсона коэффициент 77

— соотношения 76—77

— теплота 77

— явление 75, 77, 79

Туннельный диод 165, 280, 286, 364, 396—399

— эффект 11, 59, 62, 66, 193, 285, 363, 368, 373, 392, 396

Увлечение электронов фононами 299—304

Уровень (энергетический) см. энергетический

Уровень Ферми (см. также потенциал химический) 33, 34, 59, 71, 80, 81, 105, 433

— — в невырожденном полупроводнике 82, 229—232, 234

— — в невырожденном полупроводнике, температурная зависимость 35, 36, 232, 233

— —, температурная зависимость 35, 36

Фазовая скорость 177, 179, 180, 253

Фазовое пространство 265—267, 328

Ферми — Дирака статистика 199, 431

— интегралы 241, 243, 299

— поверхность 314, 339, 342

— распределение 32—34, 224—226, 328, 417, 431

Фонона энергия 53, 104, 118, 253—255

Фононы 49, 54, 104, 118, 186, 187, 255, 299—303, 317, 318, 324, 405—407

—, групповая скорость 317

Фотовольтаические эффекты 421—426

Фотодиод 425, 426

Фотоионизация 384

Фотоноситель 100, 413, 421

Фотопроводимость 100—103, 108, 409—421

— примесная, собственная 410—412

— стационарная 103

Фоточувствительность 107, 109, 110, 409, 410

Фото-э. д. с. 421

Фотоэффект 100, 109

— вентильный 423—426

—, квантовый выход 423

—, красная граница 410

— примесный 409

— «собственный» 409

Фундаментальное поглощение, см. поглощение

Химическая связь 143—147

Холла постоянная 87—91, 347, 349—352

— э. д. с. 90, 341, 349
— эффект 86, 92, 341—352, 356
Холловское поле 87, 90, 91, 346, 351, 354
Центр рекомбинации примесный 101, 105, 110
Цепочка двухатомная 180—182
Циклотронный резонанс 334, 426—430
Шокли — Рида статистика рекомбинации 416—421
Шредингера уравнение 128—134, 190, 274
— — для нестационарных процессов 131—138
— — для электронов в адиабатическом приближении 192
— — — — в одноэлектронном приближении 195
— — невозмущенное 135
Эйнштейна соотношение 69, 174, 370
Экситоны 106, 111, 196, 407—409
Экстракция 75
Электрическое поле сильное 278-291, 362
Электронный газ, вырожденное состояние 30—32, 34, 237—239, 256, 262, 406, 407, 435
— —, невырожденное состояние 30, 31, 35, 44, 77, 81, 227, 238, 239, 256, 262, 299
— — невырожденный, статистика (для полупроводников) 226—231
— —, плотность 36
Электронный газ, сохранение теплового равновесия с кристаллом 257
Электрон-фононное взаимодействие 255
Электрохимический потенциал 79, 224, 293
Элементарная ячейка — см. ячейка
Эллипсоид энергии 216
Энергетический уровень захвата 416
— — электрона в кристалле 236
Энергетический спектр электронов в кристалле 10—13, 20, 209, 236, 244, 249
Энергия активации 11, 19, 27, 31, 172, 339, 341, 407
— — оптическая 410, 412, 413
— — примесных электронов 22, 35, 120, 166, 233, 287, 368, 373
— — термическая 410, 412, 413
— тепловых колебаний атомов кинетическая средняя 49, 52
— — — — потенциальная средняя 50
— электронов в направленном потоке средняя 40, 80, 81, 293
— — тепловая (кинетическая) средняя 40, 242—243
Эттингсгаузена эффект 91, 92, 349—351
Ячейка (решетки кристаллической) элементарная 111, 152—155, 157, 160, 183

Несмотря на то, что полупроводниковые материалы и приборы уже давно вышли за пределы физических лабораторий, до настоящего времени в мировой литературе вряд ли можно найти книгу, посвященную физике полупроводников и рассчитанную на достаточно широкий круг читателей: инженеров и научных работников, занимающихся разработкой полупроводниковых материалов и приборов, и студентов соответствующих специальностей. Автор поставил своей целью восполнить этот пробел.

В гл. 1 описаны основные явления в полупроводниках и основные свойства полупроводников в качественной форме почти без математических выкладок, последующие восемь глав посвящены количественному анализу основных вопросов физики полупроводников, но и в них автор старался на первый план выдвинуть физическую сущность явлений, а затем уже вывести и основные математические соотношения.

По ряду основных проблем физики полупроводников уже имеются прекрасные монографии советских авторов и ряд переводных книг. Поэтому вопросам, рассмотренным в этих работах, в настоящей книге уделялось меньше места, чем тем, которые не освещены достаточно полно в литературе.

В основу книги положены лекции, прочитанные автором на радиотехническом факультете Ленинградского политехнического института им. М. И. Калинина, а также в ряде научно-исследовательских институтов. При работе над рукописью широко использованы монографии, список которых приведен в конце книги.

Автор считает своим приятным долгом выразить глубокую благодарность М. С. Соминскому за очень полезные советы и дискуссии при обсуждении отдельных разделов книги, а также Н. А. Пенину за ценные замечания при ее рецензировании.

Все замечания и пожелания читателей просьба посылать по адресу: г. Ленинград, Д-187, наб. Кутузова, 10, Институт полупроводников АН СССР, Л. С. Стельбансу.

- C — теплоемкость;
 c — скорость света;
 E — напряженность электрического поля;
 e — заряд электрона;
 \mathcal{E} — полная энергия;
 F — свободная энергия, сила;
 f — функция распределения, сила Лоренца, коэффициент упругой силы;
 g — кратность уровня, скорость генерации носителей;
 H — напряженность магнитного поля, оператор Гамильтона;
 h — постоянная Планка, $\hbar = h/2\pi$;
 I — ток;
 j — плотность тока;
 K — коэффициент поглощения;
 k — постоянная Больцмана;
 k — волновое число, $k = 1/\lambda$;
 L — диффузионная длина;
 l — длина свободного пробега, толщина слоя объемного заряда;
 m, M — масса электрона и атома (иона) соответственно;
 n, p — число свободных электронов и дырок в 1 см^3 ;
 Q — поток энергии;
 r — радиус-вектор;
 s, S — энтропия, площадь, путь;
 T — температура, кинетическая энергия;
 t — время;
 U — потенциальная энергия;
 u — подвижность;
 V — разность потенциалов;
 $V_{\text{к}}$ — контактная разность потенциалов;
 v — скорость частиц;
 W — работа выхода;
 ω — скорость распространения упругих колебаний, вероятность;
 α — термо-э. д. с.;
 ϵ — энергия частицы;
 κ — удельная теплопроводность, диэлектрическая проницаемость;
 λ — длина волны;
 μ — химический потенциал;

ν — частота, $\omega = 2\pi\nu$;

σ — удельная электропроводность;

τ — время релаксации, время свободного пробега;

θ — температура Дебая;

φ, ψ — угол, волновая функция;

ω — круговая частота.

ОСНОВНЫЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

1.1. НЕКОТОРЫЕ СВЕДЕНИЯ О СТРОЕНИИ АТОМА

Физические свойства полупроводников тесно связаны со структурой валентных оболочек атомов, из которых они образованы; поэтому начнем с краткого качественного рассмотрения этого вопроса.

Атом любого элемента состоит из положительно заряженного ядра с зарядом ze (где e — заряд электрона и z — порядковый номер атома в системе Менделеева) и z электронов, расположенных в виде системы оболочек с постепенно возрастающими размерами: линейные размеры оболочек соотносятся между собой как квадраты целых чисел. Первая оболочка носит название K -оболочки, вторая — L -оболочки, третья — M -оболочки и т. д.

Первая, K -оболочка, имеет только одну подоболочку, обозначаемую s , вторая (L) делится на s - и p -подоболочки, третья (M) — на s -, p - и d -подоболочки и т. д.

Мы уже упоминали о том, что размеры оболочек определяются их номером; подоболочки отличаются друг от друга конфигурацией (формой). В упрощенной модели электронного строения атома предполагается, что электрон подобен заряженному волчку, обладающему собственным и магнитным и механическим моментами (называемыми спинами), а каждый атом подобен солнечной системе, в которой электроны движутся вокруг ядра по эллиптическим орбитам, причем номер оболочки определяет длину большей полуоси орбиты, а номер подоболочки — ее форму (эксцентриситет). Каждая подоболочка, начиная со второй,

состоит из ряда орбит, отличающихся друг от друга своей ориентацией в пространстве; первая, *s*-подоболочка, имеет только одну орбиту, на которой могут разместиться два электрона с противоположно направленными спинами. Вторая подоболочка, *p*, состоит из трех орбит, на каждой из которых может разместиться по 2 электрона; таким образом, всего во второй подоболочке может разместиться 6 электронов; третья, *d*-подоболочка, состоит из пяти орбит, на которых может поместиться 10 электронов, и т. д. Номер оболочки называется также главным квантовым числом и обозначается *n*, номер подоболочки называется орбитальным или побочным квантовым числом и обозначается *l*. Побочное квантовое число определяет эксцентриситет орбиты, а следовательно, и орбитальный момент количества движения электрона, находящегося на ней. Полный момент количества движения электрона складывается из орбитального и собственного (спина) и определяется третьим квантовым числом *j*, которое в зависимости от ориентации спина может принимать значения $j_1 = l + 1/2$ и $j_2 = l - 1/2$. Наконец, ориентация орбиты в пространстве характеризуется четвертым магнитным квантовым числом *m*.

Согласно принципу Паули в атоме не может существовать двух или большего числа электронов с одинаковой четверкой квантовых чисел; этим законом и определяется число электронов, помещающихся на одной орбите (2), в одной подоболочке ($2l + 1$) и в одной оболочке — $2n^2$.

В таблице приведены максимальные числа электронов, которые могут разместиться в каждой подоболочке. В последнем столбце приведено полное число электронов, которое размещается в данной оболочке.

ТАБЛИЦА 1

| Номер оболочки (<i>n</i>) | Номер подоболочки | | | | | z_n |
|-----------------------------|-------------------|---------------|---------------|---------------|---------------|-------|
| | 0(<i>s</i>) | 1(<i>p</i>) | 2(<i>d</i>) | 3(<i>f</i>) | 4(<i>g</i>) | |
| 1 (<i>K</i>) | 2 | — | — | — | — | 2 |
| 2 (<i>L</i>) | 2 | 6 | — | — | — | 8 |
| 3 (<i>M</i>) | 2 | 6 | 10 | — | — | 18 |
| 4 (<i>N</i>) | 2 | 6 | 10 | 14 | — | 32 |
| 5 (<i>P</i>) | 2 | 6 | 10 | 14 | 18 | 50 |

Радиусы круговых орбит и энергии электрона на них могут быть найдены из двух условий:

1) условия классической устойчивости — равенства центробежной и центростремительной силы:

$$\frac{m_0 v^2}{r} = \frac{z e^2}{r^2}, \quad (1.1)$$

2) условия квантовой устойчивости. Согласно квантовой механике (см. подробнее § 2.1) электрон, движущийся со скоростью v , описывается волной, длина которой $\lambda = h/m_0 v$, где m_0 — масса электрона и h — универсальная постоянная, так называемая постоянная Планка, равная $6,57 \cdot 10^{-27}$ эрг/сек. Для того чтобы орбита была устойчива, на ней должно укладываться целое число волн (для того чтобы при многократном прохождении орбиты электронная волна находилась в фазе сама с собой, а не гасила самое себя):

$$2\pi r = n\lambda = \frac{nh}{m_0 v}. \quad (1.2)$$

Решая совместно (1.1) и (1.2), находим

$$r_n = n^2 \frac{\hbar^2}{m_0 e^2} \frac{1}{z}, \quad (1.3)$$

где $\hbar = h/2\pi$.

Величина r равна $a_0 = \hbar^2/m_0 e^2$ при $n=1$ и $z=1$, т. е. представляет собой радиус первой (самой близкой к ядру) орбиты в атоме водорода ($a_0 = 0,529 \cdot 10^{-8}$ см). Радиусы других орбит в атоме водорода или в одноэлектронных ионах с зарядом ядра z выражаются через a_0 :

$$r_n = n^2 \frac{a_0}{z}. \quad (1.4)$$

Полная энергия электрона на n -й орбите равна сумме его кинетической и потенциальной энергии:

$$\mathcal{E}_n = T_n + U_n = \frac{1}{2} m_0 v^2 - \frac{z e^2}{r}.$$

Используя (1.1) и (1.3), получаем

$$\mathcal{E}_n = -\frac{z e^2}{2r} = -\frac{m_0 e^4 z^2}{2\hbar^2 n^2}. \quad (1.5)$$

Одним из основных постулатов квантовой механики является предположение, что, находясь на стабильной орбите, электрон не излучает энергии, а при переходе с одной орбиты на другую излучает (или поглощает) квант энергии, величина которого равна разности энергий электрона в начальном и конечном состояниях:

$$h\nu = \mathcal{E}_n - \mathcal{E}_{n'} = \frac{m_0 e^4 z^2}{2\hbar^2} \left(\frac{1}{n^2} - \frac{1}{n'^2} \right). \quad (1.6)$$

Формула (1.6) дает точное значение частот спектра испускания и поглощения атома водорода; этот факт был первым триумфом квантовой теории атома Бора, основанной на модели круговых орбит. Однако теория Бора не только имела целый ряд внутренних противоречий, но и не объясняла особенностей спектров более сложных атомов, требовавших учета эллиптических орбит, релятивистских эффектов спина и магнитного момента электрона. Соответствующие дополнения к теории были сделаны Зоммерфельдом и Вильсоном, Уленбеком и Гоудсмитом; мы ограничимся здесь ссылкой на популярное изложение этих вопросов в книге Де Бройля [5].

ВАЛЕНТНОСТЬ

По мере роста порядкового номера элемента заполняется сначала первая, *K*-оболочка (самые нижние энергетические уровни), затем последовательно *s*- и *p*-уровни второй, *L*-оболочки, и т. д. В водороде имеется один электрон на *s*-уровне *) первой оболочки, что записывается следующим образом: $1s_1$. (Первая цифра означает номер оболочки, *s* определяет уровень подоболочки, и индекс справа — число электронов на ней.) В гелии имеются два электрона, которые целиком заполняют первую оболочку (в соответствии с этим его электронная структура записывается как $1s_2$).

Два электрона, движущиеся по одной и той же орбите, образуют так называемую электронную пару. Спаренное состояние электронов энергетически более устойчиво, чем неспаренное. Поэтому, вступая в соединение, атомы стремятся таким образом изменить число и характер движения электронов внешней оболочки, чтобы не осталось

*) Или, что то же самое, на орбите или в *s*-подоболочке.

неспаренных. В ряде случаев неспаренные электроны отдельных атомов вступают во взаимодействие друг с другом и изменяют характер своего движения таким образом, что начинают вращаться по новым молекулярным орбитам в виде электронных пар. Такие соединения называются ковалентными.

Однако число неспаренных электронов в изолированных атомах очень невелико — один или два; поэтому нередко и такие случаи, когда внутриатомные пары разрушаются и электроны изолированных атомов в ковалентном соединении образуют новые молекулярные пары. Поэтому в ковалентном соединении предельная валентность элемента определяется числом электронов во внешней незаполненной оболочке.

В других случаях, входя в соединение, атом либо стремится полностью отдать электроны внешней оболочки, либо увеличить их число до полного ее заполнения. Такие соединения называются ионными. Отсюда следует, что в ионных соединениях максимальная валентность электронов определяется числом электронов (валентность по кислороду) или числом пустых мест (валентность по водороду) на внешней оболочке. Этим объясняется с точки зрения современной физики периодический закон, открытый Менделеевым.

Здесь следует заметить, что между ковалентными и ионными соединениями не существует резкой границы. Подробнее этот вопрос будет рассмотрен в § 2.3.

1.2. ЭНЕРГИЯ И ДВИЖЕНИЕ ЭЛЕКТРОНА В ТВЕРДОМ ТЕЛЕ

ЭНЕРГЕТИЧЕСКИЙ СПЕКТР ЭЛЕКТРОНОВ В КРИСТАЛЛЕ

Посмотрим теперь, как изменяется характер движения и энергетический спектр валентных электронов в твердом теле. При образовании твердого тела соседние атомы настолько сближаются друг с другом, что внешние электронные оболочки не только соприкасаются, но даже перекрываются. В результате этого характер движения электронов резко изменяется: электроны, находящиеся на определенном энергетическом уровне одного атома, получают возможность переходить без затраты энергии на соответ-

ствующий уровень соседнего атома и таким образом свободно перемещаться вдоль всего твердого тела.

С точки зрения классической физики для перехода с орбиты определенного атома на соответствующую орбиту соседнего атома электрон должен преодолеть потенциальный барьер, имея необходимую для этого энергию, называемую энергией активации. Эта энергия затрачивается при отрыве электрона от первого атома и возвращается при связывании электрона со вторым атомом; таким образом, энергия электрона в начальном и конечном состояниях одинакова.

Энергия активации, или, что то же самое, высота потенциального барьера, зависит от расстояния между атомами; при больших расстояниях она равна энергии связи электрона на соответствующей орбите — потенциалу ионизации, при уменьшении расстояния эта энергия начинает уменьшаться и, наконец, когда орбиты перекрываются, обращается в нуль.

Согласно представлениям квантовой физики наряду с описанным выше возможен и другой механизм перехода электрона от атома к атому — без затраты и возврата энергии. Поскольку в этом случае электрон как бы «просачивается» через потенциальный барьер, это явление получило название туннельного эффекта. Вероятность туннельного эффекта экспоненциально возрастает по мере уменьшения высоты потенциального барьера и его ширины, т. е. расстояния между атомами. Поэтому при больших расстояниях эта вероятность исчезающе мала, а при малых (10^{-7} — 10^{-8} см) становится настолько большой, что это явление приобретает решающее значение во всех электронных переходах, осуществляющихся внутри твердого тела. Более того, если два твердых тела сближаются настолько, что расстояние между ними становится меньше чем 10^{-6} см, то в некоторых случаях надо учитывать туннельные переходы из одного тела в другое.

Вместо индивидуальных атомных орбит образуются коллективные, и подоболочки отдельных атомов объединяются в единый для всего кристалла коллектив — зону. Однако при этом число электронов, которые могут разместиться в данных подоболочках атомов, образующих кристалл, остается неизменным.

Если для изолированного атома это число было $2g$ (для s -подоболочки $g = 1$, p — $g = 3$ и т. д.) и число атомов в кристалле N , то общее число мест в зоне по-прежнему равно $2Ng$. Однако энергетический спектр электронов при этом претерпевает существенные изменения. Если в изолированном атоме каждой подоболочке соответствует строго определенная энергия, то зона состоит из Ng уровней, расположенных очень близко друг к другу, причем на каждом уровне, в соответствии с принципом Паули, могут разместиться только два электрона, движущиеся в противоположных направлениях. Различным уровням одной

зоны соответствует различная скорость поступательного движения электрона по кристаллу (кинетическая энергия) и различная потенциальная энергия, чем и объясняется различное положение уровней в зоне.

Таким образом, строго определенные значения энергии, соответствующие отдельным уровням в изолированном

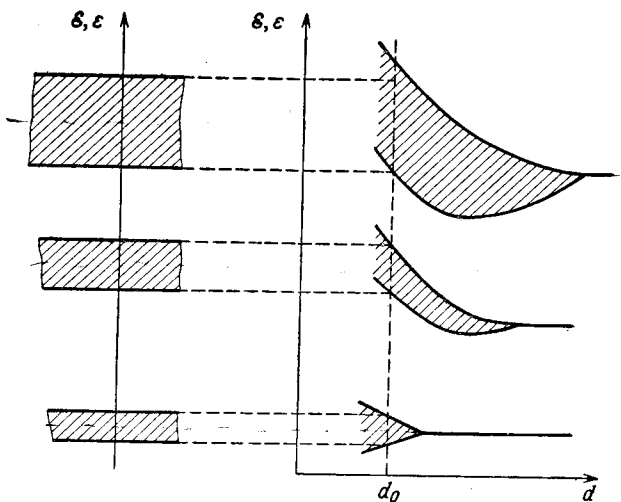


Рис. 1.1. Образование энергетических зон в кристалле из атомных энергетических уровней:

d — расстояние между соседними атомами; d_0 — равновесное расстояние между соседними атомами в кристалле.

атоме, заменяются в кристалле целым интервалом энергий — на языке квантовой механики это означает, что энергетические уровни атомов в кристалле под влиянием взаимодействия объединяются и расщепляются в зоны; в этом и состоит наиболее характерное отличие энергетического спектра кристалла от спектра изолированного атома.

Кроме этого, каждый уровень при сближении атомов несколько смещается, так как электрон теперь взаимодействует не с одним атомом, а со всеми атомами кристалла.

На рис. 1.1 представлена зависимость энергетического спектра электрона от расстояния между атомами: справа

нанесены три энергетических уровня, соответствующих изолированным атомам, левее показано, как уровни сдвигаются и расширяются по мере сближения атомов.

ДИЭЛЕКТРИКИ, ПОЛУПРОВОДНИКИ, МЕТАЛЛЫ

Мы уже говорили о том, что в изолированном атоме в каждой подоболочке помещается строго определенное число электронов; при сближении атомов спектр электрона существенно модифицируется: из локальных орбит изолированных атомов образуются зоны, единые для всего кристалла, но при этом общее число электронов, которые могут разместиться в данных подоболочках атомов, образовавших кристалл, остается неизменным. Если для изолированного атома это число составляло g (для s -подоболочки $g = 1$, для p -подоболочки $g = 3$ и т. д.) и число атомов в кристалле N , то соответствующая зона состоит из $2Ng$ состояний, расположенных попарно очень близко одно под другим, причем каждому состоянию соответствует другое, в котором электрон обладает той же энергией, но движется в противоположном направлении. Таким образом, число электронов, которые могут разместиться в данной зоне, равно общему числу мест на уровнях изолированных атомов, из которых она образовалась.

Как мы уже упоминали, внутренние оболочки в изолированных атомах целиком заполнены; так как число мест в зоне кристалла остается тем же, эти же условия должны сохраниться в соответствующих зонах кристалла.

Отсюда следует важный вывод: электроны внутренних оболочек не могут переносить электрический ток. Действительно, на каждом энергетическом уровне зоны находятся два электрона, движущиеся в противоположных направлениях, значит, суммарный переносимый ими ток равен нулю. С другой стороны, даже под воздействием поля ни один электрон не может изменить состояния своего движения, так как все соседние уровни заняты, а в силу принципа Паули на каждом уровне могут поместиться только два электрона.

Иначе может обстоять дело в самой верхней зоне, образовавшейся из уровней, на которых располагались валентные электроны; электропроводность кристаллов в основном и определяется степенью заполнения валентной зоны и ее расстоянием до следующей пустой зоны. Рассмотрим

случаи, которые встречаются в действительности и определяют разделение твердых тел на диэлектрики, полупроводники и металлы.

Валентная зона заполнена целиком, расстояние до следующей зоны велико — диэлектрики. В этом случае (рис. 1.2) валентная зона ничем не отличается от расположенных ниже; не выходя за ее пределы, электроны

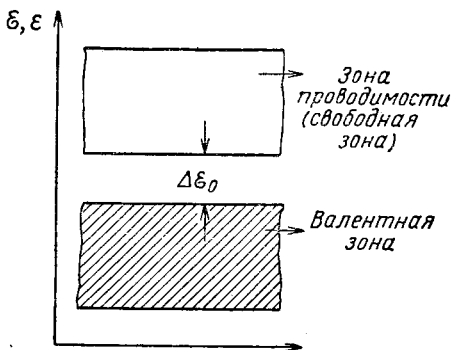


Рис. 1.2. Схема заполнения энергетических зон в диэлектрике и полупроводнике.

не могут участвовать в электрическом токе. Формулировка «расстояние до следующей зоны велико» также требует некоторого уточнения.

Характер теплового движения атомов в твердом теле и в газе существенно различается: в газе атомы хаотически движутся в пределах всего объема, в твердом теле — колеблются около равновесных положений, так как в этом случае атомы связаны друг с другом силами сцепления, колебания одного атома передаются другому и вдоль всего кристалла распространяются в самых разнообразных направлениях упругие волны. Если в кристалле существует перепад температур, то эти волны переносят тепло от горячего конца к холодному; таков механизм теплопроводности кристаллической решетки твердого тела (подробнее см. гл. 6).

Если температура всего кристалла одинакова, то одинакова и средняя кинетическая энергия тепловых колебаний атомов, и при достаточно высоких температурах она равна $\frac{3}{2} kT$, что при комнатной температуре составляет

0,04 эв *). Но кинетическая энергия атомов газа только в среднем равна $\frac{3}{2}kT$. Мгновенные же скорости распределяются по закону Максвелла; всегда имеется некоторое число атомов, скорости которых намного больше и намного меньше средних; вероятность того, что атом имеет энергию \mathcal{E}_0 , пропорциональна $e^{-\mathcal{E}_0/kT}$. Это же относится и к тепловым колебаниям твердого тела; в каждый данный момент имеется небольшое число атомов, амплитуда и энергия колебаний которых значительно больше средней.

Атомы при своих колебаниях взаимодействуют не только друг с другом, но и с электронами; колеблющийся атом может передать всю или часть своей энергии электрону. В результате этого энергия электрона увеличится и он перейдет на более высокий энергетический уровень; все это, разумеется, возможно, если уровень, на который должен перейти электрон в результате взаимодействия («столкновения») с атомом, свободен.

Этот акт взаимодействия мы будем называть в дальнейшем тепловым возбуждением электрона, в результате которого электрон может «перескочить» из заполненной валентной зоны в свободную, называемую также зоной проводимости, и уже сможет участвовать в электрическом токе. Появившееся при этом в заполненной зоне пустое место, которое мы в дальнейшем будем называть «дыркой», создает в ней возможность эстафетного (дырочного) механизма проводимости: какой-либо электрон занимает освободившееся место, его место занимает другой и т. д.

Число тепловых возбуждений в секунду (P) пропорционально числу электронов вблизи верхнего края валентной зоны C_v , числу пустых мест вблизи нижнего края свободной зоны C_c и вероятности того, что какой-либо атом приобретает при своих колебаниях энергию $\Delta\mathcal{E}_0$, достаточную для переброса электрона через запрещенную зону (рис. 1.2) в свободную. Согласно сказанному выше эта вероятность пропорциональна экспоненциальному множителю $e^{-\Delta\mathcal{E}_0/kT}$; следовательно, $P = aC_cC_v e^{-\Delta\mathcal{E}_0/kT}$, где a — некоторый коэффициент, зависящий от частоты «столкновений» атомов и электронов **).

*) 1 эв — это кинетическая энергия, которую приобретает электрон, ускоренный разностью потенциалов в 1 в.

**) Из рис. 1.2 видно, что наиболее вероятны забросы электрона с одного из верхних уровней валентной зоны на один из нижних уровней свободной зоны (так как такие забросы требуют мини-

Чем больше ширина запрещенной зоны, тем меньше вероятность этого процесса (а следовательно, и проводимость рассматриваемого кристалла).

Оценка показывает, что при ширине запрещенной зоны $\Delta \mathcal{E}_0 > 2\varepsilon \hbar$ эта вероятность остается исчезающе малой при всех доступных нам температурах и поэтому вещества, в которых выполняется это неравенство, не проводят электрический ток и называются диэлектриками; разумеется, граница $\Delta \mathcal{E}_0 = 2$ условна: и на самом деле между диэлектриками и полупроводниками, о которых речь будет ниже, нет резкой грани.

Валентная зона заполнена частично либо перекрывается со следующей свободной зоной — металлы. Мы уже упоми-

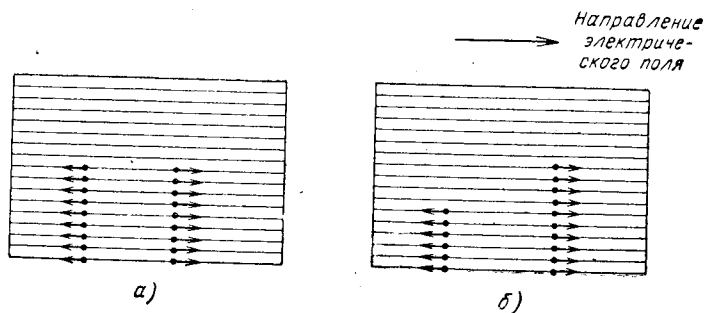


Рис. 1.3. Схема распределения электронов в валентной зоне щелочного металла:

а — в отсутствие электрического поля; б — при электрическом поле.

нали выше, что атомы благородных газов отличаются тем, что в них завершается заполнение очередной оболочки — в гелии первой, неоне второй и т. д. В щелочных металлах, следующих в периодической системе за ними, в первой s -подоболочке следующей оболочки имеется один электрон; всего в s -подоболочке есть два разрешенных места, и она в данном случае заполнена наполовину. Следовательно, валентная зона соответствующих кристаллов также заполнена наполовину; так как электроны стремятся занять энергетически наиболее выгодные уровни,

мальной затраты энергии). Поэтому C_v не полное число электронов в валентной зоне, а число электронов на ее верхних уровнях и соответственно C_p — число нижних уровней свободной зоны. (Подробнее мы рассмотрим этот вопрос в гл. 4).

то в нижней половине зоны на каждом уровне размещаются два электрона, а верхняя половина зоны оказывается пустой (рис. 1.3, а).

Так обстоит дело в отсутствие электрического поля, и поскольку на каждом уровне имеются два электрона с противоположно направленными скоростями, то суммарный ток через кристалл равен нулю. Если мы прикладываем к кристаллу разность потенциалов, то картина меняется: часть электронов, скорости которых были направ-

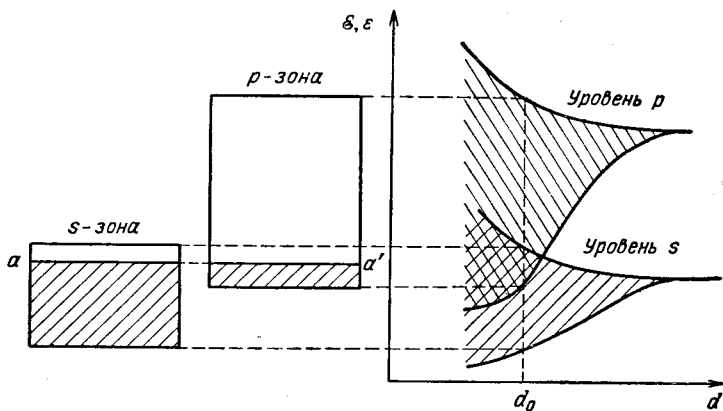


Рис. 1.4. Схема образования перекрывающихся *s*- и *p*-зон в щелочно-земельных металлах.

лены противоположно полю, под его действием изменяют направление своего движения (рис. 1.3, б), переходя на свободные уровни в верхней половине зоны, — в кристалле появляется электрический ток. Чем больше разность потенциалов, тем больше электронов изменяют направление своей скорости — электрический ток растет пропорционально приложенному напряжению. Таков механизм проводимости щелочных металлов.

В щелочно-земельных металлах на *s*-уровне находится по два электрона и поэтому *s*-зона заполнена целиком, но решетка этих металлов построена таким образом, что *s*-зона перекрывается со следующей за ней свободной *p*-зоной (рис. 1.4). Электроны из верхней части *s*-зоны переходят в *p*-зону таким образом, что обе зоны оказываются заполненными до одного и того же энергетического уровня

(a' — на рис. 1.4). Благодаря наличию пустых мест (дырок) в s -зоне и некоторого количества электронов в p -зоне кристалл становится проводником электричества.

Теория и опыт показывают, что во всех металлах и металлических сплавах валентная зона либо заполнена лишь частично, либо перекрывается следующей свободной зоной, поэтому все металлы и сплавы хорошо проводят электрический ток.

Валентная зона заполнена целиком, но расстояние до следующей свободной зоны мало (меньше $2\frac{1}{2}e\bar{v}$) — полупроводники. При абсолютном нуле валентная зона полупроводника заполнена целиком, свободная зона пуста и электропроводность равна нулю. С повышением температуры экспоненциально растет число тепловых забросов электронов, следовательно, растет число электронов в свободной зоне и дырок в заполненной и соответственно электропроводность полупроводника.

Выше мы говорили, что число тепловых забросов в секунду, которое мы в дальнейшем обозначим g_0 , растет экспоненциально с температурой:

$$g_0 = \alpha e^{-\frac{\Delta\mathcal{E}_0}{kT}}, \quad (1.7)$$

где $\alpha = aC_c C_v$; с другой стороны, если уже имеется некоторое число электронов в свободной зоне и дырок в заполненной, то одновременно идет обратный процесс — рекомбинация электронов и дырок: часть электронов падает на пустые места и сразу же исчезают один свободный электрон и одна дырка. Число рекомбинаций в секунду пропорционально вероятности «встречи», т. е. произведению числа электронов n и числа дырок p :

$$N_{\text{рек}} = \gamma np = \gamma n^2, \quad (1.8)$$

так как в рассматриваемом случае $n = p$.

Если кристалл находится при данной температуре достаточно долго, то должно установиться равновесие:

$$\alpha e^{-\frac{\Delta\mathcal{E}_0}{kT}} = \gamma n^2. \quad (1.9)$$

Отсюда

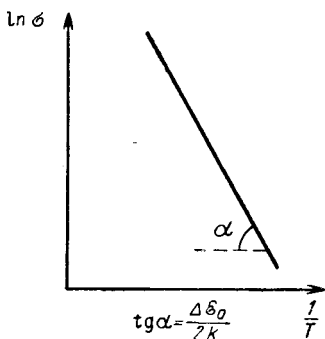
$$n = p = \sqrt{\frac{\alpha}{\gamma}} e^{-\frac{\Delta\mathcal{E}_0}{2kT}}. \quad (1.10)$$

Температурная зависимость электропроводности любого материала определяется температурными зависимостями концентрации носителей тока и их подвижности; подвижность электронов в полупроводниках, так же как и в металлах, зависит от температуры обычно по степенному закону:

$$u \sim T^{-a}, \quad (1.11)$$

где для металлов $0 \leq a \leq +1$ и для полупроводников $-3/2 \leq a \leq 3$. Но в металле концентрация свободных электронов неизменна, и поэтому температурная зависимость электропроводности целиком определяется подвиж-

Рис. 1.5. Температурная зависимость электропроводности собственного полупроводника (в полупологарифмических координатах).



ностью. В полупроводниках же n очень сильно зависит от температуры и по сравнению с этой зависимостью температурная зависимость подвижности играет слабую роль *).

Поэтому можно считать, что электропроводность полупроводников в первом приближении растет с температурой примерно по тому же закону, что и концентрация электронов и дырок:

$$\sigma \approx A e^{-\frac{\Delta \mathcal{E}_0}{2kT}}. \quad (1.12)$$

Итак, с феноменологической точки зрения полупроводники отличаются от металлов тем, что в металлах с повышением температуры электропроводность относительно медленно падает, а в полупроводниках — очень быстро растет. Температурная зависимость электропроводности полупро-

*) За исключением тех случаев, когда энергия активации $\Delta \mathcal{E}_0$ мала.

водника изображается обычно на графиках в так называемых полулогарифмических координатах; если прологарифмировать выражение (1.12), то оно примет вид

$$\ln \sigma = \ln A - \frac{\Delta \mathcal{E}_0}{2k} \frac{1}{T}. \quad (1.13)$$

Следовательно, если на графике по оси ординат откладывать $\ln \sigma$, а по оси абсцисс — $1/T$, то получим прямую с наклоном $\Delta \mathcal{E}_0/2k$ (рис. 1.5). Таким образом, зная наклон этой прямой, можно получить важнейшую характеристику полупроводника — ширину запрещенной зоны.

ПРИМЕСНЫЕ ПОЛУПРОВОДНИКИ

Полупроводники, которые мы рассмотрели в предыдущем разделе этого параграфа, называются собственными полупроводниками, а их проводимость — собственной проводимостью. Название это происходит оттого, что механизм появления электронов и дырок, рассмотренный выше, определяется характером собственного энергетического спектра кристалла и ничем не связан с содержанием примесей.

Наряду с этим имеется широкий класс веществ, в которых концентрация носителей определяется примесями; эти вещества называются примесными полупроводниками.

Рассмотрим механизм образования свободных электронов и дырок в этом случае. Если в решетку кристалла входит атом другого вещества, то часть энергетических уровней этого атома попадает в запрещенный промежуток между свободной и заполненной зоной. При этом надо различать два важных случая.

1. Энергетический уровень, занятый валентным электроном, оказался размещенным недалеко (на расстоянии $\Delta \mathcal{E}_1$) от нижнего края зоны проводимости (рис. 1.6). Электрон, находясь на этом уровне, не может участвовать в проводимости, но энергия теплового возбуждения, необходимая для перевода этого электрона в свободную зону, относительно мала; поэтому такие переходы будут играть основную роль при низких температурах, при которых вероятность теплового возбуждения из заполненной зоны, а следовательно, и собственная проводимость ничтожно малы. Можно показать*), что число электронов, попавших

*) Мы не приводим этот вывод, так как он в точности повторяет сделанный выше для собственной проводимости.

в зону проводимости с примесных уровней, выражается формулой

$$n_1 = \alpha_1 e^{-\frac{\Delta \mathcal{E}_1}{2kT}}, \quad (1.14)$$

где α_1 — коэффициент, пропорциональный корню из концентрации примесных атомов N_1 и C_c :

$$\alpha_1 \sim \sqrt{N_1 C_c}. \quad (1.14a)$$

Таким образом, при низких температурах в проводимости такого кристалла доминирующую роль играют электроны, перешедшие в зону проводимости с примесных

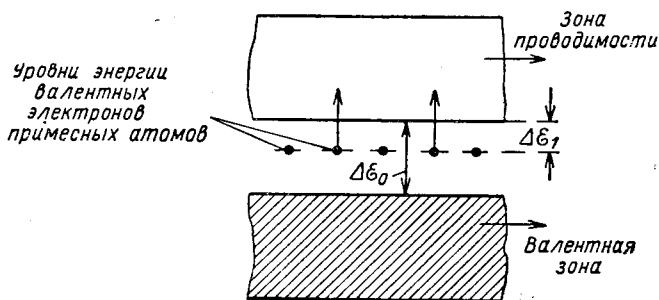


Рис. 1.6. Энергетический спектр примесного электронного полупроводника.

уровней (число дырок и электронов собственной проводимости будет во много раз меньше). Такой механизм проводимости называется *примесным*, а полупроводник в этом случае — *примесным электронным* полупроводником (в отличие от примесных дырочных полупроводников, о которых речь будет ниже). Атомы, отдающие электроны в зону проводимости, называются *донорами*, а соответствующие энергетические уровни — *донорными* уровнями.

Если прологарифмировать выражение (1.14), то оно примет вид

$$\ln n_1 = \ln \alpha_1 - \frac{\Delta \mathcal{E}_1}{2k} \frac{1}{T}, \quad (1.15)$$

соответствующий график, следовательно, имеет вид прямой с наклоном $\Delta \mathcal{E}_1 / 2k$.

Однако формула (1.15) справедлива лишь до тех пор, пока число электронов в зоне проводимости много меньше, чем общее число электронов на донорных уровнях. При достаточно высокой температуре все электроны с донорных уровней перейдут в зону проводимости и, следовательно, концентрация свободных электронов будет постоянна вплоть до тех температур, при которых начнет проявляться собственная проводимость. Таким образом, температур-

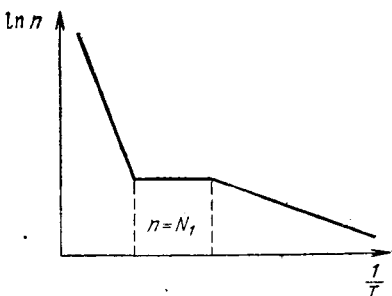


Рис. 1.7. Температурная зависимость концентрации электронов в примесном полупроводнике.

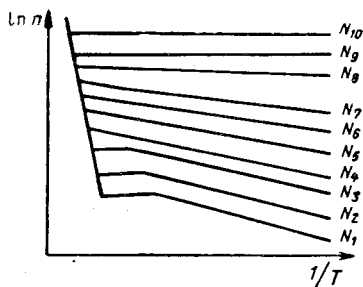


Рис. 1.8. Температурная зависимость концентрации носителей для ряда образцов с различным содержанием примесей.

ная зависимость концентрации свободных электронов в широком интервале температур будет иметь вид, представленный на рис. 1.7; при высоких температурах примесный полупроводник становится собственным.

На рис. 1.8 дана температурная зависимость концентрации носителей для ряда образцов с различным содержанием примесей ($N_1 < N_2 < N_3 \dots$).

Как видно из рисунка, при некоторой концентрации примесей (кривая N_4) горизонтальный участок кривой исчезает; это значит, что возбуждение электронов в заполненной зоне началось раньше, чем исчерпались электроны на примесных уровнях. При дальнейшем повышении содержания примесей наклон прямой на участке примесной проводимости начинает уменьшаться — энергия активации примесных электронов начинает падать и, наконец, начиная с некоторой концентрации (N_8), обращается в нуль; концентрация свободных электронов остается

постоянной начиная от самых низких температур и до температур, при которых начинается собственная проводимость. Такие вещества называются полуметаллами, при низких температурах они ведут себя как металлы (так как $n = \text{const}$, то электропроводность падает с ростом температуры за счет уменьшения подвижности), при высоких — как полупроводники.

Содержание примесей, при котором примесный полупроводник превращается в полуметалл, для различных материалов различно, но в большинстве случаев колеблется в пределах от 0,001% до 0,01%, что соответствует концентрации электронов 10^{17} — 10^{18} в 1 см^3 .

2. Уровень, занятый валентным электроном примесного атома, расположен ниже верхнего края заполненной зоны; следующий за ним свободный уровень расположен немного выше верхнего края заполненной зоны. Валентные электроны примесных атомов, попав в общую массу валентных

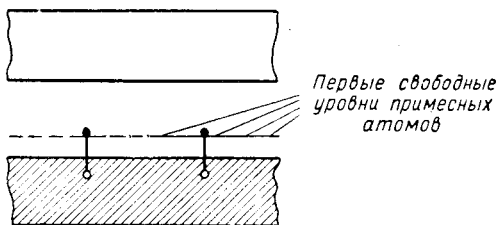


Рис. 1.9. Энергетический спектр примесного дырочного полупроводника.

электронов кристалла, не будут сколько-нибудь заметно влиять на электропроводность. В этом случае оказывается весьма существенной роль первых свободных уровней примесных атомов (рис. 1.9). При температуре, отличной от абсолютного нуля, часть электронов из заполненной зоны будет забрасываться на свободные уровни примесных атомов, в заполненной зоне появятся пустые места — дырки — и станет возможной дырочная проводимость.

В электрическом поле электроны движутся от минуса к плюсу, следовательно, дырки движутся в обратном направлении, т. е. ведут себя как положительные заряды. Таким образом, сложное эстафетное движение электронов мы можем рассматривать как перемещение фиктивной

положительно заряженной частицы — дырки. В соответствии со сказанным выше, такой механизм проводимости называется *дырочным*, а этот тип полупроводников — *примесными дырочными* полупроводниками. Энергетические уровни, на которые забрасываются электроны из заполненной зоны, называются *акцепторными*; атомы, которым принадлежат эти уровни, — *акцепторами*. Температурная зависимость концентрации дырок дырочного полупроводника имеет тот же вид, что и электронного (рис. 1.7 и 1.8). При низких температурах экспоненциально растет концентрация дырок, и затем, когда все акцепторные уровни уже забиты, она делается постоянной вплоть до той температуры, при которой начинает проявляться собственная проводимость. Точно так же при определенном содержании примесей дырочный полупроводник превращается в дырочный полуметалл.

ЭФФЕКТИВНАЯ МАССА

В предыдущих параграфах мы часто использовали термин «свободные электроны», имея в виду при этом электроны, которые могут перемещаться под действием электрического поля, т. е. принимать участие в электрическом токе. Не следует, однако, понимать этот термин буквально. Свободной мы с полным правом можем называть частицу, которая настолько удалена от других тел, что не испытывает практически никакого воздействия и находится, в согласии с первым законом Ньютона, в состоянии покоя или прямолинейного и равномерного движения. Если же частица движется в поле других частиц и испытывает их притяжение, то свободной мы ее называем в том случае, если ее кинетическая энергия больше энергии связи с частицами и поэтому она движется по незамкнутой траектории, т. е. с течением времени удалится на сколько угодно большое расстояние от частиц, в поле которых она находится.

Второй случай есть и в известном смысле обобщение первого: «когда-то давно» частица была достаточно удалена от других тел и двигалась почти прямолинейно и равномерно. При своем перемещении она приблизилась к другим телам, попала в сильные поля, траектория ее искривилась и скорость стала меняться от точки к точке. Но это лишь «эпизод» в движении частицы, пройдет некоторое время и она опять выйдет из сферы действия сильных

полей и будет двигаться прямолинейно и равномерно с той же скоростью, но в другом направлении.

Совершенно в иных условиях находится электрон проводимости в кристалле. Электрон, по сути дела, при своем движении все время связан с каким-либо атомом (причем энергия связи больше его кинетической энергии) и лишь благодаря тому, что решетка построена симметрично и из одинаковых атомов и энергия электрона не зависит от того, с каким именно атомом он связан, электрон и получает возможность перемещаться от одного атома к другому. При этом энергия связи периодически меняется: при приближении к положительно заряженному ядру она очень сильно возрастает, при удалении — убывает; поэтому периодически меняется и скорость движения. Лишь средняя скорость остается постоянной. Это следует из закона сохранения энергии: полная энергия электрона остается постоянной, так как нет внешних сил, которые бы совершали работу. В аналогичных точках кристалла потенциальная энергия электрона одинакова, следовательно, должна быть одинаковой и кинетическая энергия. Поэтому если не обращать внимания на периодические колебания скорости, а говорить о средней, то можно считать, что электрон в кристалле в отсутствие внешнего поля движется прямолинейно и равномерно, в известном смысле, как свободная частица в отсутствие внешних сил.

Когда к кристаллу прикладывается разность потенциалов, то характер движения электрона должен измениться; если бы он был «по-настоящему свободным», то вопрос бы решался просто согласно закону сохранения энергии: работа силы F на пути S должна равняться изменению кинетической энергии:

$$FS = \frac{m_0 v_2^2}{2} - \frac{m_0 v_1^2}{2} \quad (1.16)$$

или

$$F \frac{v_1 + v_2}{2} \tau = \frac{m_0 (v_2 + v_1)(v_2 - v_1)}{2}, \quad (1.17)$$

где τ — время, за которое электрон пройдет путь S ; отсюда, сократив левую и правую части на $(v_1 + v_2)/2$ и поделив на τ , сразу же получаем второй закон Ньютона:

$$F = m_0 a \quad (1.18)$$

или

$$a = \frac{F}{m_0} . \quad (1.19)$$

Разумеется, закон сохранения энергии соблюдается и для электронов в кристалле, но дело осложняется тем, что в данном случае при изменении траектории движения электрона под действием внешнего поля меняется одновременно его потенциальная и кинетическая энергия и только одна часть работы сил поля переходит в кинетическую энергию, а другая — в потенциальную.

При этом движение электрона может изменяться под действием поля таким образом, что потенциальная энергия его уменьшится, а, следовательно, увеличение его кинетической энергии будет больше работы сил поля (в кинетическую энергию перейдет также часть потенциальной); в этом случае электрон будет вести себя как очень легкая частица, т. е. частица с массой, меньшей массы свободного электрона. Может случиться, что почти вся работа пойдет на увеличение потенциальной энергии, скорость электрона увеличится так незначительно, как будто его масса очень велика. Может быть и так, что увеличение потенциальной энергии будет больше работы внешних сил, т. е. в потенциальную энергию перейдет и часть кинетической — скорость электрона уменьшится и он будет вести себя как частица с отрицательной массой.

Ускорение электрона, вообще говоря, может быть не пропорционально силе поля, а связано с ним более сложной функциональной зависимостью.

Все эти возможности действительно реализуются в различных кристаллах для электронов, находящихся на различных уровнях в зоне проводимости и валентной зоне; из этого можно заключить, насколько сложен характер движения электрона в кристалле. Однако существуют и здесь некоторые общие закономерности, которые позволяют упростить анализ механизма электропроводности.

Для электронов, находящихся вблизи нижнего края зоны проводимости, ускорение (на длине свободного пробега) пропорционально приложенной силе

$$F = ma$$

или

$$a = \frac{F}{m} . \quad (1.20)$$

где m — некоторый положительный коэффициент пропорциональности, различный для различных кристаллов.

В силу формальной аналогии между выражением (1.20) и законом Ньютона (1.19) m называют эффективной массой электрона в кристалле.

Из всего сказанного следует, что эффективная масса может быть и меньше и больше массы свободного электрона, что действительно имеет место в различных кристаллах. Из сказанного выше следует также, что эффективная масса может зависеть от направления движения электрона, что также наблюдается в большинстве полупроводниковых кристаллов. В этих случаях зависимость энергии электрона от его импульса*) изображается не сферой $\varepsilon = p^2/2m$, как для свободного электрона, а эллипсоидом

$$\varepsilon = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + \frac{p_3^2}{2m_3}. \quad (1.21)$$

Во многих случаях зонная структура носит еще более сложный характер. Так, в зоне проводимости германия имеется четыре одинаковых эллипсоида вращения, в зоне проводимости кремния — три. В ряде кристаллов валентные и свободные электроны размещаются в нескольких зонах, имеющих различные параметры (т. е. абсолютные значения и анизотропию эффективных масс).

В последние годы было проведено весьма большое число работ, и теоретических и экспериментальных, посвященных исследованию энергетического спектра электронов в полупроводниках. Эти исследования показали, что при определенных условиях, даже вблизи нижнего края зоны, масса не остается постоянной, а растет с ростом энергии электрона. Эффективная масса зависит от температуры, от давления, может зависеть от характера и числа примесей и от концентрации электронов.

Впоследствии мы покажем (см. гл. 3), что скорость электронов обращается в нуль у верхнего края зоны. Электроны, находящиеся вблизи верхнего края зоны, двигаясь в направлении поля, увеличивают свою энергию, т. е. поднимаются на энергетической диаграмме вверх и таким образом приближаются к «потолку» зоны. Но это значит в свете сказанного выше, что скорость их при этом

*) Точнее, «квазиимпульса», см. гл. 3.

уменьшается; двигаясь в направлении приложенной силы, они не ускоряются, а замедляются, т. е. ведут себя как частицы с отрицательной массой. Пустые места (дырки), напротив, ведут себя здесь как нормальные положительно заряженные частицы с положительной массой. Для дырок вблизи верхнего края зоны точно так же, как для электронов, вблизи нижнего края зоны ускорение пропорционально приложенной силе

$$a = \frac{F}{m_p}, \quad (1.22)$$

где m_p — коэффициент, называемый эффективной массой дырок.

Все сказанное выше об эффективной массе электронов относится и к эффективной массе дырок; поэтому мы не будем здесь повторяться.

Для электронов, находящихся в середине зоны, обычно не существует простой пропорциональности между силой и ускорением. Мы не будем на этом случае останавливаться подробнее, так как он наиболее сложен и сравнительно редко встречается в полупроводниках.

Уже упоминалось, что общее число электронных состояний в любой зоне $2Ng$ равно сумме числа состояний на атомарных уровнях, из которых она образовалась. Однако ширина валентных и свободных зон в разных кристаллах колеблется в очень широких пределах — от долей электрон-вольта до нескольких электрон-вольт.

Чем шире зона, тем меньше число электронных состояний dG , приходящихся на определенный интервал энергии $d\varepsilon$. Величину $g(\varepsilon) = dG/d\varepsilon$ называют плотностью электронных состояний. Теория показывает, что вблизи края зоны $g(\varepsilon)$ определяется эффективной массой m и выражается через нее следующим образом:

$$g(\varepsilon) = 4\pi (2m)^{\frac{3}{2}} \varepsilon^{\frac{1}{2}}, \quad (1.23)$$

где m — соответственно эффективная масса электрона или дырки.

ЭЛЕКТРОНЫ В ТЕПЛОМ РАВНОВЕСИИ

При своем движении в зоне проводимости полупроводника электроны все время испытывают соударения с ато-

мами, совершающими тепловые колебания, и обмениваются с ними кинетической энергией. Именно поэтому кинетическая энергия этих электронов в среднем, так же как кинетическая энергия атомов, равна $\frac{3}{2}kT$ (при комнатной температуре приблизительно равна 0,04 эв).

Таким образом, энергия теплового движения электронов так же, как энергия частиц обычного газа, растет пропорционально температуре, скорость — пропорционально квадратному корню из температуры. Совершенно по-иному ведут себя валентные электроны в металле. Вернемся опять к примеру щелочных металлов. Как уже упоминалось, в этом случае валентная зона кристалла оказывается наполовину заполненной. Ширина валентной зоны составляет несколько электрон-вольт, и, таким образом, все энергетические уровни в ней до уровня с энергией по крайней мере в 1 эв заполнены уже при абсолютном нуле. Поэтому тепловые колебания атомов не могут оказать значительного влияния на распределение электронов по энергиям.

Действительно, электроны, находящиеся не на самых верхних уровнях, а в нижней половине зоны, не могут изменить своей энергии под действием столкновений с атомами, так как при этом они должны перейти на другой уровень, а все соседние уровни заняты. Электроны же, находящиеся на верхних уровнях, имеют энергию порядка 1 эв, в то время как средняя энергия тепловых колебаний атомов составляет при комнатной температуре всего лишь 0,04 эв, при температуре 1000° К (700° К) — 0,1 эв, т. е. во много раз меньше. Поэтому при столкновении с атомом электрон не может увеличить свою энергию; с другой стороны, он не может также и передать свою энергию атому, так как при этом он должен был бы перейти в энергетически более низкое состояние, а все нижние энергетические уровни заняты.

Впрочем, все сказанное относится лишь к столкновениям электронов с атомами со средней энергией. Наряду с этим всегда имеется некоторое число атомов, энергия которых значительно больше средней; под действием столкновений с такими атомами электроны могут переходить в верхнюю половину зоны. Так как число таких атомов очень мало, вероятность этих переходов тоже очень мала, но все же благодаря им при температуре, отличной от абсолютного нуля, всегда имеется некоторое число

электронов в верхней половине валентной зоны и некоторое число пустых мест в нижней.

Итак, в силу того что в щелочных металлах половина валентной зоны заполнена электронами, средняя энергия электронов уже при абсолютном нуле значительно выше, чем энергия тепловых колебаний атомов при любых температурах вплоть до точки плавления металла; поэтому энергия электронов не может значительно изменяться под действием соударений с атомами. Это относится и к щелочно-земельным металлам, так как в них s -зона заполнена целиком и они проводят электрический ток лишь благодаря тому, что верхний край s -зоны перекрывается нижним краем следующей, пустой p -зоны.

Такое состояние свободных электронов, когда их энергия почти не зависит от температуры, называется *вырожденным*. В полупроводниках энергия электронов растет пропорционально температуре; такое состояние электронного газа называется *невырожденным*.

Из сказанного выше ясно, что степень вырождения зависит от числа электронов в зоне проводимости. В полупроводниках это число обычно мало, и все электроны могли бы разместиться на самых низких уровнях в виде тонкого и плотного слоя «электронной жидкости». Но под действием теплового движения электроны «испаряются», т. е. забрасываются на более высокие уровни, расположенные в энергетическом интервале порядка kT , и так как число уровней в этом интервале во много раз больше, чем число электронов, то «электронная жидкость» превращается в разреженный газ. В металлах же зона заполнена до половины плотной «электронной жидкостью», над которой при температуре, отличной от нуля, имеется тонкий слой «электронного пара». Для того чтобы вся эта «жидкость» превратилась в пар (или газ), нужна очень высокая температура (в зависимости от ширины зоны — от 5 000 до 10 000°), и фактически сам металл испарится раньше, чем это должно было бы произойти.

В 1 см^3 металла содержится приблизительно 10^{22} свободных электронов; в полупроводниках концентрация свободных электронов колеблется обычно в зависимости от чистоты, ширины запрещенной зоны и температуры в пределах от 10^{12} до 10^{17} см^{-3} . Наконец, в полуметаллах и в очень «грязных» полупроводниках мы имеем дело с промежуточным случаем: в них число свободных электронов

в 1 см^3 меняется в зависимости от содержания примесей от 10^{18} до 10^{21} . Предположим для определенности, что оно порядка 10^{19} . Такое число электронов при плотной упаковке и эффективной массе $m = m_0$ заняло бы все уровни в зоне приблизительно вплоть до энергий $0,03 \text{ эв}$. Как мы уже упоминали, при комнатной температуре энергия тепловых колебаний атомов составляет примерно столько же. Поэтому в данном случае при температурах значительно ниже комнатной электронный газ будет находиться в вырожденном состоянии, при температурах значительно выше комнатной — в невырожденном. В интервале температур приблизительно от 100 до 400° К электроны будут находиться в состоянии частичного вырождения, т. е. постепенно переходить из вырожденного состояния в невырожденное. В соответствии с этим энергия электронов при низких температурах будет очень слабо зависеть от температуры, в промежуточной области зависимость постепенно усилится и, наконец, при дальнейшем нагревании средняя энергия электронов начнет расти пропорционально температуре.

Все сказанное выше о полуметаллах с концентрацией электронов 10^{19} см^{-3} будет относиться ко всем полуметаллам (в том числе и дырочным) с той лишь разницей, что в зависимости от концентрации электронов или дырок переходная область будет сдвигаться соответственно в сторону либо более высоких, либо более низких температур.

Примесный полупроводник с большой концентрацией примесей и малой энергией активации также может при низких температурах находиться в вырожденном состоянии.

Качественно это можно объяснить следующим образом (на примере электронного примесного полупроводника). При температуре, близкой к абсолютному нулю, число электронов в зоне проводимости ничтожно и поэтому они находятся в невырожденном состоянии.

При увеличении температуры начинают конкурировать два процесса: заброс электронов в зону проводимости и «расселение» их по зоне в силу растущего теплового движения этих электронов. Если первый процесс будет преобладать, то в некотором интервале температур электроны в зоне могут «забыть» все нижние уровни и наступит вырождение. При дальнейшем увеличении температуры вырождение всегда снимается.

Более того, вырожденное состояние может иметь место и в собственном полупроводнике, если плотность состояний в валентной зоне во много раз больше, чем в свободной, и ширина запрещенной зоны невелика.

При этих условиях относительно небольшое число электронов, перешедших из валентной зоны в свободную, может заполнить последнюю до высокого уровня и, таким образом, создать вырождение. Это действительно имеет место в антимониде индия.

ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ ФЕРМИ

Выше мы упоминали, что в щелочных металлах при абсолютном нуле все состояния, расположенные ниже середины зоны, заняты, а все состояния, расположенные выше середины зоны, свободны. При температуре, отличной от нуля, под влиянием столкновения с колеблющимися атомами часть электронов испаряется в верхнюю половину зоны, а в нижней половине появляются пустые места — дырки.

Нечто подобное происходит и в собственном полупроводнике: при абсолютном нуле вся валентная зона заполнена, а зона проводимости пуста; при температуре, отличной от нуля, часть электронов «испаряется» из валентной зоны в зону проводимости; отличие от металла заключается в том, что валентная зона и зона проводимости разделены запрещенным промежутком, а верхняя и нижняя половины зоны металла непосредственно примыкают друг к другу.

В примесном электронном полупроводнике в зону проводимости переходят электроны с донорных уровней, а на них появляются неподвижные дырки. В дырочном полупроводнике электроны забрасываются на акцепторные уровни из валентной зоны и «застревают» там, а в зоне появляются подвижные дырки.

Для анализа подавляющего большинства явлений недостаточно только этих качественных представлений и даже знания того, сколько электронов при данной температуре находится в зоне проводимости полупроводника (или в верхней половине валентной зоны металла), а необходимо знать, как эти электроны распределены по энергиям. Для этого вводится так называемая функция распределения $f_0(\epsilon)$ (названная функцией распределения Ферми),

характеризующая вероятность того, что состояние с данной энергией ϵ занято:

$$f_0(\epsilon) = \frac{n(\epsilon)}{g(\epsilon)}. \quad (1.24)$$

Как показывает теория, $f_0(\epsilon)$ является универсальной функцией; плотность состояний также может быть более или менее точно вычислена для полупроводников и металлов; таким образом, распределение электронов по энергиям $n(\epsilon)$ может быть получено из (1.23) и (1.24):

$$n(\epsilon) = g(\epsilon) f_0(\epsilon). \quad (1.25)$$

Вычисление плотности состояний и вывод функции распределения отложим до гл. 4, здесь же приведем выражение для $f_0(\epsilon)$ без вывода:

$$f_0(\epsilon) = \frac{1}{e^{\frac{\epsilon - \mu}{kT}} + 1}, \quad (1.26)$$

где μ — величина, зависящая от энергетического спектра электронов в данном материале, числа электронов в нем и температуры.

Вычисление μ мы также отложим до гл. 4; пока же ограничимся качественным рассмотрением вопроса.

Как видно из (1.26), μ имеет размерность энергии, поэтому на энергетической диаграмме μ характеризует определенный уровень, который называют уровнем химического потенциала или уровнем Ферми.

Исследуем графически, как зависит функция распределения от энергии и температуры. Для этого будем откладывать, как это мы всегда делаем, по оси ординат энергию ϵ , а по оси абсцисс — $f_0(\epsilon)$ (рис. 1.10). Нуль отсчета энергии, как всегда, произволен; уровень μ на рис. 1.10 обозначен горизонтальной штрих-пунктирной линией.

Рассмотрим сначала вид $f_0(\epsilon)$ при абсолютном нуле. При $T \rightarrow 0$ и при $\epsilon > \mu$ значение $(\epsilon - \mu)/kT \rightarrow \infty$ и величина $f_0(\epsilon) \rightarrow 0$; при $\epsilon < \mu$ $(\epsilon - \mu)/kT \rightarrow -\infty$ и $f_0(\epsilon) \rightarrow 1$; при $\epsilon = \mu$ $f_0(\epsilon)$ скачком переходит от значения $f_0(\epsilon) = 1$ к значению $f_0(\epsilon) = 0$ и, следовательно, изображается ломаной кривой, представленной на рис. 1.10, а. Таким образом, все состояния ниже уровня Ферми заполнены, а выше него пусты. В соответствии с изложенным это означает, что при абсолютном нуле уровень Ферми

проходит в щелочном металле как раз посредине валентной зоны.

При температуре, отличной от абсолютного нуля, согласно (1.26) будет по-прежнему $f_0(\epsilon) = 0$ при $(\epsilon - \mu)/kT \gg 0$ и $f_0(\epsilon) = 1$ при $(\epsilon - \mu)/kT \ll 0$, но для значений энергий, близких к μ , положение существенно изменяется (рис. 1.10, б). Теперь при $\epsilon = \mu$ значение $f_0(\epsilon) = 1/2$, при $\epsilon < \mu$ $1/2 < f_0(\epsilon) < 1$ и при $\epsilon > \mu$

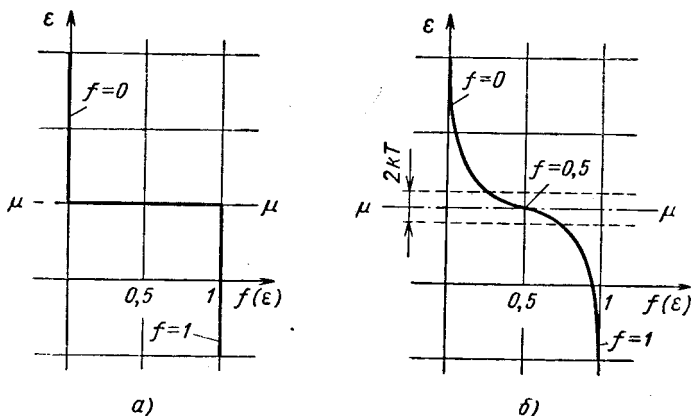


Рис. 1.10. Функция распределения Ферми при двух различных температурах: а) $T = 0^\circ \text{K}$; б) $T > 0^\circ \text{K}$.

имеет место неравенство $0 < f_0(\epsilon) < 1/2$; функция распределения теперь уже плавно переходит от значения $f_0(\epsilon) = 1$ к значению $f_0(\epsilon) = 0$, причем интервал энергий, в котором происходит этот постепенный переход, порядка $4kT$ (рис. 1.10, б). Чем выше температура, тем больше «расплывается» функция распределения Ферми. Качественно это означает, что часть электронов забрасывается из нижней половины зоны в верхнюю. Так как все это происходит в интервале энергий порядка $4kT$, что при комнатной температуре составляет всегда лишь 0,1 эв, то в среднем энергия всех свободных электронов в металле почти не изменится; как мы упоминали, такое состояние электронов, в котором их энергия почти не зависит от температуры, называется вырожденным.

Рассмотрим теперь, как обстоит дело в полупроводниках. Общее число состояний в свободной зоне полупровод-

ника равно $2Ng$, где N — число атомов и g — кратность уровня, из которого образовалась зона. Число атомов в 1 см^3 твердого тела $N \approx 10^{22}$, следовательно, такого же порядка будет и число состояний. Число же свободных электронов в полупроводниках колеблется обычно в пределах 10^{12} — 10^{18} см^{-3} . Это значит, что доля занятых состояний ничтожно мала или, иными словами, $f_0(\epsilon) \ll 1$ для всех состояний, включая самые нижние, примыкающие ко дну зоны *).

Но если $f_0 \ll 1$, то это значит, что $e^{(\epsilon - \mu)/kT} \gg 1$ или $(\epsilon - \mu)/kT \gg 1$, т. е. уровень химического потенциала должен проходить ниже дна зоны. В этом случае для всех состояний в зоне проводимости мы можем пренебречь единицей в выражении для $f_0(\epsilon)$ (1.26) и, следовательно,

$$f_0(\epsilon) = e^{\frac{\mu - \epsilon}{kT}}. \quad (1.27)$$

Приближенное выражение (1.27) называется распределением Максвелла — Больцмана. Состояние электронного газа, в котором $f_0(\epsilon) \ll 1$ и которое описывается поэтому приближенной формулой (1.27), называется невырожденным. Мы увидим ниже, что в этом случае энергия в отличие от вырожденного состояния растет пропорционально температуре. Анализ положения и температурной зависимости уровня Ферми (гл. 4) позволяет сделать следующие выводы.

1. В области собственной проводимости уровень химического потенциала обычно (если эффективные массы электронов и дырок примерно одинаковы) проходит середине зоны.

2. В примесном (некомпенсированном) электронном полупроводнике при температуре абсолютного нуля уровень химического потенциала проходит середине между дном зоны проводимости и примесными уровнями, а затем начинает подниматься. При этом конкурируют два процесса: заброс электронов в зону проводимости и расселение (или «испарение» со дна зоны) заброшенных электронов. **)

*) Это условие не выполняется в вырожденных полупроводниках и полуметаллах.

**) Если энергия активации примесных уровней мала, а число их велико, то уровень химического потенциала при этом может пересечь дно зоны, и неравенство $f_0(\epsilon) \ll 1$ нарушится. В некотором интервале температур полупроводник будет частично вырожден.

До тех пор пока первый процесс преобладает, плотность электронного газа в зоне проводимости растет и уровень Ферми поднимается. Когда начинает преобладать второй процесс, уровень Ферми начинает опускаться; когда половина примесных электронов перешла в зону, уровень Ферми совпадает с примесными уровнями, затем пересекает их и приближается к середине зоны — начинается область собственной проводимости.

3. Аналогично все происходит в дырочном (нескомпенсированном) полупроводнике. При абсолютном нуле уровень Ферми проходит посередине между верхним краем валентной зоны и акцепторными уровнями, затем сначала опускается и приближается ко дну валентной зоны, потом начинает подниматься, пересекает акцепторные уровни и приближается к середине запрещенной зоны.

В некоторых случаях мы еще более упростим вопрос и будем считать, что уровень химического потенциала во всей области примесной проводимости проходит посередине между примесными уровнями и краем соответствующей зоны.

1.3. ЭЛЕКТРОПРОВОДНОСТЬ ПОЛУПРОВОДНИКОВ

ФОРМУЛЫ ДЛЯ ПОДВИЖНОСТИ И ЭЛЕКТРОПРОВОДНОСТИ

В отсутствие поля скорости свободных электронов имеют всевозможные направления. Под действием электрического поля E электроны получают ускорение $a = -eE/m_n$ и направленную по полю добавку к скорости $\Delta v = a\tau_n$, где τ_n — время, в течение которого действует ускорение a . Если бы τ_n не было ничем ограничено, то и скорость электрона в направлении поля неограниченно возрастала бы. В действительности, однако, это не так. Теория показывает, что в строго периодическом поле, т. е. в идеальной кристаллической решетке с покоящимися атомами, электроны действительно двигались бы ускоренно на протяжении очень больших промежутков времени. Но идеальных кристаллов не существует. В реальных же кристаллах всегда существуют нарушения периодического потенциала, обусловленные, во-первых, тепловыми колебаниями атомов и, во-вторых, всевозможными дефектами решетки:

— так называемыми точечными дефектами: пустыми узлами и атомами в междуузлиях, нейтральными и заряженными примесными атомами (которые также могут размещаться в узлах и междуузлиях);

— линейными дефектами — краевыми и винтовыми дислокациями;

— плоскими дефектами — границами зерен, блоков, двойников;

— объемными дефектами — трещинами, полостями, включениями (в том числе зародышами) другой фазы.

В силу этого электрон лишь на сравнительно небольшом отрезке пути l_n , называемом длиной свободного пробега, движется ускоренно, затем испытывает соударение, теряет при этом свою направленную скорость, и весь процесс начинается сначала. При высоких температурах длину свободного пробега ограничивают тепловые колебания атомов, при низких — примеси и другие дефекты кристаллов.

Время τ_n , входящее в выражение для Δv , которое называется временем свободного пробега, мы можем, зная длину свободного пробега l_n , вычислить по формуле

$$\tau_n = \frac{l_n}{v}. \quad (1.28)$$

Опыт показывает, что длина свободного пробега электрона обычно очень мала, $l_n < 10^{-5}$ см, и в силу этого направленная добавка к скорости электрона Δv при не очень сильных полях мала по сравнению со скоростью хаотического теплового движения электронов v_0 в отсутствие поля *), поэтому в (1.28) вместо v может быть поставлено v_0 :

$$\tau_n \approx \frac{l_n}{v_0}. \quad (1.29)$$

Средняя скорость, с которой электрон будет двигаться вдоль поля,

$$\bar{v} = \frac{\overline{\Delta v}}{2} = \frac{e}{2m_n} \frac{l_n}{v_0} E = u_n E, \quad (1.30)$$

где отношение скорости дрейфа электронов к электрическому полю обозначено буквой u :

$$u_n = \frac{e}{2m_n} \frac{l_n}{v_0}. \quad (1.31)$$

*) Условия, при которых добавка к скорости становится сравнимой с тепловой скоростью, рассмотрены в § 5 гл. 5.

Это отношение называется подвижностью электрона. Иными словами, подвижность — это скорость дрейфа электронов в поле напряженностью 1 в/см . Более строгий расчет дает в два раза большее значение подвижности:

$$u_n = \frac{e}{m_n} \frac{l_n}{v_0} = \frac{e}{m_n} \tau_n. \quad (1.32)$$

Если в 1 см^3 содержится n свободных электронов и все они движутся в направлении поля со скоростью \bar{v} , то через площадку в 1 см^2 за секунду пройдут все электроны, находящиеся в объеме параллелепипеда длиной \bar{v} ; число таких электронов будет $n\bar{v}$; заряд, перенесенный этими электронами, называется плотностью тока

$$j = en\bar{v} = \frac{ne^2\bar{l}_n}{m_n v_0} E.$$

Следовательно, электропроводность материала

$$\sigma_n = \frac{j}{E} = \frac{ne^2\bar{l}_n}{m_n v_0} = enu_n. \quad (1.33)$$

Совершенно аналогично вычисляется подвижность и электропроводность дырочного полупроводника; поэтому мы приведем здесь соответствующие выражения, не повторяя проделанных элементарных вычислений:

$$u_p = \frac{e}{m_p} \frac{\bar{l}_p}{v_0} \quad (1.34)$$

и

$$\sigma_p = \frac{e^2 p \bar{l}_p}{m_p v_0}. \quad (1.35) *$$

В области собственной и смешанной (т. е. переходной от примесной к собственной) проводимости электронная и дырочная проводимости складываются:

$$\sigma = e(nu_n + pu_p); \quad (1.36)$$

при наличии нескольких сортов носителей (дырок и электронов) с различными массами и подвижностями выражение (1.36) легко обобщается:

$$\sigma = \Sigma \sigma_i = \Sigma en_i u_i, \quad (1.37)$$

*) В дальнейшем, в тех случаях, когда это не вызывает недоразумений, мы индексы n и p будем опускать.

где суммирование проводится по всем сортам носителей (и дыркам и электронам).

**ЗАВИСИМОСТЬ ДЛИНЫ СВОБОДНОГО ПРОБЕГА
ЭЛЕКТРОНА ОТ ТЕМПЕРАТУРЫ. ТЕМПЕРАТУРНАЯ
ЗАВИСИМОСТЬ ПОДВИЖНОСТИ**

Выше уже упоминалось, что длина свободного пробега, а следовательно, и подвижность электрона в реальных кристаллах ограничивается нарушениями периодического потенциала, обусловленными тепловыми колебаниями и дефектами решетки; дефекты играют основную роль при низких температурах, а тепловые колебания — при высоких. Мы должны рассмотреть теперь каждый механизм рассеяния электронов подробнее.

Рассеяние на дефектах решетки. На стр. 37 мы перечислили ряд причин, нарушающих периодичность потенциала в реальных кристаллах.

Для того чтобы электрон изменил направление своего движения в результате взаимодействия с нейтральным дефектом, необходим акт столкновения в буквальном смысле этого слова — траектория электрона должна проходить через место расположения дефекта либо через непосредственно примыкающую к нему область решетки, в которой им вызваны искажения.

Электрическое поле ионизированного примесного атома распространяется значительно дальше, и электрон, проходя на значительном расстоянии от него, изменит под действием этого поля направление своего движения. Поэтому из всех перечисленных выше дефектов основную роль в рассеянии электронов играют ионизированные примеси и только их мы сейчас будем рассматривать.

Чем больше тепловая скорость движения электрона, тем меньше времени он находится в непосредственной близости от примесного иона и тем меньшее отклонение он испытывает; поэтому длина свободного пробега электрона очень резко растет с увеличением его скорости. Теория показывает, что в этом случае длина свободного пробега пропорциональна четвертой степени скорости или, что то же самое, квадрату кинетической энергии:

$$\bar{l} \sim v_0^4 \sim \bar{\epsilon}^2. \quad (1.38)$$

Из (1.38) следует два важных вывода:

1. При рассеянии на ионах примеси доля участия электронов с различными скоростями в электрическом токе различна — основную роль играют быстрые электроны, так как их длина свободного пробега больше. Отсюда также следует, что средняя энергия в потоке (электрическом токе) электронов будет отлична (в большую сторону) от средней тепловой энергии.

2. Согласно (1.34) подвижность электрона пропорциональна отношению длины свободного пробега к средней скорости: $u \sim \bar{l}/v_0$, и согласно формуле (1.38) $u \sim v_0^3$. Но для невырожденного полупроводника средняя кинетическая энергия пропорциональна температуре, а средняя тепловая скорость пропорциональна квадратному корню из температуры $v_0 \sim T^{1/2}$; следовательно, в рассматриваемом нами случае рассеяния электронов на ионах примеси

$$u \sim T^{\frac{3}{2}}. \quad (1.38a)$$

Таким образом, здесь осуществляется очень редкий случай, когда подвижность электронов растет с температурой.

Рассеяние электронов на тепловых колебаниях решетки.

С повышением температуры увеличиваются тепловые скорости электронов и уменьшается вероятность их рассеяния на примесных ионах. Одновременно при этом растут амплитуды тепловых колебаний атомов и растет роль этого механизма рассеяния. Поэтому при высоких температурах рассеяние на тепловых колебаниях играет решающую роль, а рассеянием на ионах примеси можно пренебречь. В этом случае зависимость длины свободного пробега электрона от температуры обусловлена двумя причинами: во-первых, как уже упоминали, с ростом температуры увеличиваются амплитуды тепловых колебаний атомов, что непосредственно ведет к уменьшению длины свободного пробега; во-вторых, длина свободного пробега электрона зависит от его кинетической энергии ϵ (или, что то же самое, от v_0), которая, в свою очередь, зависит от температуры.

Теория показывает, что в этом случае выражение для длины свободного пробега имеет вид

$$l = \epsilon^r l_0(T), \quad (1.39)$$

причем показатель степени r и температурная зависимость $l_0(T)$ различны в кристаллах различного типа и в различных интервалах температур.

В *ионных полупроводниках* при рассеянии электронов на оптических колебаниях в области высоких температур *) $r = 1$ и

$$l_0(T) = \frac{a}{T}, \quad (1.40)$$

где a — некоторая константа; таким образом,

$$\bar{l} \sim \frac{\bar{\epsilon}}{T} \sim \frac{v_0^2}{T} \quad (1.41)$$

и, следовательно, подвижность

$$u \sim \tau \sim \frac{\bar{l}}{v_0} \sim \bar{\epsilon}^{-1/2} T^{-1}. \quad (1.42)$$

В невырожденных полупроводниках $\bar{\epsilon} \sim T$, и, следовательно,

$$u \sim T^{-\frac{1}{2}}. \quad (1.43)$$

В тех же кристаллах, но при температурах ниже температуры Дебая $r = \frac{1}{2}$, $l_0(T) \sim e^{\theta/T}$, где θ — температура Дебая; следовательно,

$$\bar{l} \sim \frac{1}{\bar{\epsilon}^{\frac{1}{2}}} e^{\theta/T}, \quad (1.44)$$

время релаксации $\bar{\tau} = \bar{l}/v_0$ не зависит от энергии и

$$u \sim e^{\theta/T}. \quad (1.45)$$

В ковалентных полупроводниках при рассеянии на акустических колебаниях $r = 0$

$$\bar{l} \sim T^{-1} \quad (1.46)$$

и не зависит от энергии; таким образом,

$$u \sim \bar{\epsilon}^{-\frac{1}{2}} T^{-1} \sim T^{-3/2}. \quad (1.47)$$

*) Т. е. температур, выше температуры Дебая, см. гл. 2, § 4 и 5.

Теория показывает также, что подвижность и прямо [см. формулу (1.32)] и косвенно (через v_0 и l_0) зависит от эффективной массы по двум причинам: так как тепловая средняя скорость зависит от эффективной массы $1/2 \overline{mv_0^2} = = 3/2 kT$ и так как вероятность столкновений, а следовательно, и длина свободного пробега l зависят от плотности состояний $g(\epsilon)$, а плотность состояний зависит от эффективной массы.

Экспериментальные и теоретические исследования, проведенные за последние годы, показали, что в целом ряде полупроводников (в первую очередь в полупроводниках с узкой запрещенной зоной) эффективная масса зависит от температуры.

Причины этого заключаются в следующем. Эффективная масса в полупроводниках с узкой запрещенной зоной растет по мере роста расстояния от дна зоны. Так как с ростом температуры средняя энергия электронов растет и они удаляются от дна зоны, то соответственно растет и средняя эффективная масса, входящая в выражение (1.32). Эффективная масса зависит также от межатомных расстояний: с ростом межатомных расстояний она, как правило, увеличивается; таким образом, она возрастает из-за теплового расширения кристалла. В некоторых случаях на эффективную массу оказывают также прямое влияние тепловые колебания.

В силу температурной зависимости эффективной массы температурная зависимость подвижности в области высоких температур (т. е. рассеяния на тепловых колебаниях решетки) оказывается часто более крутой, чем это предсказывают формулы, приведенные выше. В ряде полупроводников $\mu \sim T^{-2}$; $\mu \sim T^{-2.5}$ и даже $\mu \sim T^{-3}$.

Для дальнейшего важно, однако, что длина свободного пробега в большинстве случаев является степенной функцией от энергии электрона:

$$l \sim \epsilon^r, \quad (1.48)$$

где показатель степени r зависит от механизма рассеяния электронов; в атомной решетке при расстоянии на акустических колебаниях $r = 0$; $r = 1$ при $T > \theta$ и $r = 1/2$ при $T < \theta$ в ионной решетке; при рассеянии электронов на ионах примеси $r = 2$.

За последние годы обнаружен довольно широкий класс полупроводников (окислы переходных металлов и др.),

в которых при низких температурах перемещение электронов носит характер активационных скачков и подвижность экспоненциально растет с повышением температуры. Этот механизм электропроводности теоретически еще мало изучен, несмотря на то, что на этих материалах основан важный класс полупроводниковых приборов — терморезисторы.

1.4. ТЕПЛОПРОВОДНОСТЬ ПОЛУПРОВОДНИКОВ

В твердом теле существуют два основных механизма переноса тепла: упругими колебаниями решетки и свободными электронами (разумеется, если таковые имеются)*). В соответствии с этим теплопроводность κ можно разбить на две части — решеточную (κ_p) и электронную ($\kappa_{эл}$):

$$\kappa = \kappa_p + \kappa_{эл}. \quad (1.49)$$

В металлах число свободных электронов очень велико (порядка 10^{22} в 1 см^3) и велика электронная теплопроводность. В диэлектриках свободных электронов почти нет и $\kappa = \kappa_p$. В полупроводниках обычно решеточная теплопроводность много больше электронной, однако в тех случаях, когда число свободных электронов относительно велико (порядка 10^{18} — 10^{20} в 1 см^3), $\kappa_{эл}$ и κ_p могут быть одного и того же порядка.

Рассмотрим обе составляющие теплопроводности отдельно.

ЭЛЕКТРОННАЯ ТЕПЛОПРОВОДНОСТЬ

Механизм электронной теплопроводности ничем не отличается от механизма теплопроводности обычного газа, и можно для ее анализа воспользоваться формулой для теплопроводности газа, вывод которой приводится в любом учебнике физики:

$$\kappa_{эл} = \frac{1}{3} C_v \bar{v}_0 \bar{l}, \quad (1.50)$$

где C_v — теплоемкость 1 см^3 при постоянном объеме; \bar{v}_0 — средняя тепловая скорость и \bar{l} — средняя длина свободного пробега молекул газа.

*) Мы здесь не останавливаемся на переносе тепла излучением, который не имеет большого значения в большинстве полупроводников (см. гл. 6).

В рассматриваемом нами случае

$$C_v = \frac{3}{2} nk, \quad (1.51)$$

где k — постоянная Больцмана и n — число свободных электронов в единице объема. Следовательно,

$$\kappa_{эл} = \frac{1}{2} knv_0\bar{l}. \quad (1.52)$$

Как и следовало ожидать, теплопроводность электронов, так же как и электропроводность, пропорциональна их концентрации и длине свободного пробега. Но только эти две величины и могут различаться в различных материалах (средняя скорость v_0 для невырожденного электронного газа зависит только от температуры). Следовательно, отношение теплопроводности к электропроводности должно быть некоторой универсальной постоянной.

Действительно, согласно формуле (1.33)

$$\sigma = \frac{ne^2\bar{l}}{2mv_0} \quad (1.53)$$

и, следовательно,

$$\frac{\kappa_{эл}}{\sigma} = \frac{k}{e^2} mv_0^2. \quad (1.53a)$$

Полагая $\frac{1}{2}mv_0^2 = \frac{3}{2}kT$, получаем окончательно

$$\frac{\kappa_{эл}}{\sigma} = 3 \left(\frac{k}{e} \right)^2 T. \quad (1.54)$$

Более строгий вывод (с учетом распределения электронов по энергиям) дает формула для невырожденного электронного газа

$$\frac{\kappa_{эл}}{\sigma} = (r + 2) \left(\frac{k}{e} \right)^2 T = LT^*), \quad (1.55)$$

где r — показатель степени в зависимости длины свободного пробега от энергии, и для сильновырожденного электронного газа (т. е. в металлах)

$$\frac{\kappa_{эл}}{\sigma} = \frac{\pi^2}{3} \left(\frac{k}{e} \right)^2 T. \quad (1.56)$$

*) Коэффициент в отношении теплопроводности к электропроводности, который мы обозначим буквой L , называют числом Лоренца.

Если выразить теплопроводность $\kappa_{эл}$ в *вт/см·град* и электропроводность σ в *ом⁻¹·см⁻¹*, то согласно (1.56) для металлов

$$\frac{\kappa_{эл}}{\sigma} = 2,44 \cdot 10^{-8} T \quad (1.57)$$

и для полупроводников согласно (1.55)

$$\frac{\kappa_{эл}}{\sigma} = 0,7 \cdot 10^{-8} T (r + 2). \quad (1.58)$$

Соотношения (1.55) и (1.56) носят название закона Видемана и Франца по имени ученых, впервые установивших, что отношение теплопроводности к электропроводности металлов есть величина постоянная (зависящая только от температуры).

Соотношение (1.55) было также подтверждено на полупроводниках советскими учеными А. Ф. Иоффе и А. В. Иоффе, Е. Д. Девятковой и П. В. Гультяевым.

Однако они же и одновременно независимо от них иностранные ученые установили, что закон Видемана и Франца соблюдается лишь в области примесной проводимости. Уже при приближении к температуре, при которой начинает проявляться собственная проводимость, теплопроводность начинает расти значительно быстрее, чем электропроводность, и в области собственной проводимости достигает значений, на много больших, чем следует из (1.55). Причина этого явления заключается в следующем.

Формула (1.53), выведенная для обычного газа, учитывает лишь тепло, переносимое электронами за счет того, что их кинетическая энергия растет с повышением температуры. Но в области собственной проводимости число электронов и дырок на горячем конце образца больше, чем на холодном; в результате этого возникает диффузия электронно-дырочных пар в направлении, обратном градиенту температуры, и на горячем конце будет поглощаться энергия, необходимая для образования такой пары, а на холодном пары будут рекомбинировать и та же энергия будет выделяться. Эта энергия состоит из кинетической энергии электрона и дырки и энергии $\Delta \mathcal{E}_0$, необходимой для того, чтобы электрон был переброшен из заполненной зоны в свободную. Так как обычно $\Delta \mathcal{E}_0 \gg kT$, то энергия, переносимая такой парой, во много раз больше, чем переносимая электроном или дыркой в области примесной

проводимости. Таким образом, возникает дополнительный поток тепла и дополнительная теплопроводность $\Delta\kappa_{эл}$. Если число пар велико, то будет велика и дополнительная теплопроводность $\Delta\kappa_{эл}$. Но для этого необходимо, чтобы ширина запрещенной зоны была не слишком велика, так как в области собственной проводимости $n = p \sim e^{-\Delta\mathcal{E}_0/2kT}$. Поэтому $\Delta\kappa_{эл}$ велико в полупроводниках со сравнительно узкой запрещенной зоной (так, например, в теллуриде висмута, в котором это явление было впервые обнаружено, $\Delta\mathcal{E}_0 \approx 0,2$ эв). Б. И. Давыдовым и И. М. Шмушкевичем [3] была выведена формула для электронной теплопроводности в области собственной проводимости с учетом $\Delta\kappa_{эл}$ за много лет до того, как это явление было обнаружено на опыте:

$$\kappa_{pn} = L(\sigma_p + \sigma_n) + L \frac{2\sigma_p\sigma_n}{\sigma_p + \sigma_n} \left(\frac{\Delta\mathcal{E}_0}{2kT} + r + 2 \right)^2, \quad (1.59)$$

где r — показатель степени в зависимости длины свободного пробега электрона от энергии.

ТЕПЛОПРОВОДНОСТЬ КРИСТАЛЛИЧЕСКОЙ РЕШЕТКИ*)

Атомы твердого тела связаны друг с другом упругими силами. Поэтому тепловые колебания любого атома передаются соседнему, и таким образом по всему телу во всех возможных направлениях распространяются упругие волны. Эти волны отличаются друг от друга не только направлением распространения, но и своей длиной, **): наиболее короткие имеют длину λ , равную удвоенному расстоянию между соседними атомами; наиболее длинные имеют длину λ , равную удвоенной длине кристалла L . Скорости их распространения ***) ω_ϕ связаны с длиной волны и частотой колебаний ν соотношением

$$\omega_\phi = \lambda\nu, \quad (1.60)$$

*) Нарисованная здесь картина очень упрощена, см. подробнее гл. 2 и 6.

**) Кроме этого, волны отличаются друг от друга поляризацией; для заданной длины волны и направления существует три вида колебаний: два поперечных и одно продольное.

***) Речь идет о фазовой скорости.

т. е. частота каждого колебания обратно пропорциональна длине волны:

$$v = \frac{\omega_{\phi}}{\lambda}. \quad (1.61)$$

В теории колебаний обычно вводится понятие волнового числа $q = 1/\lambda$, равного числу волн, укладываемых на единице длины, и волнового вектора q , равного по абсолютной величине волновому числу и совпадающего с направлением распространения волны.

Остановимся коротко на факторах, определяющих теплопроводность решетки. Если бы кристалл был совершенно идеальным (т. е. не имел бы никаких дефектов), а атомы были бы связаны друг с другом силами F , строго подчиняющимися закону Гука:

$$F = -f \frac{x - x_0}{x_0} = -f \frac{\Delta x}{x_0}, \quad (1.62)$$

где f — коэффициент упругой связи; x_0 — нормальное расстояние между покоящимися атомами (т. е. постоянная решетки $x_0 = a$) и Δx — изменение этого расстояния, то колебания были бы строго периодическими (гармоническими); при этих условиях упругие волны совершенно не взаимодействовали бы при встрече друг с другом и проходили бы «одна сквозь другую» так же свободно, как лучи света в пустоте.

Если в таком идеальном кристалле можно было бы создать перепад температур, то атомы, находящиеся на горячем конце и колеблющиеся с наибольшими амплитудами, передавали бы свою энергию соседним и фронт тепловой волны распространялся бы со скоростью звука вдоль всего кристалла. Так как при этом тепловая волна не встречала бы на своем пути никакого сопротивления, то для создания бесконечно большого потока тепла достаточно было бы бесконечно малого градиента температуры; теплопроводность такого кристалла была бы бесконечной*). Но таких идеальных кристаллов не существует. В реальных же кристаллах сила взаимодействия соседних атомов лишь в первом приближении (при очень небольших смещениях Δx) может быть представлена уравнением (1.62).

*) Точнее, при отсутствии взаимодействия между колебаниями нельзя бы было ввести понятие температуры, а следовательно — и теплопроводности.

В действительности же F выражается через Δx бесконечным рядом

$$F = -f \frac{\Delta x}{x_0} + g \left(\frac{\Delta x}{x_0} \right)^2 + \dots \quad (1.63)$$

и чем больше изменение расстояния (амплитуда колебаний) Δx , тем большее значение приобретают последующие члены ряда.

При упругих деформациях кристалла под влиянием внешнего воздействия смещения Δx бывают обычно очень малы; следовательно, вторым членом в (1.63) можно пренебречь, и поэтому в данном случае соблюдается закон Гука (1.62).

При рассмотрении тепловых колебаний необходимо (и обычно достаточно) учитывать второй член в разложении (1.63); при этом значение его тем больше, чем больше коэффициент g , называемый коэффициентом ангармоничности, и чем больше амплитуда колебаний (т. е. чем выше температура). Наличие этого члена приводит к двум важным следствиям:

— среднее расстояние между атомами в реальном кристалле растет с повышением температуры, причем коэффициент теплового расширения пропорционален коэффициенту g ;

— колебания атомов перестают быть строго гармоническими и в результате этого не могут распространяться независимо, а напротив, при встрече друг с другом рассеиваются, т. е. меняют направление своего движения и обмениваются энергией.

В реальных твердых телах, кроме этого, всегда имеются дефекты (примеси, пустые узлы и атомы в междуузлиях, дислокации, границы зерен в поликристаллических материалах), на которых также рассеиваются упругие волны.

Поэтому теплопроводность реальных кристаллов имеет конечную величину, зависящую от того, в какой мере перечисленные выше факторы (ангармоничность колебаний и дефекты) затрудняют распространение тепловых волн.

Для любых волн, точно так же, как и для частиц, можно ввести понятие длины свободного пробега l ; в данном случае это расстояние, на котором волна один раз рассеивается, т. е. изменяет направление своего движения *).

*) Точнее, это то расстояние, на котором амплитуда волны в результате рассеяния уменьшится в e раз.

При низких температурах длина свободного пробега упругих волн ограничивается рассеянием на дефектах, при высоких — рассеянием друг на друга (вследствие их ангармоничности).

Дебай в 1915 г. вывел формулу для теплопроводности кристаллической решетки, совершенно аналогичную по внешнему виду формуле (1.52) для теплопроводности газа:

$$\kappa_p = \frac{1}{3} C \omega_{\phi} \bar{l}_{\phi}, \quad (1.64)$$

где C — теплоемкость 1 см^3 решетки ($C = c_v \rho$, c_v — удельная теплоемкость и ρ — плотность);

ω_{ϕ} — скорость звука *);

\bar{l}_{ϕ} — средняя длина свободного пробега упругих волн — фононов) см. ниже).

Рассмотрим каждый сомножитель, входящий в правую часть (1.64), отдельно.

Скорость звука ω_{ϕ} выражается через плотность ρ , и модуль сжимаемости — через g_0 :

$$\omega_{\phi} = \sqrt{\frac{g_0}{\rho}}. \quad (1.65)$$

Эта величина сравнительно слабо отличается в различных твердых телах и обычно приблизительно равна 10^5 см/сек .

Рассмотрим первый множитель в (1.64) — теплоемкость единицы объема. Для этого представим себе, что рассматриваемый нами кристалл находится в тепловом равновесии с окружающим его газом (полученные таким образом результаты будут справедливы и для твердого тела в вакууме). Молекулы газа, бомбардируя поверхность кристалла, будут обмениваться с его атомами энергией, поэтому в состоянии равновесия средние кинетические энергии атомов твердого тела и газа будут одинаковы и равны $\frac{3}{2}kT$. Но энергия колеблющихся атомов состоит из кинетической и потенциальной; в момент максимального отклонения вся энергия переходит в потенциальную, в момент прохождения через положение равновесия вся энергия переходит в кинетическую, а в среднем (во времени) они должны быть равны. Следовательно, среднее значение

* Точнее, в (1.64) вместо ω_{ϕ} должно быть поставлено $\omega_{\text{гр}}$ — групповая скорость тепловых колебаний (см. гл. 2 и 6).

потенциальной энергии колеблющегося атома должно быть тоже $\frac{3}{2}kT$, полная энергия $\mathcal{E} = 3kT$; полная энергия \mathcal{E} грамм-атома любого вещества, таким образом, должна быть одна и та же и равна $\mathcal{E}_N = \mathcal{E}N = 3RT$ (где N — число Авогадро и R — универсальная газовая постоянная). Тогда теплоемкость грамм-атома

$$C = \frac{d\mathcal{E}}{dT} = 3R \approx 6 \text{ кал/моль.} \quad (1.66)^*$$

Соотношение (1.66), носящее название закона Дюлонга и Пти, в память ученых, установивших его эксперимен-

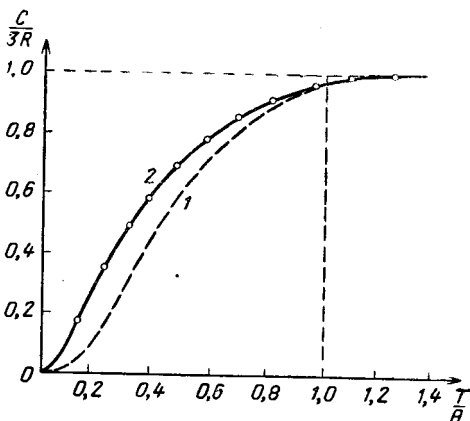


Рис. 1.11. Температурная зависимость теплоемкости твердого тела:

1 — функция Эйнштейна; 2 — функция Дебая. Кружками отмечены экспериментальные точки.

тально, нашло блестящее подтверждение в измерениях теплоемкости большого количества материалов при комнатной и более высоких температурах. Однако в дальнейшем исследования в области низких температур показали, что теплоемкость всех тел здесь падает по кривой, аналогичной кривой 2, представленной на рис. 1.11, причем

*) Для твердых тел теплоемкости при постоянном давлении (c_p) и постоянном объеме (c_v) очень мало отличаются друг от друга; это связано с тем, что при изменении давления объем твердого тела меняется очень слабо (по сравнению с газом), и поэтому работа при этом совершается очень малая; поэтому в дальнейшем мы не будем для твердого тела делать различия между c_v и c_p и индекс при c будем опускать.

температура θ , при которой начинается это падение (названная в дальнейшем температурой Дебая), для различных тел различна *).

Для объяснения аномального поведения теплоемкости Эйнштейн в 1911 г. предположил, что энергия колеблющегося атома (осциллятора) не может меняться непрерывно,

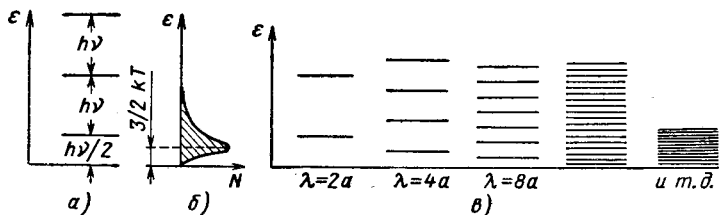


Рис. 1.12. Энергетический спектр квантованного осциллятора (а), распределение атомов газа по энергиям (б) и энергетический спектр цепочки атомов (в).

а изменяется лишь определенными порциями — квантами, причем величина этой порции связана с частотой колебаний соотношением

$$\epsilon_0 = h\nu, \quad (1.67)$$

где h — некоторая универсальная постоянная, равная $h = 6,57 \cdot 10^{-27}$ эрг·сек **).

На рис. 1.12, а изображен энергетический спектр такого квантованного осциллятора. В соответствии со сказанным выше он состоит из серии линий, отстоящих друг от друга на равное расстояние $\epsilon_0 = h\nu$. Самый нижний уровень соответствует состоянию с энергией $h\nu/2$ ***). Если частота колебаний равна, например, 10^{13} сек, то $\epsilon_0 \approx 10^{-13}$ эрг.

Представим снова, что твердое тело, состоящее из таких квантованных осцилляторов, окружено газом и находится в тепловом равновесии с ним, причем температура газа

*) Но если чертить эту зависимость в относительных координатах T/θ , как это сделано на рис. 1.11, то $c = f(T/\theta)$ сливаются для различных тел в одну универсальную кривую.

***) Понятие кванта и величина h (названная постоянной Планка) уже за десять лет до этого были введены Планком в теорию электромагнитного излучения.

****) Это состояние (так называемых нулевых колебаний) соответствует температуре абсолютного нуля.

100° К. При такой температуре средняя кинетическая энергия атомов газа ($\frac{3}{2}kT \approx 0,015 \text{ эв}$) будет много меньше ϵ_0 . (Значение энергии, равное $\frac{3}{2}kT$, нанесено на графике штриховой линией.) Если бы все атомы газа обладали такой энергией, недостаточной для перевода осцилляторов в первое возбужденное состояние, то колебания твердого тела вовсе не возбуждались бы и его теплоемкость равнялась бы нулю.

В действительности энергии атомов газа лишь в среднем равны $\frac{3}{2}kT$, и всегда имеется небольшое число атомов, энергия которых много больше средней. Эти атомы при соударениях с твердым телом будут заставлять колебаться небольшое число его атомов, поэтому теплоемкость твердого тела будет отлична от нуля, но много меньше величины, предсказываемой законом Дюлонга и Пти. Чем выше температура, тем большее количество молекул газа будет в состоянии возбуждать колебания и тем больше будет теплоемкость твердого тела. Наконец, при некоторой температуре, при которой средняя энергия молекул газа станет больше ϵ_0 , практически все атомы твердого тела перейдут в первое возбужденное состояние; при более высоких температурах дискретность энергетического спектра уже не будет чувствоваться, энергия осциллятора будет возрастать линейно и теплоемкость его достигнет значения $3R$, предсказываемого классической теорией. Более строгое рассмотрение показывает, что это произойдет при температуре θ , определяемой уравнением

$$k\theta = h\nu \text{ или } \theta = \frac{h\nu}{k}. \quad (1.68)$$

Эта температура называется температурой Дебая. Эйнштейн оформил изложенные выше качественные соображения количественно и вывел формулу для молекулярной теплоемкости твердого тела

$$C = \frac{3R \left(\frac{h\nu}{kT} \right)^2 e^{-\frac{h\nu}{kT}}}{\left(e^{\frac{h\nu}{kT}} - 1 \right)^2}. \quad (1.69)$$

Как видно из (1.69), при $kT \ll h\nu$

$$C \approx 3R \left(\frac{h\nu}{kT} \right)^2 e^{-\frac{h\nu}{kT}} \ll 3R, \quad (1.70)$$

а при $kT \gg hv$

$$C \approx 3R.$$

(1.71)

На рис. 1.11 штриховой линией показана температурная зависимость теплоемкости, построенная согласно (1.69). Как видно из рисунка, при высоких температурах она дает хорошее согласие с опытными данными, а при низких — слишком крутое падение.

Дебай в 1912 г., анализируя причины расхождения формулы Эйнштейна с опытными данными, обратил внимание на то, что в твердом теле распространяются упругие колебания не с одной длиной волны, а с самими различными — от $\lambda = 2a$ (где a — постоянная решетки) до $\lambda = 2L$ (L — длина всего кристалла). В соответствии с (1.61) частота таких колебаний также колеблется в очень широких пределах; частота колебаний самых коротких волн равна частоте ν_0 колебаний в модели Эйнштейна, частота самых длинных в отношении a/L ниже. Энергия каждой упругой волны, по Дебаю, может меняться также только дискретно на величину $\pm hv$, где ν — частота колебаний. Следовательно, энергетические уровни для длинных волн будут расположены во много раз чаще, чем для коротких (рис. 1.12, в). Поэтому длинноволновые колебания будут возбуждаться при сравнительно низких температурах и возрастание теплоемкости начнется раньше, чем это следует из теории Эйнштейна. Дебай вывел формулу для теплоемкости твердого тела с учетом этих соображений, которая дала прекрасное совпадение с опытными данными (см. гл. 2).

Приведенные выше представления о квантовании энергии упругих волн позволяют нам сделать еще один шаг во введение понятий квантовой механики. Оказывается, что изменение энергии колебаний (волны) на один квант можно рассматривать как появление или исчезновение особой частицы — фонона, имеющего энергию $\epsilon = hv$ и движущегося со скоростью звука. Если колебание с частотой ν_i возбуждено до какого-либо n -го уровня, то на корпускулярном языке это означает, что в кристалле имеется n частиц i -го сорта, движущихся в направлении распространения данной волны. В дальнейшем выяснилось, что фонону можно также приписать и импульс, связанный с волновым вектором волны соотношением

$$p = \hbar q.$$

(1.72)

Таким образом, фононы приобрели не только ряд свойств, похожих на свойства обычной частицы (наличие энергии, импульса), но и свойства, во многом отличные (постоянство скорости, энергии и значения импульса, способность «исчезать» и появляться при изменении энергии колебаний). Направление импульса фонона может меняться, и это соответствует изменению направления распространения волны после того, как произошел акт рассеяния.

Эти представления, введенные вначале чисто формально, в дальнейшем оказались чрезвычайно плодотворными. Оказалось, что почти все процессы, связанные с упругими колебаниями решетки, можно перевести на фононный язык: рассеяние электрона на тепловых колебаниях решетки можно рассматривать как столкновение с фононом, причем при таком столкновении соблюдаются законы сохранения энергии и импульса. Точно так же рассеяние (благодаря ангармоничности) одной упругой волны на другой можно рассматривать как столкновение фонона с фононом, рассеяния волн на дефектах — как столкновение фононов с дефектами, явление теплопроводности — как поток фононов от горячего конца тела к холодному и т. д. При этом все закономерности приобретают более простой и удобный для вычислений вид. Более того, ряд явлений оказалось возможным объяснить только с помощью фононных представлений. Поэтому эти представления теперь следует рассматривать не как формальный прием, а наоборот, как одно из проявлений сложной (двойственной) природы материи.

Величина третьего множителя в выражении (1.64) для теплопроводности кристаллической решетки — длины свободного пробега \bar{l} упругой волны (или, как теперь мы можем сказать, фонона), как уже упоминалось, ограничивается при низких температурах рассеянием фононов на дефектах, а при высоких — фононов на фононах. Чем больше коэффициент ангармоничности g в (1.63), тем больше вероятность актов рассеяния; на фононном языке это означает: тем больше поперечное сечение фонона.

С другой стороны, чем ниже температура Дебая для данного тела, тем больше будет число фононов при какой-то заданной температуре, тем больше будет, следовательно, актов соударения. Поэтому длина свободного пробега фононов в данном материале будет тем меньше, чем больше

коэффициент ангармоничности в (1.63) и чем ниже температура Дебая.

Эти качественные соображения, развитые А. Ф. Иоффе, удовлетворительно согласуются с опытными данными.

1.5. КОНТАКТНЫЕ ЯВЛЕНИЯ

РАБОТА ВЫХОДА

Для того чтобы оторвать от изолированного атома валентный электрон, необходимо затратить некоторую работу, т. е. сообщить электрону энергию, необходимую для преодоления сил притяжения. Эта энергия, выраженная в электрон-вольтах, называется (для газов) ионизационным потенциалом. Соответствующее понятие для твердого тела носит название работы выхода. Качественно происхождение работы выхода можно объяснить следующим образом.

На рис. 1.13, *a* изображены схематически энергетические спектры трех изолированных атомов одновалентного металла; штриховой линией обозначена потенциальная энергия электрона в электрическом поле ядра, сплошными горизонтальными линиями — энергетические уровни валентного электрона и следующий за ним свободный уровень. Стрелкой обозначен ионизационный потенциал.

На рис. 1.13, *б* показано, как модифицируется спектр электрона в металле. В этом случае энергетические барьеры между соседними атомами настолько снижаются, что валентные электроны приобретают возможность свободно перемещаться вдоль всего кристалла; из валентных уровней образуется наполовину заполненная валентная зона, из свободных уровней — следующая за ней свободная зона.

Однако, как видно из рисунка, ход потенциала на поверхности металла остается приблизительно таким же, как у изолированного атома: для того чтобы «вырвать» валентный электрон из металла, необходимо затратить значительную работу (обозначенную на чертеже стрелкой). Как уже упоминалось выше, эту работу и называют работой выхода.

Нарисованная выше картина чрезвычайно схематична. В действительности, в работу выхода входит целый ряд дополнительных составляющих: энергия взаимодействия отрываемого электрона со всеми (и в первую очередь с ближайшими) атомами кристалла (это взаимодействие,

как уже говорилось, приводит к смещению и расщеплению атомных уровней), энергия взаимодействия со свободными электронами (которая, в свою очередь, состоит из ряда компонент: энергии преодоления на поверхности двойного слоя, образующегося за счет давления электронного газа,

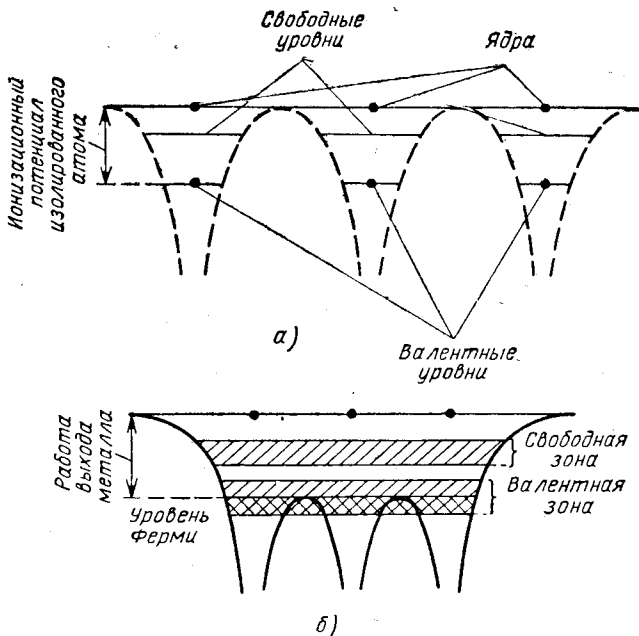


Рис. 1.13. Схема возникновения работы выхода:
 а — энергетический спектр электрона в изолированных атомах;
 б — энергетический спектр электрона в кристалле.

энергии преодоления сил зеркального изображения, энергии отталкивания за счет кулоновских сил и принципа Паули), состояние поверхности и др.

Упомянутые выше вопросы детально рассматриваются в курсах электроники; для нас же сейчас важно, что в результате перечисленных выше причин на поверхности металла для электрона образуется потенциальный барьер, для преодоления которого необходимо совершить некоторую работу.

Итак, согласно рис. 1.13, б работой выхода мы называем минимальную энергию, необходимую для удаления

электрона из металла, и измеряется она расстоянием от самого верхнего уровня, занятого электронами в металле, т. е. от уровня химического потенциала до потенциала электрона в пустоте.

Однако это определение будет строгим лишь при абсолютном нуле. При температуре, отличной от абсолютного нуля, на любом уровне выше уровня химического потенциала скапливается некоторое число электронов *), и такое определение работы выхода становится неточным. Однако если мы удалим некоторое количество электронов с какого-либо уровня, находящегося ниже уровня химического потенциала, то равновесие электронов в металле нарушится: начнутся перебросы электронов с более высоких уровней на освободившиеся и металл нагреется за счет освободившейся при этом энергии. Поэтому затраченную в этом случае работу нельзя называть работой выхода, так как часть ее пойдет на нагрев металла. Если же мы будем удалять электроны с уровня, находящегося выше уровня Ферми, процесс пойдет в обратном направлении: электроны будут переходить в большем количестве снизу вверх, чем сверху вниз, и в процессе восстановления равновесия металл охладится. Затраченная в этом случае работа будет меньше работы выхода, так как на вырывание электрона из металла, кроме этого, уйдет часть его внутренней энергии.

Учитывая все вышесказанное не только при абсолютном нуле, но и при температуре, отличной от абсолютного нуля, работу выхода отсчитывают от уровня химического потенциала и называют ее *изотермической* работой выхода.

Еще более сложным становится вопрос об определении работы выхода полупроводника: на вырывание электрона в вакуум из зоны проводимости требуется одна работа, из заполненной зоны — другая, с примесных уровней — третья. Однако и в этом случае можно показать, что если

*) При температуре, отличной от абсолютного нуля, в вакууме вокруг металла тоже имеется некоторое количество электронов — металл заряжен положительно и окружен отрицательной электронной атмосферой, плотность которой убывает по закону Больцмана по мере удаления от поверхности. При этом в условиях равновесия уровень химического потенциала для электронов и в металле и в вакууме будет одинаков, а потенциал будет возрастать по мере удаления от металла на расстояние порядка дебаевской длины.

мы будем удалять электроны с любого уровня, находящегося выше уровня химического потенциала, то материал будет охлаждаться; при удалении электронов с уровня, находящегося ниже уровня химического потенциала, он будет нагреваться. И только в том случае, когда мы будем одновременно удалять и те и другие электроны, причем в такой пропорции, чтобы их энергия в среднем равнялась энергии на уровне химического потенциала, температура останется неизменной. Поэтому и в полупроводниках изотермическую работу выхода отсчитывают от уровня химического потенциала, несмотря на то, что на этом уровне большей частью нет ни одного электрона, так как обычно он проходит где-то в запрещенной зоне (см. гл. 4).

Электрон может получить энергию, необходимую для того, чтобы покинуть металл (или полупроводник), различными способами: от падающего на поверхность металла света (на этом основаны вакуумные и газонаполненные фотоэлементы), от «бомбардирующих» поверхность быстрых электронов (на этом основаны фотоумножители) или других частиц и т. д. Эта энергия может быть также получена за счет внутренней энергии самого металла. Действительно, при достаточно высокой температуре распределение Ферми настолько «расплывзется», что в нем появится значительное количество электронов, энергия которых больше работы выхода; если какой-либо из таких электронов появляется вблизи поверхности металла и его скорость при этом направлена нормально к поверхности, то он может покинуть металл.

На этом явлении основана эмиссия электронов катодов электронных ламп. Работа выхода оксидных катодов значительно меньше, чем вольфрамовых, поэтому они требуют значительно меньшего подогрева, и в этом их основное преимущество.

КОНТАКТНАЯ РАЗНОСТЬ ПОТЕНЦИАЛОВ

Рассмотрим теперь, что происходит, когда мы приближаем друг к другу два различных металла. Не только при высокой, но и при любой температуре всегда имеется некоторое количество электронов, энергия которых достаточна для того, чтобы покинуть металл; чем меньше работа выхода, тем больше число таких электронов. Поэтому при соприкосновении металлов начнется обмен электро-

нами между ними *). Предположим, для определенности, что работа выхода первого металла меньше, чем второго (рис. 1.14, а) **). Тогда электронный ток слева направо $I_{s1} = Ae^{-W_1/kT}$ будет больше, чем ток в обратном направлении $I_{s2} = Ae^{2W_2/kT}$, поверхность второго металла будет заряжаться отрицательно, а на поверхности первого будет оставаться нескомпенсированный положительный заряд ***). Этот процесс будет продолжаться до тех пор,

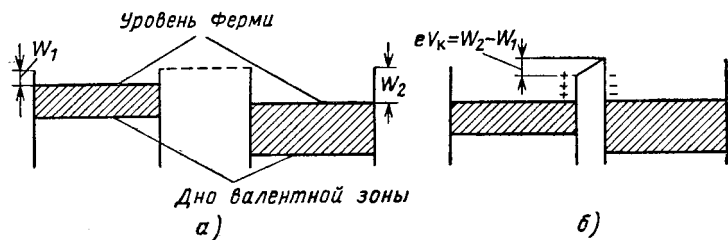


Рис. 1.14. Схема возникновения контактной разности потенциалов

пока возникшая вследствие этого в зазоре (который становится подобен плоскому конденсатору) разность потенциалов не компенсирует разность работ выхода $eV_k = W_2 - W_1$ (рис. 1.14, б).

Оценим количество электронов, которое должно перейти из одного металла в другой, чтобы создать равновесную разность потенциалов V_k . Предположим, что разность работ выхода $V_k = 1$ в и зазор между металлами $d = 10^{-7}$ см. Тогда поле в зазоре $E = 10^7$ в/см и поверхностная плотность заряда, необходимая, чтобы создать такое поле в зазоре, $q = E/4\pi$. Следовательно, количество электронов, которое должно перейти с 1 см² одного металла в другой, чтобы создать это поле, будет

$$n = \frac{q}{e} = \frac{E}{4\pi e} \approx 10^{12}. \quad (1.73)$$

*) Кроме этого, при достаточном сближении металлов появляется возможность перехода электронов из одного металла в другой за счет туннельного эффекта.

**) Здесь скачок потенциала на поверхности металлов для простоты показан вертикальным. Свободные зоны на рисунке не показаны.

***) Коэффициент зависит от плотности состояний в обоих металлах, от температуры и ряда других факторов; в данном случае существенно, что он одинаков в $I_{s1} = I_{s2}$.

Число атомов на поверхности металла $\sim 10^{16}$; следовательно, в первом металле всего лишь приблизительно 1% поверхностных атомов будет ионизован, а во втором такое же количество атомов получит избыточный заряд. После этого уровни химического потенциала сравняются. При этом электронный ток справа налево останется прежним, а поток электронов слева направо уменьшится, так как, для того чтобы перейти из первого металла во второй, электрону надо преодолеть не только работу выхода W_1 , но и разность потенциалов в зазоре V_K :

$$I_{s1} = A \exp\left(-\frac{eV_K + W_1}{kT}\right) = I_{s2}, \quad (1.74)$$

и результирующий ток $I = I_{s1} - I_{s2} = 0$.

Посмотрим теперь, что произойдет, когда к контакту будет приложена некоторая разность потенциалов V от постороннего источника электродвижущей силы.

ВЫПРЯМЛЕНИЕ НА КОНТАКТЕ ДВУХ МЕТАЛЛОВ

Рассмотрим два случая: первый, когда внешняя разность потенциалов V складывается с контактной (рис. 1.15, б) и полное падение напряжения на зазоре равно их сумме $V_1 = V_K + V$, и второй, когда внешняя разность потенциалов приложена в обратном направлении $V_2 = V_K - V$ (рис. 1.15, в) *).

Как видно из рисунка, в обоих случаях электронный ток справа налево останется неизменным $I_2 = I_s$, а ток слева направо в первом случае уменьшится:

$$I_1 = A \exp\left(-\frac{W_1 + e(V_K + V)}{kT}\right) = I_s e^{-\frac{eV}{kT}}, \quad (1.75)$$

а во втором — увеличится:

$$I_1 = A \exp\left[-\frac{W_1 + e(V_K - V)}{kT}\right] = I_s e^{\frac{eV}{kT}}; \quad (1.76)$$

результирующий ток в первом случае будет направлен справа налево:

$$I = I_1 - I_2 = I_s (1 - e^{-\frac{eV}{kT}}), \quad (1.77)$$

а во втором — слева направо:

$$I = I_2 - I_1 = I_s (e^{\frac{eV}{kT}} - 1). \quad (1.78)$$

*) В последнем случае мы будем предполагать, что $V < V_K$.

Можно объединить формулы (1.77) и (1.78), если току и напряжению будем приписывать положительный знак, когда они направлены слева направо, и отрицательный — для обратного направления; тогда в обоих случаях

$$I = I_s \left(e^{\frac{eV}{kT}} - 1 \right). \quad (1.79)$$

Как видно из (1.79), контакт обладает выпрямляющими свойствами: при положительном знаке V ток экспонен-

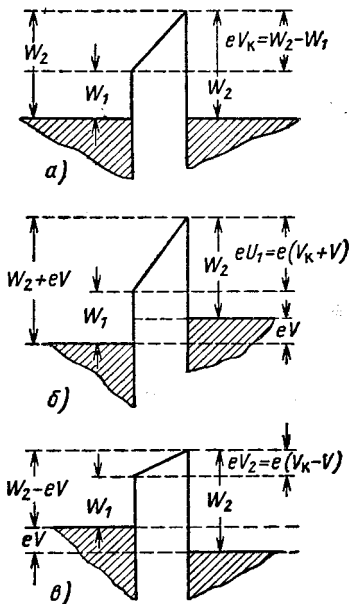


Рис. 1.15. Схема контакта двух металлов:

a — в отсутствие внешнего напряжения; $б$ — при напряжении, приложенном в запиорном направлении $eU_1 = eV_K + eV$; $в$ — при напряжении, приложенном в пропускном направлении $eV_2 = eV_K - eV$.

циально возрастает, при отрицательном — асимптотически приближается к I_s .

Полученная выше формула, по крайней мере качественно, описывает выпрямление на контакте полупроводника и металла и на контакте двух полупроводников.

Единственный случай, когда она до недавнего времени не подтверждалась на практике, это тот, для которого мы ее вывели: контакт двух металлов при обычных условиях не обладает выпрямляющим свойством.

Объясняется это тем, что при соприкосновении двух металлов зазор между ними становится очень тонким

и поэтому прозрачным для туннельного эффекта; при этих условиях сопротивление контакта становится чрезвычайно малым и одинаковым для обоих направлений тока.

Если между металлами поместить очень тонкий слой диэлектрика и таким образом увеличить ширину зазора, то туннельный эффект можно исключить, и тогда все выводы, сделанные выше, будут справедливы. Такие опыты действительно производились и подтвердили теорию, развитую выше. За последнее время начали разрабатываться специальные типы выпрямителей и нелинейных сопротивлений, основанных на выпрямляющих свойствах контакта двух металлов, разделенных изоляционной прослойкой. Слой, обедненный электронами, достаточно толстый, чтобы исключить туннельный эффект, образуется при определенных условиях на контакте полупроводника и металла. Этот случай мы и рассмотрим ниже.

КОНТАКТ ПОЛУПРОВОДНИКА И МЕТАЛЛА

Условия равновесия. Рассмотрим контакт электронного полупроводника и металла, причем предположим, что работа выхода полупроводника меньше работы выхода металла (рис. 1.16, а).

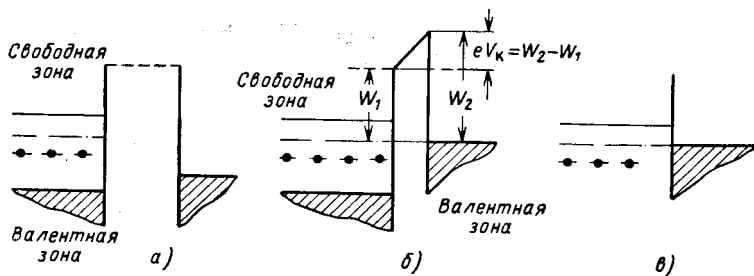


Рис. 1.16. Контакт электронного полупроводника и металла:

а — положение энергетических уровней полупроводника и металла до установления равновесия; б — схема распределения потенциала на контакте полупроводник — металл в равновесии без учета объемного заряда; в — контакт электронный полупроводник — металл в наиболее схематичном виде (без объемного заряда и вакуумного промежутка).

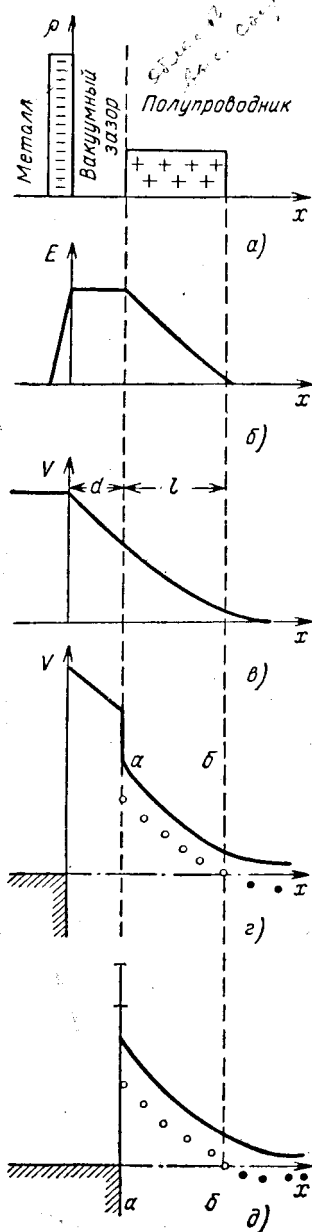
В этом случае в первый момент поток электронов из полупроводника будет больше, чем из металла, и металл начнет заряжаться отрицательно, а полупроводник —

положительно; этот процесс будет продолжаться до тех пор, пока возникшая вследствие этого разность потенциалов не скомпенсирует разность работ выхода. Таким образом, схематически процесс установления равновесия будет происходить так же, как на контакте двух металлов (рис. 1.16, б). Однако картина, представленная на этом рисунке, является лишь очень грубой схемой, применимой, впрочем, для рассмотрения ряда контактных явлений (например, явления Пельтье *) , но недостаточной для понимания явления выпрямления.

Существенное отличие от контакта двух металлов возникает в данном случае вследствие того, что количество свободных электронов в полупроводнике (даже если включить сюда все электроны, локализованные на донорах) во много раз меньше, чем в металле, и вследствие этого слой, обедненный электронами, в полупроводнике будет во много раз толще.

Проиллюстрируем это следующим примером. Пусть контактная разность потенциалов и толщина зазора будут те же, что и в рассмотренном выше случае двух металлов $V_K = 1$ в и $d = 10^{-7}$ см. Тогда, повторив те же расчеты, мы получим число электронов, перешедших из полупроводника в металл, равное $5 \cdot 10^{12}$. Предположим, что число доноров в единице объема полупроводника $5 \cdot 10^{16}$, тогда, для того чтобы создать такой заряд, нужно ионизировать все доноры на значительном расстоянии от поверхности ($d_1 = 10^{-4}$ см), в 1000 раз большем, чем толщина зазора. Следовательно, в полупроводнике возникнет значительный слой объемного заряда, в котором плотность заряда будет постоянна, поле будет возрастать линейно, потенциал — квадратично (рис. 1.17). Основное падение потенциала будет, таким образом, сосредоточено не в зазоре, а в слое объемного заряда. Учитывая то, что зазор прозрачен для туннельного эффекта, мы в дальнейшем будем рисовать схему контакта полупроводника и металла так, как она представлена на рис. 1.17. На этом рисунке слой *a* — *б* настолько удален от уровня Ферми, что в нем почти нет свободных электронов. Этот слой, обладающий очень большим сопротивлением, называют запорным слоем.

*) Так как зазор между полупроводником и металлом прозрачен для туннельного эффекта, то при рассмотрении явления Пельтье схему контакта можно еще более упростить (рис. 1.16, в).



В том случае, когда работа выхода полупроводника больше, чем работа выхода металла, электроны переходят из металла в полупроводник и на контакте электронного полупроводника и металла образуется слой $a - б$ (рис. 1.18), обогащенный носителями, — антизапорный слой. Ход потенциала в этом случае изображен на рис. 1.18.

Антизапорный слой не обладает выпрямляющими свойствами, и мы его в дальнейшем рассматривать не будем. На границе дырочного полупроводника и металла запиорный слой образуется в том случае, когда работа выхода полупроводника больше, чем металла, и антизапорный — при обратном соотношении.

Во всех рассмотренных выше случаях запиорные слои образовались за счет ухода электронов (или дырок) в металл. Такие запиорные слои называются физическими. В ряде случаев за счет химического взаимодействия полупроводника и металла или специальной обработки поверхности полупроводника на контакте образуется слой нового

Рис. 1.17. Контакт электронного полупроводника и металла (с учетом объемного заряда):

a — распределение заряда; $б$ — ход электрического поля; $в$ — ход электрического потенциала (без учета скачков потенциала на поверхности полупроводника и металла); $г$ — ход потенциала с учетом скачков потенциала на поверхности полупроводника и металла; $д$ — упрощенная зонная диаграмма контакта (без учета вакуумного зазора).

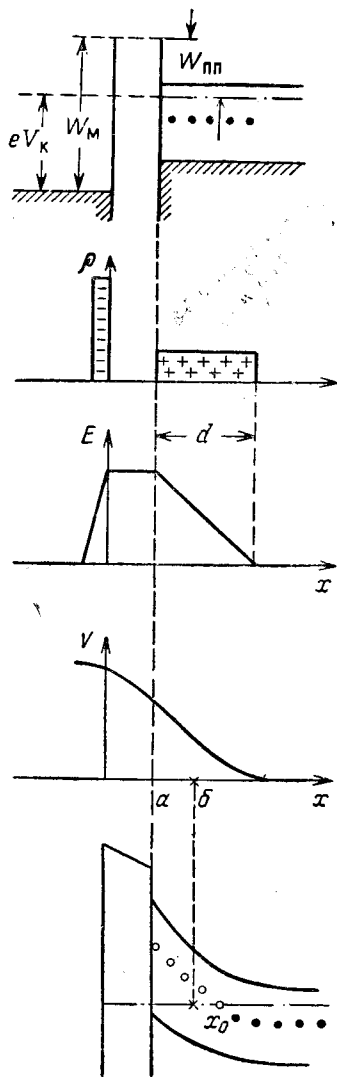
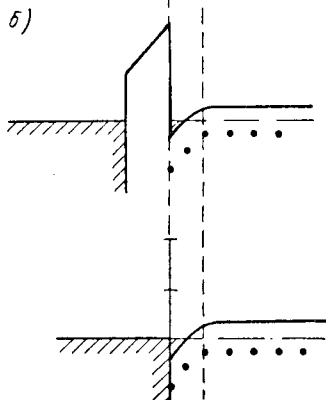
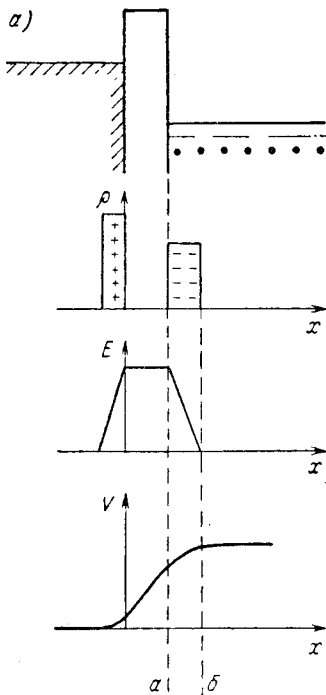


Рис. 1.18. Антизапорный слой на границе электронный полупроводник — металл:
 а — до установления равновесия; б — после установления равновесия.

Рис. 1.19. Схема образования p-n перехода на границе полупроводник — металл.

вещества, по всей природе являющегося диэлектриком. Такие запирающие слои называются химическими.

В заключение этого раздела еще раз вернемся к рассмотрению запирающего слоя на контакте электронного полупроводника и металла, но предположим при этом, что разность работ выхода больше ширины запрещенной зоны полупроводника (рис. 1.19).

При этих условиях на границе с металлом образуется слой дырочного полупроводника (слой $a - b$ на рис. 1.19) с относительно низким сопротивлением; на некотором расстоянии от границы полупроводник становится электронным, в точке x_0 имеет место переход от дырочной проводимости к электронной — так называемый $p-n$ переход. В этом случае выпрямление будет происходить не на границе полупроводника с металлом, а на $p-n$ переходе.

ВЫПРЯМЛЕНИЕ НА КОНТАКТЕ ПОЛУПРОВОДНИКА И МЕТАЛЛА

Процесс протекания тока через выпрямляющий контакт существенно зависит от того, испытывают ли электроны соударения при прохождении через запирающий слой или нет. Первый случай реализуется тогда, когда длина свободного пробега l электрона меньше толщины запирающего слоя d , второй — при обратном соотношении.

В соответствии с этим существует две теории выпрямления: диффузионная ($l \ll d$) и диодная ($l \gg d$) (см. гл. 8).

Начнем с рассмотрения тонкого запирающего слоя $d \ll l$. В этом случае механизм протекания тока через запирающий слой ничем не отличается от протекания тока через вакуумный промежуток между двумя металлами *). Поэтому теория выпрямления имеет точно такой же вид, как и в рассмотренном выше случае контакта двух металлов. Введем обозначения: I_1 — ток справа налево (из полупроводника в металл); I_2 — ток слева направо (из металла в полупроводник).

В отсутствие внешней разности потенциалов оба тока равны:

$$I_{s1} = I_{s2} = I_s,$$

и результирующий ток $I = I_{s1} - I_{s2} = 0$ (рис. 1.20, а).

*) С тем лишь отличием, что в данном случае тонкий запирающий слой все же достаточно толст, чтобы исключить туннельный эффект.

Рассмотрим, как и ранее, два случая: первый, когда внешняя разность потенциалов складывается с контактной (рис. 1.20, б), и второй, когда внешняя разность потенциалов приложена в противоположную сторону (рис. 1.20, в). Как видно из рисунков, ток из металла в полупроводник в обоих случаях остается неизменным

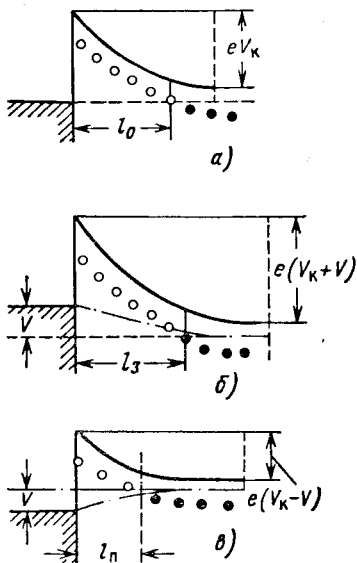


Рис. 1.20. Схема выпрямления тока на контакте полупроводник — металл.

$I_2 = I_s$ (так как в металле высота барьера, отсчитываемая от уровня Ферми, остается неизменной).

Для электронов, переходящих из полупроводника в металл, барьер в первом случае повышается: $V_1 = V_k + V$, и ток

$$I_1 = I_s e^{-\frac{eV}{kT}}, \quad (1.80)$$

во втором высота барьера понижается и ток

$$I_1 = I_s e^{\frac{eV}{kT}}. \quad (1.81)$$

Повторив те же выводы, которые делались при рассмотрении контакта двух металлов (и сохранив обозначения для полярности тока и напряжения), получим выра-

жение, аналогичное (1.79), для вольтамперной характеристики контакта:

$$I = I_s \left(e^{\frac{eV}{kT}} - 1 \right). \quad (1.82)$$

Диодная теория (см. гл. 8) дает следующее выражение для тока насыщения:

$$I_s = \frac{1}{4} enu_0 e^{-\frac{eV_K}{kT}}, \quad (1.83)$$

где n — концентрация носителей в объеме полупроводника;
 $n_0 = n e^{-eV_K/kT}$ — концентрация электронов вблизи поверхности;
 u_0 — их средняя тепловая скорость;
 eV_K — высота потенциального барьера.

Таким образом,

$$I = \frac{1}{4} enu_0 e^{-\frac{eV_K}{kT}} \left(e^{\frac{eV}{kT}} - 1 \right). \quad (1.84)$$

В случае, когда длина свободного пробега электрона меньше толщины запирающего слоя d (или точнее $l \ll d$), феноменологическая вольтамперная характеристика контакта остается прежней и выражение (1.82) для тока остается в силе.

Однако механизм прохождения тока через контакт при этом совершенно иной *): в этом случае электроны испытывают в запирающем слое множество соударений и результирующий электрический ток является разностью тока, создаваемого электрическим полем в запирающем слое, которое «гонит» электроны по направлению от контакта, и диффузионного тока, возникающего в результате наличия градиента концентрации носителей. Так как концентрация носителей тока вблизи контакта меньше, чем в объеме полупроводника, то диффузия создает ток в обратном направлении.

Теория дает следующее выражение для диффузионного тока электронов:

$$I_d = -eDn \frac{\partial n}{\partial x}, \quad (1.85)$$

*) А поэтому и выражение для тока насыщения будет другим — см. ниже.

где D_n — коэффициент диффузии электронов, связанный с подвижностью соотношением Эйнштейна:

$$eD_n = kTu_n. \quad (1.86)$$

Таким образом, при наличии электрического поля и градиента концентрации полный электрический ток складывается из дрейфового и диффузионного:

$$I_n = I_0 + I_d = e \left(Enu_n - D_n \frac{\partial n}{\partial x} \right)^*. \quad (1.87)$$

При отсутствии внешнего поля дрейфовый и диффузионный ток равны друг другу и результирующий ток равен нулю.

Если к контакту приложена внешняя разность потенциалов V в заперном направлении, т. е. внешнее поле складывается с контактным, то омический ток увеличивается и в предельном случае, когда он становится много больше диффузионного, ток через контакт равен

$$I = I_s = en_0 \bar{u}_n E = en_0 \bar{u}_n \frac{V_K + V}{l}, \quad (1.88)$$

где l — толщина заперного слоя.

При противоположном направлении поля дрейфовый ток уменьшается и разность диффузионного и дрейфового тока экспоненциально возрастает (см. гл. 8). Как мы уже упоминали, феноменологическое выражение для вольтамперной характеристики и в этом случае остается прежним (1.82), но (см. гл. 8)

$$I_s = en_0 \bar{u}_n E = en_0 \bar{u}_n \frac{V_K \mp V}{l}. \quad (1.89)$$

Учитывая, что произведение $\bar{u}_n E = v_d$ равно дрейфовой скорости электронов в поле заперного слоя, можно придать формуле (1.89) вид, идентичный (1.83), с точностью до коэффициента $1/4$:

$$I_s = en_0 v_d. \quad (1.90)$$

Однако существенное качественное и количественное различие выражений (1.83) и (1.89) заключается в следующем:

— при всех полях, которые могут существовать в заперном слое (вплоть до таких, при которых наступает

*) При наличии двух знаков носителей для дырочного тока выражение для I_p принимает вид, аналогичный (1.87).

пробой выпрямителя), дрейфовая скорость электронов меньше тепловой. Поэтому ток насыщения толстого запорного слоя много меньше, чем тонкого;

— с другой стороны, согласно (1.89) и (1.90) ток насыщения толстого запорного слоя зависит от приложенного напряжения и растет для запорного направления.

КОНТАКТ ДЫРОЧНОГО И ЭЛЕКТРОННОГО ПОЛУПРОВОДНИКА

В подавляющем большинстве выпрямителей, применяемых в настоящее время, выпрямление происходит не на контакте полупроводника и металла, а на контакте дырочного и электронного полупроводника — $p-n$ переходе. Поэтому этот случай имеет наибольшее научное и практическое значение.

Можно создать $p-n$ переход различными путями:

1) вблизи контакта полупроводника и металла, если ширина запрещенной зоны полупроводника меньше разности работ выхода полупроводника и металла (см. выше);

2) на поверхности полупроводника за счет примесных уровней и дефектов;

3) на границе двух различных полупроводниковых материалов с различным характером проводимости;

4) в объеме одного и того же материала, если в одной части он легирован донорной примесью, а в другой — акцепторной (или, если мы имеем дело с образцом с компенсированными примесями, в одной части объема имеется избыток доноров, а в другой — акцепторов).

Предположим вначале для простоты, что $p-n$ переход образуется путем создания механического контакта между двумя слитками (дырочного и электронного германия *).

Итак, пусть два слитка германия (дырочный и электронный) введены в соприкосновение друг с другом (рис. 1.21). Так как работа выхода для электронов в n -германии меньше, чем в p -германии, то с момента создания контакта возникает поток электронов слева направо

*) В реальных выпрямителях $p-n$ переход создается путем введения (вплавлением или диффузией) в примесный электронный полупроводник акцепторной примеси или в примесный дырочный полупроводник — донорной примеси. Однако избранный путь рассуждения нам представляется более наглядным.

и в n -германии создается слой положительного объемного заряда (с полностью ионизованными донорами), а в p -германии — слой отрицательного объемного заряда (с полностью занятыми акцепторами) (рис. 1.21, б). Переход электронов будет продолжаться до тех пор, пока создавшаяся таким образом разность потенциалов (рис. 1.21, г) не скомпенсирует разность работ выхода и уровни Ферми p - и n -слитков не сравняются друг с другом (рис. 1.21, д) (при этом заряды отрицательного и положительного слоя будут одинаковы, но толщина их может быть различной). Основное отличие p - n перехода от рассмотренного раньше контакта полупроводника и металла заключается в том, что здесь возможны и дырочный и электронный токи, причем в равновесии потоки электронов и слева направо и справа налево (и соответствующие потоки дырок) должны быть одинаковы. Нетрудно убедиться, что при равенстве уровней химического потенциала оба эти условия одновременно удовлетворяются. Действительно, каждый электрон из p -области может беспрепятственно перейти в n -область (так как контактное поле будет гнать его в этом направлении). Но число электро-

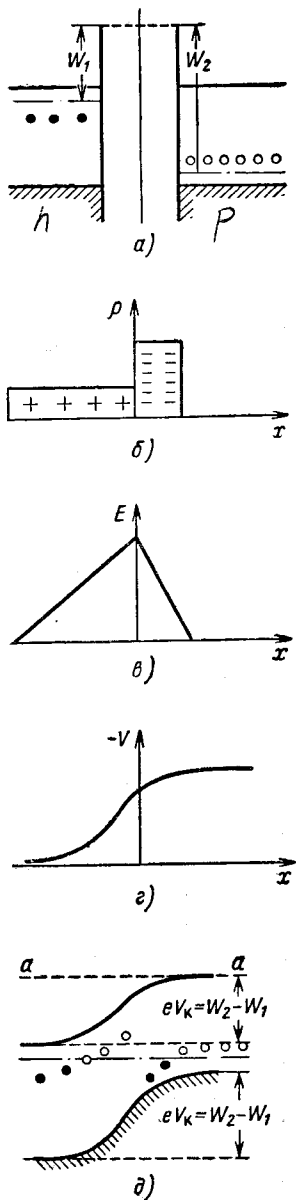


Рис. 1.21. Контакт дырочный полупроводник — электронный полупроводник (p - n переход):
 а — до установления равновесия; б, в, г, д — после установления равновесия (б — распределение зарядов; в — поле; г — ход потенциала; д — зонная диаграмма).

нов в p -области очень мало. Для того чтобы перейти из n -области в p -область, электрон должен иметь энергию, достаточную, чтобы преодолеть потенциальный барьер, и число таких электронов опять-таки будет определяться расстоянием от дна зоны до уровня $a - a$. Совершенно аналогичные рассуждения можно привести и для дырок.

ВЫПРЯМЛЕНИЕ НА p - n ПЕРЕХОДЕ

Диффузия неравновесных носителей. Представим себе, что в какой-либо точке или плоскости полупроводника каким-то образом создается избыточная по сравнению с равновесной концентрация носителей тока Δn_0 ; тогда избыточные носители будут диффундировать вглубь и по мере удаления от этой точки их концентрация будет спадать за счет рекомбинации по экспоненциальному закону

$$\Delta n(x) = \Delta n_0 e^{-\frac{x}{L}}, \quad (1.91)$$

где константа L — диффузионная длина, т. е. расстояние, на котором неравновесная концентрация уменьшается в e раз.

Теория показывает, что диффузионная длина выражается через коэффициент диффузии и время жизни носителей τ следующим образом:

$$L = \sqrt{D\tau}. \quad (1.92)$$

Упрощающие предположения. Для облегчения дальнейшего анализа предположим, что так как уровень Ферми в запиорном слое удален от свободной и заполненной зон, то концентрация и электронов и дырок в нем (запиорном слое) значительно меньше, чем во всем объеме полупроводника. Поэтому мы будем считать, что он представляет основное сопротивление для тока, а падением напряжения в толще полупроводника можно пренебречь.

Предположим далее, что диффузионные длины дырок и электронов значительно больше толщины запиорного слоя и что рекомбинация в нем, а также поверхностная рекомбинация на границе p - и n -области не играют существенной роли *).

*) Последнее условие не всегда соответствует реальной картине явления. На практике различают тонкие p - n переходы, для которых предположения, сделанные выше, правомерны, и толстые, для которых эти условия не соблюдаются.

При этих условиях можно упростить схему зонной структуры p - n перехода и представить ее так, как она изображена на рис. 1.22 (т. е. сжать всю область объемного заряда в одну вертикальную линию).

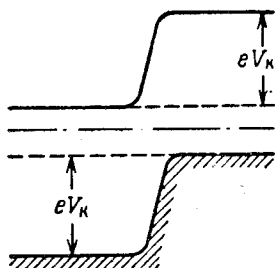


Рис. 1.22. Упрощенная схема зонной структуры p - n перехода.

Вольтамперная характеристика. Представим теперь, что к p - n переходу приложено напряжение V в пропускном

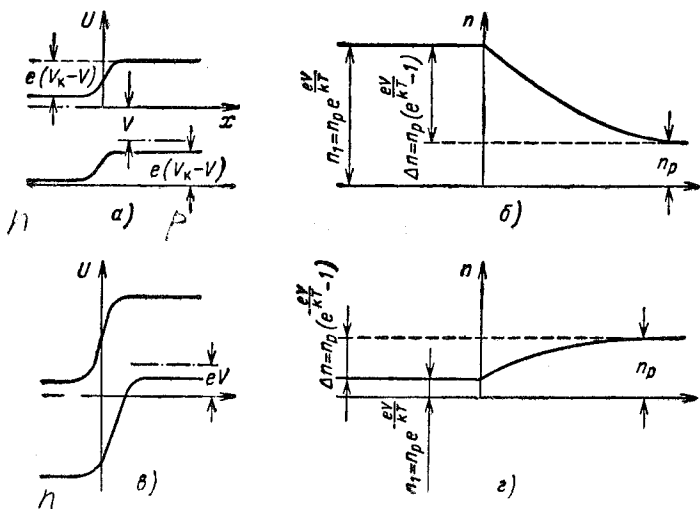


Рис. 1.23. Выпрямление на p - n переходе.

направлении, т. е. внешнее напряжение вычитается из контактной разности потенциалов и падение напряжения на нем $eV = e(V_K - V)$ (рис. 1.23, а).

Как видно из рис. 1.23, а и б, число электронов (n_p) в зоне проводимости p -области останется прежним, а сле-

довательно, не изменится и электронный ток справа налево, число же электронов n_1 в n -области, энергия которых будет достаточна для того, чтобы перейти в p -область, увеличится (так как уменьшится барьер):

$$n_1 = n_p e^{\frac{eV}{kT}}. \quad (1.93)$$

Вследствие этого электронное равновесие на границе p - и n -области нарушится. В результате наличия градиента концентрации в зоне проводимости возникнет диффузионный ток слева направо, электроны будут диффундировать в p -область и там постепенно рекомбинировать с дырками. Согласно (1.91) избыточная концентрация электронов в p -области будет убывать по закону

$$\Delta n(x) = (n_1 - n_p) e^{-\frac{x}{L_n}} = n_p \left(e^{\frac{eV}{kT}} - 1 \right) e^{-\frac{x}{L_n}}. \quad (1.94)$$

Выражение (1.94) позволяет нам вычислить диффузионный электронный ток:

$$I_n = e D_n \frac{dn}{dx} = e \frac{D_n n_p}{L_n} \left(e^{\frac{eV}{kT}} - 1 \right). \quad (1.95)$$

Это и будет полный электронный ток в p -области вблизи p - n перехода.

Аналогичным образом можно вычислить дырочный ток и n -области:

$$I_p = e \frac{D_p p_n}{L_p} \left(e^{\frac{eV}{kT}} - 1 \right). \quad (1.96)$$

Таким образом, полный ток через p - n переход будет равен

$$I = e \left(\frac{D_p p_n}{L_p} + \frac{D_n n_p}{L_n} \right) \left(e^{\frac{eV}{kT}} - 1 \right). \quad (1.97)$$

На достаточном расстоянии от границы концентрации электронов и дырок приходят к своим равновесным значениям, т. е. $dn/dx = dp/dx = 0$, и весь ток становится дрейфовым:

$$I = eE (u_n n + u_p p). \quad (1.98)$$

Однако в n -области $n_n \gg p_n$ и $I = eE u_n n_n$; в p -области, наоборот, $I = eE u_p p_n$. Возникает вопрос: каким обра-

*) Увеличение концентрации неосновных носителей при приложении напряжения в пропускном направлении называют инжекцией.

зом дрейфовый ток в толще полупроводника превращается в диффузионный в области p - n перехода? Качественно суть дела заключается в следующем. Электронный дрейфовый ток в n -области при подходе на диффузионное расстояние к p - n переходу начинает уменьшаться за счет рекомбинации с инжектированными в n -область дырками и до области объемного заряда уже доходит часть его $I_1 = I - I_p$; эта часть электронного тока инжектируется в виде диффузионного тока в p -область и там рекомбинирует с подходящим к границе дрейфовым дырочным током. Аналогичное рассуждение можно провести для дырочного тока в p -области.

Рассмотрим теперь, что происходит, если к p - n переходу приложено напряжение в запертом направлении (рис. 1.23, b и z). В этом случае (как и при пропускном направлении) концентрация электронов n_p в зоне проводимости p -полупроводника вдали от p - n перехода остается неизменной, концентрация же электронов n_1 в n -области с энергией, достаточной для перехода в p -область, уменьшается:

$$n_1 = n_p e^{-\frac{eV}{kT}}.$$

В этом случае электроны будут вытягиваться (экстрагироваться) из p -области в n -область, на границе возникает таким образом градиент концентрации и диффузионный ток в обратном направлении

$$I = e \left(\frac{D_p p_n}{L_p} + \frac{D_n n_p}{L_n} \right) \left(1 - e^{-\frac{eV}{kT}} \right). \quad (1.99)$$

Сохранив принятые выше обозначения для полярности тока, можно объединить формулы (1.97) и (1.99) в одну, как это делалось выше.

1.6. ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ

ТЕРМОДИНАМИЧЕСКИЕ СООТНОШЕНИЯ И КАЧЕСТВЕННАЯ КАРТИНА ТЕРМОЭЛЕКТРИЧЕСКИХ ЯВЛЕНИЙ

Термоэлектрическими явлениями называются явления термо-э. д. с. (Зеебека), Пельтье и Томсона.

Сущность явления термо-э. д. с. (открытого Зеебеком в 1821 г.) состоит в том, что в электрической цепи, состоя-

щей из последовательно сдвинутых разных материалов, возникает электродвижущая сила (термо-э. д. с.), если места контактов поддерживаются при различных температурах. В простейшем случае, когда такая цепь состоит из двух различных материалов, она носит название термоэлемента или термопары. Экспериментальные исследования показали, что термоэлектродвижущая сила термопары зависит от температуры горячего (T) и холодного (T_0) спаев и от состава материалов, образующих термопару.

В небольшом интервале температур величину термо-э. д. с. E можно считать, с достаточной для практических целей точностью, пропорциональной разности температур и некоторому коэффициенту α , называемому коэффициентом термо-э. д. с.:

$$E = \alpha_{1,2} (T - T_0), \quad (1.100)$$

более точно, $dE = \alpha_{1,2} dT$. Коэффициент термо-э. д. с. α определяется в первую очередь материалами ветвей термоэлемента, для обозначения которых мы ввели индексы 1 и 2, но зависит также и от интервала температур, в котором используется или исследуется термопара; во многих случаях с изменением температуры он даже меняет знак.

Явление Пельтье (открытое в 1834 г.) представляет собой эффект, обратный явлению Зеебека; сущность его состоит в том, что при прохождении тока в цепи, состоящей из различных проводников, в местах контакта в дополнение к теплоте Джоуля выделяется или поглощается (в зависимости от направления тока) некоторое количество тепла Q_{II} , пропорциональное прошедшему через контакт количеству электричества (т. е. силе тока и времени) и некоторому коэффициенту Π , зависящему от природы находящихся в контакте материалов, названному коэффициентом Пельтье:

$$Q_{II} = \Pi It. \quad (1.101)$$

Согласно (1.93) коэффициент Пельтье можно определить как количество тепла, выделяющегося (или поглощающегося) на контакте при прохождении через него единицы заряда.

В 1856 г. Томсон, применив к термоэлектрическим явлениям первое и второе начала термодинамики, вывел соотношение между коэффициентом термо-э. д. с. α и коэффициентом Пельтье Π :

$$\alpha = \frac{\Pi}{T}, \quad (1.102)$$

и предсказал существование третьего явления, названного в дальнейшем явлением Томсона.

Явление Томсона (установленное экспериментально Леру в 1867 г.) заключается в следующем: если вдоль проводника, по которому протекает электрический ток, существует перепад температур, то в дополнение к теплоте Джоуля в объеме проводника выделяется или поглощается, в зависимости от направления тока, некоторое количество тепла Q_τ (теплота Томсона), пропорциональное силе тока I , времени t , перепаду температур $(T - T_0)$ и коэффициенту (названному коэффициентом Томсона), зависящему от природы материала:

$$Q_\tau = \tau (T - T_0) It. \quad (1.103)$$

Согласно теории Томсона коэффициент термо-э. д. с. пары проводников связан с коэффициентами Томсона соотношением

$$\frac{d\alpha_{1,2}}{dT} = \frac{\tau_1 - \tau_2}{T}. \quad (1.104)$$

Оба соотношения Томсона (1.102) и (1.104) многократно подтверждались экспериментальными исследованиями. Качественно термоэлектрические явления можно объяснить следующим образом. Энергия свободных невырожденных электронов растет пропорционально температуре. Если вдоль проводника существует перепад температур, то электроны на горячем конце приобретают более высокие энергии и скорости, чем на холодном; в полупроводнике концентрация свободных электронов также растет с температурой и, следовательно, будет на горячем конце больше, чем на холодном *); в результате всего этого возникает поток электронов от горячего конца к холодному и на холодном конце накапливается отрицательный заряд, а на горячем остается нескомпенсированный положительный заряд.

Этот процесс накопления заряда будет продолжаться до тех пор, пока возникшая таким образом разность потенциалов не создаст поток электронов, движущийся в обратном направлении, равный первичному, и благодаря этому не

*) Температурная зависимость концентрации носителей дает равные и противоположные по знаку вклады в объемную и контактную термо-э. д. с.; поэтому из полной термо-э. д. с. цепи эти члены исключаются.

установится стационарное состояние. Алгебраическая сумма таких разностей потенциалов в цепи дает одну составляющую термоэлектродвижущей силы.

Вторая составляющая есть следствие температурной зависимости контактной разности потенциалов.

Если оба спая термоэлемента находятся при одной и той же температуре, то контактные разности потенциалов равны и направлены в противоположные стороны; в этом случае они компенсируют друг друга. Если же температура спаев различна, то разность ΔV_k на контактах дает свой «контактный» вклад в термо-э. д. с., который может быть сравним с объемной термо-э. д. с. *).

В металлах скорости электронов из-за вырождения очень слабо зависят от температуры, уровень химического потенциала и работа выхода также почти не зависят от температуры. Поэтому термо-э. д. с. металлов очень мала.

Описанные выше явления в дырочных проводниках протекают аналогично тому, как в электронных, с той лишь разницей, что на холодном конце скапливаются положительно заряженные дырки, а на горячем остается нескомпенсированный отрицательный заряд. В силу этого в термопаре, состоящей из дырочного и электронного проводника, термоэлектродвижущие силы ветвей складываются.

Наряду с примесными полупроводниками, в которых ток переносится носителями одного знака (либо только дырками, либо только электронами), существует большое количество веществ (и полупроводников, и металлов) со смешанной проводимостью, т. е. веществ, в которых электрический ток переносится одновременно и электронами и дырками. В последнем случае к холодному концу одновременно диффундируют и электроны и дырки, и их заряды компенсируют друг друга. Если концентрации и подвижности электронов и дырок равны, то электрическое поле вовсе при этом не образуется и термо-э. д. с. равна нулю **); таким материалом с нулевой термо-э. д. с. является, в частности, свинец. Если же числа свободных электронов и дырок либо их подвижности отличаются, то термо-э. д. с. возникает как разностный эффект, и величина ее обычно относительно мала.

*) Деление термо-э. д. с. на объемную и контактную не строго (см. гл. 6).

**) Положение уровня химического потенциала в этом случае также не меняется с изменением температуры.

Если вдоль проводника, по которому протекает электрический ток, существует перепад температур, причем направление тока соответствует движению электронов от горячего конца к холодному, то, переходя из более горячего участка в более холодный, электроны передают избыточную энергию окружающим атомам, чем вызывают нагрев проводника (выделение тепла)*); при обратном направлении тока электроны, проходя из более холодного участка в более горячий, пополняют свою энергию за счёт окружающих атомов (поглощение тепла). Этим и объясняется в первом приближении явление Томсона. Для более точного описания явления следует учесть также, что в первом случае электроны тормозятся, а во втором — ускоряются полем термоэлектродвижущей силы, что изменяет значение коэффициента Томсона, а в некоторых случаях приводит даже к перемене его знака.

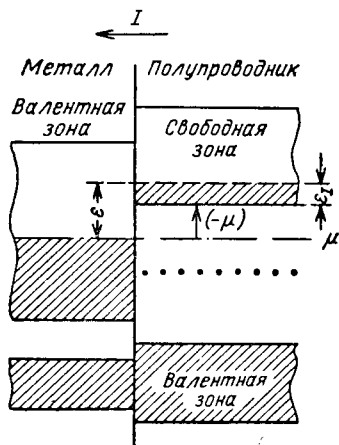


Рис. 1.24. Схема расположения энергетических уровней на контакте полупроводника с металлом (ϵ — энергия, выделяющаяся в виде тепла при переходе одного электрона из полупроводника в металл).

Причина возникновения явления Пельтье заключается в том, что средняя энергия электронов, участвующих в переносе тока в двух материалах, находящихся в контакте, различна, несмотря на то, что уровни их электрохимического потенциала совпадают (подробнее об этом см. гл. 6). Это особенно наглядно видно на контакте электронного полупроводника и металла (рис. 1.24). Предположим, что направление поля соответствует направлению перехода электронов из полупроводника в металл.

Если бы электроны, находящиеся на примесных уровнях, могли бы точно так же перемещаться под действием

*) Не следует смешивать эту теплоту с теплотой Джоуля; последняя есть следствие того, что электроны при столкновениях с атомами передают им энергию eEl , приобретенную в электрическом поле E на длине свободного пробега l .

электрического поля, как электроны в зоне проводимости, то в среднем энергия электронов, участвующих в электрическом токе, в полупроводнике равнялась бы энергии электронов на уровне Ферми в металле. При таких условиях переход электронов в металл не нарушил бы теплового равновесия в последнем; но электроны на примесных уровнях неподвижны, а энергия свободных электронов в зоне проводимости значительно выше энергии на уровне химического потенциала. Перейдя в металл, эти электроны опускаются до уровня Ферми, отдавая при столкновениях свою избыточную энергию атомам металла. Выделяющаяся при этом теплота и есть теплота Пельтье. Так как электроны приходят в тепловое равновесие в результате нескольких десятков соударений в непосредственной близости от контакта, то и вся теплота Пельтье выделяется почти на самом контакте *).

При противоположном направлении тока весь процесс идет в обратном направлении. В полупроводник из металла могут перейти электроны с энергетических уровней, расположенных выше дна зоны проводимости, т. е. значительно выше уровня Ферми. Тепловое равновесие в металле при этом нарушается и восстанавливается за счет энергии тепловых колебаний решетки. При этом поглощается теплота Пельтье.

ОСНОВНЫЕ ФОРМУЛЫ

Получив выражение для одного (любого) из термоэлектрических коэффициентов, остальные можно вывести, воспользовавшись соотношениями Томсона (1.102) и (1.104).

Рассмотрим для этого количественно энергетический баланс при переходе электрона из полупроводника в металл. Как видно из рис. 1.24, каждый электрон при таком переходе будет переносить избыточную энергию, состоящую из двух частей: $\varepsilon = \varepsilon_1 + \varepsilon_2$.

1. $\varepsilon_1 = \varepsilon_j$ — кинетической энергии в зоне проводимости. Значок j означает, что это энергия электрона, участвующего в электрическом токе. Следует сразу же подчеркнуть, что среднее значение этой энергии $\bar{\varepsilon}_j$ не равно средней тепловой энергии: $\bar{\varepsilon}_j \neq \frac{3}{2} kT$.

*) Более строгое рассмотрение показывает, что причиной всех трех термоэлектрических явлений является то, что средняя энергия носителей в потоке отличается от средней энергии в состоянии равновесия (см. гл. 6).

Объясняется это тем, что доля участия электронов с различной энергией в электрическом токе зависит от механизма рассеяния. Так, например, при рассеянии на ионах примеси (см. гл. 5) длина свободного пробега пропорциональна четвертой степени скорости. Это значит, что в данном случае в электрическом токе (или в любом другом, например, диффузионном потоке электронов) будут играть основную роль быстрые электроны и средняя энергия в потоке будет намного больше средней тепловой. В случае невырожденного электронного газа теория (см. гл. 6) дает следующее выражение для $\bar{\epsilon}_j$:

$$\bar{\epsilon}_j = (r + 2) kT, \quad (1.105)$$

где r — показатель степени в зависимости длины свободного пробега l электрона от энергии *).

2. Второе слагаемое (ϵ_2) в энергии, переносимой электроном из полупроводника в металл, как видно на рис. 1.24, измеряется расстоянием от дна зоны проводимости до уровня Ферми. Если за нуль отсчета энергии взять дно зоны проводимости полупроводника, то

$$\epsilon_2 = |\mu| = -\mu, \quad (1.106)$$

так как μ в данном случае лежит ниже дна зоны и, следовательно, отрицательно, а ϵ_2 — положительный вклад в выделяющуюся теплоту.

Таким образом, согласно (1.105) и (1.106)

$$\bar{\epsilon} = \bar{\epsilon}_j + (-\mu) = \bar{\epsilon}_j - \mu. \quad (1.107)$$

Опускаясь в металле до уровня Ферми, электрон передает избыточную энергию атома и металла, и она, таким образом, выделяется в виде тепла вблизи контакта.

Точно такое же количество тепла поглотится на контакте при противоположном направлении тока. Коэффициент Пельтье, представляющий собой отношение количества выделившегося тепла к перешедшему через контакт заряду,

*) Напомним читателю, что зависимость длины свободного пробега электронов l от энергии ϵ в общем случае имеет вид: $l \sim \epsilon^r$, где показатель степени r зависит от механизма рассеяния электронов. При рассеянии на ионах примеси $r = 2$, при рассеянии на оптических колебаниях решетки $r = 1/2$, при температуре ниже температуры Дебая и $r = 1$ при температуре выше температуры Дебая; при рассеянии на акустических колебаниях $r = 0$.

выразится, следовательно, формулой

$$\Pi_{\text{ПМ}} = \frac{\bar{\varepsilon}}{-e} = -\frac{1}{e} (\bar{\varepsilon}_j - \mu) \quad (1.108)$$

и согласно (1.102) коэффициент термо-э. д. с. полупроводника по отношению к металлу равен

$$\alpha_{\text{ПМ}} = \frac{\Pi}{T} = -\frac{k}{e} \left(\frac{\bar{\varepsilon}_j}{kT} - \mu^* \right), \quad (1.109)$$

где k — постоянная Больцмана.

Так как термо-э. д. с. металлов очень мала, то мы вправе в дальнейшем индексы при α опускать и всю термо-э. д. с. относить к полупроводнику.

Как будет показано в гл. 4, величина μ для невырожденного электронного полупроводника с простой параболической зоной определяется выражением

$$\mu = kT \ln \frac{h^3 n}{2 (2\pi m kT)^{3/2}}. \quad (1.110)$$

Таким образом, получаем окончательно

$$\alpha_n = -\frac{k}{e} \left(r + 2 + \ln \frac{2 (2\pi m_n kT)^{3/2}}{h^3 n} \right). \quad (1.111)$$

Формула (1.111) была впервые выведена в 1940 г. советским физиком Писаренко и носит его имя. Совершенно аналогичный, с точностью до знака, вид имеет выражение для термо-э. д. с. дырочного полупроводника

$$\alpha_p = \frac{k}{e} \left(r + 2 + \ln \frac{2 (2\pi m_p kT)^{3/2}}{h^3 p} \right). \quad (1.112)$$

Как уже упоминалось, в случае смешанной проводимости, когда электрический ток переносится электронами и дырками, термо-э. д. с., как правило, значительно ниже. В этих условиях уровень химического потенциала обычно находится приблизительно посредине запрещенной зоны и почти не смещается с изменением температуры (за исключением того сравнительно редкого случая, когда эффективные массы электронов и дырок сильно различаются)*; поэтому контактный член в термо-э. д. с. α близок к нулю. Что касается объемного члена, то он обычно тоже очень

*) Такая «биполярная» диффузия, как уже упоминалось, вносит весьма существенный вклад в теплопроводность (см. гл. 2).

мал; если подвижности и концентрации электронов и дырок одинаковы, то объемное поле вообще не возникает, так как оба рода носителей диффундируют от горячего участка к холодному в равном количестве.

Если же концентрация или подвижность носителей одного знака, например, электронов, больше, чем другого, то они диффундируют на холодный конец в большом количестве до тех пор, пока возникшее вследствие этого поле (тормозящее электроны и ускоряющее дырки) не уравнивает оба потока*).

Формула для термо-э. д. с. выглядит в этом случае следующим образом:

$$\alpha = \frac{\alpha_p u_p p - \alpha_n u_n n}{u_p p + u_n n}, \quad (1.113)$$

где u_p , u_n , p и n — соответственно подвижности и концентрации дырок и электронов;

α_p и α_n — «парциальные» термоэлектродвижущие силы, вычисленные согласно (1.111) и (1.112).

1.7. ГАЛЬВАНОМАГНИТНЫЕ И ТЕРМОМАГНИТНЫЕ ЯВЛЕНИЯ

ВВОДНЫЕ ЗАМЕЧАНИЯ

Как читателю известно из курса общей физики, на проводник длиной dl с током I , помещенный в магнитное поле H , действует так называемая пондермоторная сила, величина которой определяется законом Био—Савара:

$$dF = \frac{IH \, dl \sin(\widehat{H, I})}{c} = \frac{jH \sin(\widehat{H, j})}{c} dV, \quad (1.114)$$

где $j = I/s$ — плотность тока;

$dV = s \, dl$ — элемент объема;

в векторной форме

$$d\mathbf{F} = \frac{[jH]}{c} dV. \quad (1.115)$$

*) В этом случае (который реализуется в частности, в InSb) термо-э. д. с. и в области собственной проводимости может достигать больших значений.

Напомним также и закон индукции: в проводнике, перемещающемся в магнитном поле, возникает электродвижущая сила, пропорциональная скорости пересечения силовых линий магнитного поля. В простейшем случае, когда проводник расположен перпендикулярно полю, эта электродвижущая сила выражается формулой

$$dE = \frac{evH \, dl \sin(\widehat{\mathbf{H}, \mathbf{v}})}{c} \quad (1.116)$$

или в векторной форме

$$dE = \left| \frac{e [\mathbf{vH}] \, dl}{c} \right|. \quad (1.117)$$

Нетрудно показать, что пондерomotorные силы и явления индукции не что иное, как различные макроскопические проявления одного и того же микроскопического явления, сущность которого заключается в следующем.

Если в магнитном поле перемещается заряд e со скоростью v , то на него действует сила

$$f = \frac{e}{c} vH \sin(\widehat{\mathbf{H}, \mathbf{v}}) \quad (1.118)$$

или в векторной форме

$$\mathbf{f} = \frac{e [\mathbf{vH}]}{c}. \quad (1.119)$$

Убедимся, что (1.115) есть действительно следствие (1.118). Для этого выразим плотность тока через скорость электронов, и их число в единице объема; согласно выражению для плотности тока $\mathbf{j} = env$, следовательно, (1.115) можно переписать в виде

$$\mathbf{F} = e \left[\frac{\mathbf{vH}}{c} \right] n \, dV = \mathbf{f} \, dN, \quad (1.120)$$

где $dN = n \, dV$ — число электронов в элементе объема dV .

Точно так же можно показать, что закон индукции (1.117) есть следствие (1.119). В этом случае электроны вместе с проводником движутся в направлении его перемещения и сила \mathbf{F} , перпендикулярная \mathbf{H} и \mathbf{v} , заставляет перемещаться электроны вдоль проводника и создает электродвижущую силу и ток.

Если учесть еще силу eE , действующую на электрический заряд в электрическом поле, то общая сила, испытывае-

мая зарядом e в произвольном электромагнитном поле, выразится формулой

$$\mathbf{f} = e \left[\mathbf{E} + \frac{1}{c} [\mathbf{v}\mathbf{H}] \right]. \quad (1.121)$$

Соотношение (1.121) было выведено немецким ученым Лоренцом и f поэтому носит название силы Лоренца. Формула (1.121) лежит в основе расчета всех современных ускорителей (циклотрона, бетатрона, синхротрона), масс-спектрографов (т. е. приборов для разделения потока заряженных частиц по их массам), β -спектрометров (приборов для измерения энергий β -частиц) и др.

Сила, с которой действует магнитное поле на электрон [см. (1.119)], не имеет своего специального названия, и для краткости в дальнейшем будем называть ее магнитной силой.

После этих предварительных замечаний можно перейти непосредственно к основному содержанию главы.

Если проводник (или полупроводник), по которому протекает электрический ток, поместить в магнитное поле, то в нем возникнет ряд эффектов: уменьшится его электро- и теплопроводность; в направлении, перпендикулярном магнитному полю и току, возникнет разность потенциалов и температур. Каждое из этих явлений имеет свое название (большая часть связано с именем открывшего его ученого); и все они вместе называются гальваномагнитными явлениями.

Иными словами, гальваномагнитными явлениями называют явления, возникающие в проводниках первого рода*) при одновременном воздействии электрического и магнитного полей.

Аналогично этому термомагнитными явлениями называют эффекты, вызванные одновременным воздействием на проводник магнитного и температурного полей. Наиболее сильно гальвано- и термомагнитные эффекты проявляются тогда, когда электрическое или температурное поле направлено перпендикулярно магнитному. Поэтому только этот случай, имеющий наибольшее теоретическое и практическое значение, мы и будем рассматривать в дальнейшем.

*) Т. е. в материалах, в которых ток переносится электронами. В проводниках второго рода, в которых ток переносится ионами, гальваномагнитные явления не обнаруживаются.

Наиболее простым из гальваномагнитных эффектов является так называемый эффект Холла. В то же время это явление уже давно и широко применяется учеными для исследования электрических свойств твердого тела, и за последнее время нашло ряд важных применений в технике. Поэтому мы начнем с рассмотрения эффекта Холла и остановимся на нем наиболее подробно.

ЭФФЕКТ ХОЛЛА

Случай примесной проводимости. Рассмотрим образец примесного (для определенности — электронного) полупроводника, по которому протекает электрический ток, поме-

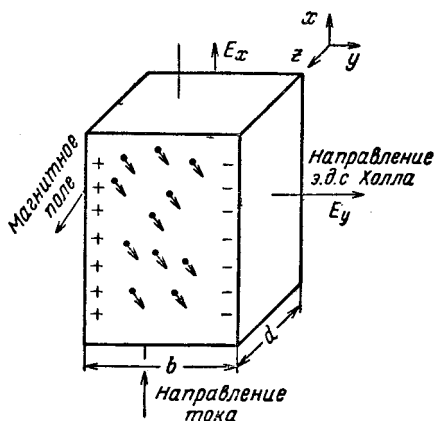


Рис. 1.25. Схема возникновения э. д. с. Холла.

щенный в магнитное поле, перпендикулярное направлению тока (рис. 1.25).

Под действием «магнитной силы» (1.119) электроны (направление их движения указано на чертеже стрелками) будут отклоняться на боковую грань образца; таким образом, на одной из граней будет скапливаться отрицательный заряд, а на противоположной будет оставаться нескомпенсированный положительный заряд. Возникшие вследствие этого поперечное электрическое поле и поперечная разность потенциалов и называются эффектом Холла.

Процесс накопления заряда будет продолжаться до тех пор, пока поперечное электрическое поле не уравновесит «магнитную силу»; после этого электроны будут двигаться вдоль образца, накопление заряда прекратится и устано-

вится стационарное состояние. Таким образом, поперечное холловское поле E_y может быть вычислено из условия равенства магнитной и электрической сил, действующих на электрон *):

$$\frac{ev_x H}{c} = eE_y, \quad (1.122)$$

откуда

$$E_y = \frac{v_x H}{c} \quad (1.123)$$

или, так как $v_x = uE_x$,

$$E_y = \frac{uH}{c} E_x. \quad (1.124)$$

При опыте измеряется не поперечное поле, а разность потенциалов (э. д. с. Холла, см. стр. 90)

$$V_x = E_y b, \quad (1.125)$$

и не скорость электронов v_x , а сила тока I , связанная с v_x соотношениями

$$j_x = env_x \text{ и } I = j_x S = j_x bd, \quad (1.126)$$

где bd — поперечное сечение образца (рис. 1.25).

Выразив E_y в (1.124) через V_x (1.125) и v_x через I , согласно (1.126) получим окончательно

$$V_x = \frac{1}{enc} \frac{HI}{d} = R_x \frac{HI}{d}, \quad (1.127)$$

где введено обозначение

$$R_x = \frac{1}{enc} \quad (1.128)$$

и d — размер образца в направлении магнитного поля. Величина R_x , являющаяся коэффициентом, связывающим поперечную разность потенциалов с напряженностью магнитного поля и силой тока, называется постоянной Холла.

Измерив четыре величины, входящие в (1.127), V_x , H , I и d , можно вычислить R_x :

$$R_x = \frac{V_x d}{HI}, \quad (1.129)$$

*) В действительности это неточно (см. гл. 7).

а зная R_x , мы можем определить из (1.128) концентрацию свободных электронов в исследуемом материале:

$$n = \frac{1}{eR_x c}. \quad (1.130)$$

Измеряя одновременно постоянную Холла и электропроводность, можно, кроме этого, вычислить и подвижность электронов; действительно, в случае примесного полупроводника

$$\sigma = en\mu$$

и согласно (1.128)

$$R_x \sigma = \frac{\mu}{c}. \quad (1.131)$$

Таким образом, одновременное измерение электропроводности и эффекта Холла позволяет получить важные сведения о примесном полупроводнике: знак носителей тока *) (по знаку э. д. с. Холла), их число и подвижность. В этом и заключается значение измерения эффекта Холла как одного из основных методов исследования электрических свойств проводников и полупроводников.

В заключение этого раздела мы должны сделать несколько замечаний.

Математический вывод э. д. с. Холла и постоянной Холла, приведенный выше, весьма нестрог; действительно, в выражение для магнитной силы входит полная скорость электрона, состоящая в полупроводниках из двух частей: скорости беспорядочного теплового движения v_0 и направленной добавки за счет электрического поля v_d ($v = v_0 + v_d$); в выражение же для плотности тока входит только скорость дрейфа в электрическом поле. отождествив в нашем выводе v_d и v , мы тем самым пренебрегли хаотическим тепловым движением электронов и их распределением по скоростям, поэтому значение постоянной Холла (1.128), полученное нами, не точно. Более строгий вывод (с учетом теплового движения (см. гл. 7) дает несколько другое значение для постоянной Холла:

$$R_x = \frac{A}{enc}, \quad (1.132)$$

*) Знак носителей тока, впрочем, проще определить по закону термо-э. д. с.

где постоянная A зависит от механизма рассеяния электронов, т. е. от показателя степени в выражении для зависимости длины свободного пробега электрона от его энергии.

Перечислим наиболее часто встречающиеся на практике случаи:

— в ковалентных кристаллах при рассеянии электронов на акустических колебаниях длина свободного пробега электрона не зависит от энергии ($r = 0$) и

$$A = \frac{3\pi}{8} \text{ и } R_{\text{X}} = \frac{3\pi}{8} \frac{1}{\text{enc}} \approx 1,17 \frac{1}{\text{enc}}; \quad (1.133)$$

— в полупроводниках с ионной решеткой и при рассеянии электронов на оптических колебаниях при температуре выше температуры Дебая длина свободного пробега пропорциональна энергии ($r = 1$) и

$$A \approx 1,11 \text{ или } R_{\text{X}} \approx 1,11 \frac{1}{\text{enc}}; \quad (1.134)$$

— при рассеянии на ионах примеси длина свободного пробега электрона пропорциональна квадрату энергии ($r = 2$) и

$$R_{\text{X}} \approx 1,93 \frac{1}{\text{enc}}, \quad (1.135)$$

— в металлах и сильно вырожденных полупроводниках в электропроводности могут принимать участие только электроны, находящиеся на самых верхних уровнях распределения Ферми; энергия и скорость у этих электронов вполне определены, и в данном случае никакого распределения по скоростям учитывать не нужно; в соответствии с этим вывод, приведенный нами, оказывается строгим для металлов и вырожденных полупроводников $A = 1$ и постоянная Холла в этом случае равна

$$R_{\text{X}} = \frac{1}{\text{enc}}. \quad (1.136) *$$

Случай смешанной проводимости. Если в переносе электричества участвуют и дырки и электроны, то картина эффекта Холла значительно усложняется. В выражение

*) Точно так же $A = 1$ при рассеянии электронов на оптических колебаниях при температуре ниже температуры Дебая, при этом $r = 1/2$ и время свободного пробега для всех электронов одинаково; это приводит к тому, что разброс тепловых скоростей электронов не играет существенной роли (см. гл. 7).

для магнитной силы входит произведение скорости частицы на заряд; так как для электронов и дырок, участвующих в переносе тока, и знак (направление) скорости и знак заряда противоположны, то направления сил, действующих на них, будут одни и те же и они будут отклоняться в одну и ту же сторону.

Если и подвижности и концентрации электронов и дырок одинаковы, то их заряды будут полностью компенсировать друг друга и холловское поле будет равно нулю *). Если же это равенство не имеет места и концентрация или подвижность носителей одного знака (для определенности предположим, что электронов) больше, чем другого, то на одной грани будет скапливаться отрицательный заряд, а на другой будет оставаться нескомпенсированный положительный заряд. Таким образом, возникает поперечное холловское поле, тормозящее движение электронов в этом направлении и ускоряющее движение дырок. Процесс накопления заряда будет продолжаться до тех пор, пока это поле не уравнивает потоки электронов и дырок. Расчет дает в этом случае следующее выражение для постоянной Холла:

$$R_x = \frac{A}{ec} \frac{u_p^2 p - u_n^2 n}{(u_p p + u_n n)^2}, \quad (1.137)$$

где u_p , p , u_n , n — соответственно подвижности и концентрации дырок и электронов.

Постоянная A в (1.137), так же как и в случае одного знака носителей, определяется механизмом рассеяния: $A = 3\pi/8$ в атомной решетке и т. д.

Нетрудно убедиться, что: а) если $n = 0$ или $p = 0$, (1.137) переходит в (1.132), б) если $n = p$ и $u_p = u_n$, постоянная Холла, а следовательно, и э. д. с. Холла равны нулю.

Согласно (1.137) в области собственной проводимости знак э. д. с. Холла соответствует знаку носителей, подвижность которых больше (обычно электронов); поэтому в примесном дырочном проводнике э. д. с. Холла при переходе

*) В данном случае влияние магнитного поля проявится лишь в том, что на одной боковой грани образца концентрация носителей, а вследствие этого и локальная электропроводность повысятся, а на другой понизятся. Впервые на существование этого эффекта указал А. Ф. Иоффе, и затем этот эффект был обнаружен на опыте рядом советских и, независимо, иностранных ученых.

к собственной проводимости обычно проходит через нуль и меняет знак.

В случае смешанной проводимости выражение для электропроводности имеет вид

$$\sigma = e (u_p p + u_n n). \quad (1.138)$$

Как видно из (1.137) и (1.138), в этом случае измерение постоянной Холла и электропроводности не дает достаточных данных для определения подвижности и концентрации электронов и дырок, так как мы имеем два уравнения с четырьмя неизвестными. Можно выйти из положения, найдя значения подвижности дырок или электронов экстраполяцией из области примесной проводимости. Задача может быть также решена, если к выражению (1.137) и (1.138) добавить еще независимые уравнения, т. е. одновременно произвести измерения термо-э. д. с. и еще каких-либо эффектов, представляющих дополнительные независимые данные. Вопрос несколько упрощается в области собственной проводимости, где $n = p$.

Перейдем к анализу следующего гальваномагнитного явления — так называемого эффекта Эттингсгаузена.

ЭФФЕКТ ЭТТИНГСГАУЗЕНА

Мы уже упоминали выше, что основная погрешность при выводе холловской разности потенциалов заключалась в том, что не было учтено распределение электронов по скоростям. Если же мы учтем эту ошибку, то сразу же станет ясно, что равенство (1.122)

$$\frac{ev_x H}{c} = eE_y \quad (1.139)$$

не может удовлетворяться для всех электронов сразу. Действительно, в (1.122) справа стоят вполне определенные величины — поперечное холловское поле и заряд электрона, слева же в качестве одного из сомножителей стоит скорость электрона, которая, если учесть ее тепловую составляющую, колеблется в весьма широких пределах. Поэтому равенство (1.122) может удовлетворяться лишь для электронов с некоторой средней скоростью; для электронов же, скорость которых больше средней, магнитная сила $e [v_x H] / c$ будет превышать электрическую eE_y , и они будут

отклоняться на одну грань образца; электроны же, скорость которых меньше средней, отклоняются под действием электрической силы в противоположную сторону. При этом быстрые электроны, скапливаясь на одной грани образца, будут передавать свою избыточную энергию кристаллической решетке, и эта грань образца будет нагреваться; медленные электроны на противоположной грани будут пополнять свою энергию за счет решетки и эта грань будет охлаждаться. Таким образом, наряду с поперечной разностью потенциалов — эффектом Холла — возникнет поперечная разность температур — так называемый эффект Эттингсгаузена.

По мере нагревания одной грани и охлаждения другой появится и будет расти поток тепла от нагретой грани к холодной. При некоторой разности температур установится стационарное состояние, когда отток тепла от горячей грани за счет теплопроводности сбалансирует приток тепла за счет быстрых электронов; эта разность температур и будет «равновесным» эффектом Эттингсгаузена.

Мы приведем здесь выражение для поперечного градиента температуры в электронном невырожденном полупроводнике без вывода:

$$\frac{\partial T}{\partial y} = \frac{kT}{e} \frac{2r-1}{2} \frac{R_{\chi}\sigma}{\chi_p + \chi_{эл}} j_x H, \quad (1.140)$$

где χ_p и $\chi_{эл}$ — соответственно теплопроводность решетки и электронного газа.

В случае смешанной проводимости выражение для поперечной разности температур сложнее, и мы его здесь приводить не будем.

Эффект Холла и эффект Эттингсгаузена называются поперечными гальваномагнитными эффектами — разность потенциалов и разность температур появляются в направлении, перпендикулярном направлению электрического тока.

Теперь перейдем к рассмотрению так называемых продольных гальваномагнитных эффектов.

ИЗМЕНЕНИЕ СОПРОТИВЛЕНИЯ В МАГНИТНОМ ПОЛЕ

Выясним, по какой траектории будет двигаться электрон, попав в однородное магнитное поле (H_z), перпендикулярное направлению его начальной скорости. Согласно (1.119) на

электрон будет действовать сила $f = evH_z/c$, и траектория его будет искривляться; однако, так как ускорение $a = f/m$ в каждый данный момент [согласно (1.119)] будет перпендикулярно направлению движения электрона, скорость по абсолютной величине не будет изменяться: $v = \text{const}$. Радиус кривизны траектории мы можем найти из условия равенства центробежной и центростремительной сил:

$$\frac{mv^2}{r} = \frac{evH}{c}, \quad (1.141)$$

откуда

$$r = \frac{mvc}{eH}. \quad (1.142)$$

Так как все входящие в (1.142) величины постоянны, радиус траектории будет тоже постоянен; иными словами, электрон будет двигаться по окружности. Период обращения электрона согласно (1.142)

$$T = \frac{2\pi r}{v} = \frac{2\pi mc}{eH} \quad (1.143)$$

и угловая скорость

$$\omega = \frac{2\pi}{T} = \frac{eH}{mc}. \quad (1.144)$$

В теории гальвано- и термомагнитных явлений различают случай сильных и слабых магнитных полей. Слабыми называют такие магнитные поля, в которых радиус кривизны r траектории электрона много больше его средней длины свободного пробега \bar{l} :

$$r \gg \bar{l}. \quad (1.145)$$

При этих условиях действие магнитного поля на электрон за время его свободного пробега будет относительно слабым: оно выразится в том, что траектория электрона незначительно искривится и направление его движения отклонится от первоначального на небольшой угол $\bar{\varphi}$ (рис. 1.26, а):

$$\bar{\varphi} = \frac{\bar{l}}{r} \ll 1. \quad (1.146)$$

Но подвижность электрона u выражается через его длину свободного пробега формулой

$$u = \frac{e}{m} \frac{\bar{l}}{v}. \quad (1.147)$$

Подставляя в (1.146) значение r из (1.142), получим

$$\bar{\varphi} = \frac{e}{m} \frac{\bar{l}}{v} \frac{H}{c} \ll 1 \quad (1.148)$$

или

$$\bar{\varphi} = \frac{uH}{c} \ll 1. \quad (1.149)$$

Следовательно, неравенство (1.146), положенное нами в основу определения слабого поля, может быть заменено эквивалентным:

$$\frac{uH}{c} \ll 1. \quad (1.150)$$

Последнее определение слабого поля (1.150) является наиболее удобным, так как в него входят величины u и H , непосредственно измеряемые на опыте.

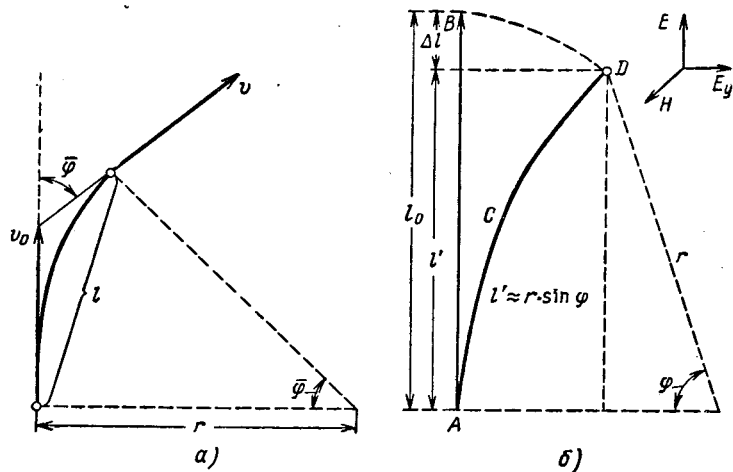


Рис. 1.26. Искривление траектории электрона в слабом магнитном поле (а) и схема, поясняющая уменьшение эффективной длины свободного пробега электрона в слабом магнитном поле (б).

Сильными магнитными полями мы называем такие поля, в которых выполняется неравенство, противоположное (1.145) или (1.150):

$$r \ll l \quad (1.151)$$

$$\frac{uH}{c} \gg 1. \quad (1.152)$$

В этом случае характер движения электрона существенно изменяется: в промежутке между столкновениями он уже не движется по почти прямолинейной траектории (как в случае слабых магнитных полей), а проходит ряд циклов либо винтовой линии, либо циклоиды, либо еще более сложной траектории.

Из сказанного выше читателю должно быть ясно, что деление на слабые и сильные поля не является абсолютным: при одном и том же значении поля для одного материала (с малой подвижностью) может выполняться неравенство (1.145), а для другого (с большой подвижностью) может удовлетворяться противоположное неравенство (1.151). Более того, для одного и того же материала при высоких температурах данное поле может быть слабым, а при низких температурах — сильным. Так, например, для чистого германия подвижность электронов составляет при комнатной температуре $u = 3000 \text{ см}^2/\text{в} \cdot \text{сек} = 9 \cdot 10^5 \text{ CGSE}$. Поле напряженностью 10^4 э будет приблизительно удовлетворять неравенству (1.150):

$$\frac{9 \cdot 10^5 \cdot 10^4}{3 \cdot 10^{10}} = 0,3 < 1.$$

При температуре 10° К подвижность для того же германия будет 10^4 — $10^5 \text{ см}^2/\text{в} \cdot \text{сек}$ и то же самое магнитное поле 10^4 э мы должны будем считать сильным *). Ограничимся пока рассмотрением слабых полей, случай сильных магнитных полей будет рассмотрен в гл. 7.

Рассмотрим изменение сопротивления в магнитном поле.

Собственная проводимость. В области собственной проводимости, где число дырок и электронов можно считать одинаковым, $n = p$, холловское поле проявляется как разностный эффект вследствие их различных подвижностей; если же $u_p = u_n$, то $E_x = 0$. В этом случае действие магнитной силы ничем не компенсируется и электроны и дырки отклоняются на своем пути от направления продольного электрического поля на угол φ (рис. 1.26):

$$\varphi = \frac{uH}{c}. \quad (1.153)$$

*) В ряде случаев одно и то же поле может быть слабым для одной группы носителей и сильным для другой.

Если в отсутствие магнитного поля электрон проходил без столкновений расстояние $AB = l_0$ в направлении электрического поля E , то теперь он пройдет такое же расстояние по дуге ACD . Следовательно, эффективная длина свободного пробега электрона в направлении электрического тока уменьшится: это будет l' — проекция ACD на AB .

Но подвижность электрона пропорциональна пути, пройденному на длине свободного пробега в направлении поля, а электропроводность пропорциональна подвижности, следовательно, соответствующим образом уменьшается и подвижность и электропроводность:

$$\frac{\Delta u}{u} = \frac{\Delta \sigma}{\sigma} = \frac{\Delta l}{l_0}. \quad (1.154)$$

Из рис. 1.26, б не представляет труда вычислить Δl :

$$\Delta l = l_0 - l' \approx l_0 - r \sin \varphi;$$

но

$$r = \frac{l_0}{\varphi}; \quad \sin \varphi \approx \varphi - \frac{\varphi^3}{3}, \quad (1.155)$$

следовательно,

$$\frac{\Delta l}{l_0} \approx \frac{\varphi^2}{3}$$

и согласно (1.154)

$$\frac{\Delta \sigma}{\sigma} \approx -\frac{1}{3} \left(\frac{uH}{c} \right)^2. \quad (1.156)$$

В приведенном выше выводе мы допустили те же погрешности, что и при выводе э. д. с. Холла, т. е. не учли теплового разброса скоростей электрона. Поэтому и формула (1.156), точно так же, как формула (1.127) для э. д. с. Холла, верна с точностью до коэффициента. Строгий вывод с учетом статистического разброса скоростей дает выражение, аналогичное (1.156):

$$\frac{\Delta \sigma}{\sigma} = -B \left(\frac{uH}{c} \right)^2, \quad (1.157)$$

где коэффициент B зависит от механизма рассеяния электронов.

В атомной решетке ($r = 0$) $B = 9\pi/16$; в ионной решетке при температуре ниже температуры Дебая $B = 1$ и при температуре выше температуры Дебая $B = 27\pi/64$.

Выше мы рассмотрели случай собственной проводимости, т. е. считали, что концентрации электронов и дырок одинаковы, и кроме того предположили, что и подвижности их равны. Эти оба условия (особенно второе) сравнительно редко реализуются на практике. Если же $n \neq p$ и $u_n \neq u_p$, выражения для $\Delta\sigma/\sigma$ приобретают весьма сложный вид.

Примесная проводимость. В случае примесной проводимости в образце возникает холловская разность потенциалов и «электрическая» сила eE_x противодействует «магнитной» evH/c . Как уже упоминалось выше, в этом случае для электронов с некоторой средней скоростью соблюдается равенство (1.122): $ev_xH/c = eE_y$.

Такие электроны будут двигаться вдоль образца и для них при включении магнитного поля эффективная длина свободного пробега не изменится. Но электроны со скоростью меньше средней будут отклоняться в сторону «электрической» силы, электроны со скоростью больше средней будут отклоняться в противоположную сторону, и для тех и для других эффективная длина свободного пробега уменьшится.

Следовательно, и в этом случае электропроводность образца уменьшится, но эффект будет меньше, чем в области собственной проводимости, так как холловское поле в известной степени «выправляет» траектории электронов.

Для примесных проводников теория дает выражение для изменения электропроводности, совершенно аналогичное (1.157):

$$\frac{\Delta\sigma}{\sigma} = B_1 \left(\frac{uH}{c} \right)^2, \quad (1.158)$$

где коэффициент B_1 , так же как и B в (1.157), зависит от механизма рассеяния: в атомной решетке $B_1 = \pi/10$, в ионной решетке при температуре ниже температуры Дебая $B_1 = 0$, при температуре выше температуры Дебая $B_1 = 0,96$, при рассеянии на ионах примеси $B_1 = 1$. Как и следует из качественных представлений, развитых выше, во всех случаях $B_1 < B$.

Как видно из (1.157) и (1.158), измерение изменения сопротивления в магнитном поле позволяет непосредственно определить подвижность носителей тока в исследуемом материале, если известен механизм рассеяния, поэтому измере-

ние этого эффекта стало за последнее время довольно распространенным *).

В заключение этого раздела остановимся на втором продольном гальваномагнитном эффекте — так называемом эффекте Нернста. Как уже упоминалось выше, траектории медленных электронов сильнее «закручиваются» в магнитном поле, чем траектории быстрых. В силу этого при включении магнитного поля доля участия в электрическом токе электронов с различной энергией изменится. Электроны с большой энергией будут почти беспрепятственно проходить через образец и, накапливаясь на том торце образца, к которому они движутся, вызовут его нагревание; противоположный торец, на котором «застраиваются» медленные электроны, охладится. Таким образом, возникает продольная разность температур, называемая эффектом Нернста.

ТЕРМОМАГНИТНЫЕ ЯВЛЕНИЯ

Рассмотрим полупроводниковый образец, на торцах которого поддерживаются различные температуры. Обра-

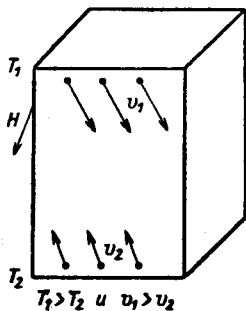


Рис. 1.27. Схема возникновения поперечных термомагнитных явлений.

зец помещен в магнитное поле, перпендикулярное направлению теплового потока. Как показано на рис. 1.27, «горячие» электроны, т. е. движущиеся от горячего конца к холодному, будут отклоняться вправо, а «холодные», т. е. движущиеся от холодного конца, — влево и, следовательно, правая грань будет нагреваться, а левая охлаждаться. Это явление носит название эффекта Риги — Ледюка.

*) Кроме этого, и эффект Холла и изменение сопротивления в магнитном поле широко используется в технике (см. [29]).

Перейдем теперь ко второму термомагнитному эффекту. Тепловая скорость электронов, движущихся от горячего конца, больше, чем скорость холодных электронов; поэтому радиус кривизны их траектории будет больше, т. е. они будут слабее отклоняться магнитным полем. Следовательно, потоки электронов на боковые грани не будут одинаковы: левая грань будет заряжаться отрицательно по отношению к правой; таким образом, возникнет поперечная разность потенциалов — эффект Нернста — Эттингсгаузена. Согласно теории, в стационарных условиях поперечное поле выражается следующей формулой (при наличии одного сорта носителей):

$$E_y = \frac{2r-1}{2} \frac{A}{e^2} \mu H \frac{\partial T}{\partial x}, \quad (1.159)$$

где A — коэффициент, входящий в выражение (1.132) для постоянной Холла.

Как следует из (1.159), исследование этого эффекта при наличии носителей одного знака позволяет непосредственно определить их подвижность. Из теории также следует, что э. д. с. Нернста — Эттингсгаузена чрезвычайно чувствительна к появлению небольшой добавки носителей другого знака. Кроме упомянутых выше двух поперечных термомагнитных эффектов существуют два продольных:

ТАБЛИЦА 2

Таблица гальвано- и термомагнитных эффектов

| Эффекты | Поперечные | Продольные |
|-------------------|-------------------------------------------------------------------|------------------------------------------------------------------------------|
| Гальваномагнитные | Эффект Холла (поперечная разность потенциалов) | Изменение сопротивлений в магнитном поле |
| Термомагнитные | Эффект Эттингсгаузена (поперечная разность температур) | Эффект Нернста (продольная разность температур) |
| | Эффект Риги — Ледюка (поперечная разность температур) | Эффект Маджи — Риги — Ледюка (изменение теплопроводности в магнитном поле) |
| | Эффект Нернста — Эттингсгаузена (поперечная разность потенциалов) | Продольный эффект Нернста — Эттингсгаузена (продольная разность потенциалов) |

— эффект Маджи — Риги — Ледюка, заключающийся в изменении (уменьшении) электронной теплопроводности, есть следствие того, что эффективная длина свободного пробега электронов вдоль теплового потока уменьшается вследствие закручивания электронных траекторий;

— эффект Нернста — Эттингсгаузена (изменение термо-э. д. с. в магнитном поле) возникает также в результате закручивания траекторий электронов; по этой причине изменяется средняя энергия электронов в потоке, а следовательно, кинетический член в термо-э. д. с. (ϵ_j).

Выше мы приводим таблицу всех гальвано- и термомагнитных явлений, которая поможет читателю запомнить их и разобраться в их классификации.

1.8. ФОТОПРОВОДИМОСТЬ

В явлении фотопроводимости, т. е. изменении сопротивления вещества под действием света, различают нормальный фотоэффект — увеличение проводимости, и аномальный — обратное явление. Аномальный фотоэффект — явление редкое, поэтому мы на нем останавливаться не будем.

В принципе увеличение электропроводности под действием освещения могло бы объясняться и увеличением подвижности носителей, и увеличением их концентрации. Однако прямыми опытами было показано, что во всех случаях мы имеем дело с увеличением концентрации, а подвижность световых и темновых носителей остается одной и той же *).

Существуют три пути для увеличения концентрации носителей под действием света:

1) кванты света вырывают электроны из заполненной зоны и забрасывают их в зону проводимости (рис. 1.28—1);

*) Время, которое проводит носитель, освобожденный светом, в зоне проводимости, колеблется в пределах 10^2 — 10^{-9} сек, а время между столкновениями в зоне составляет 10^{-10} — 10^{-13} сек. Следовательно, за время своего пребывания в зоне фотоноситель испытывает миллионы столкновений.

Непосредственно после заброса энергия фотоносителя определяется энергией светового кванта. Однако теория и опыт показывают, что нескольких десятков столкновений вполне достаточно, чтобы электрон приобрел равновесную тепловую энергию. Поэтому подавляющую часть времени жизни фотоноситель движется с тепловой скоростью и его подвижность равна подвижности тепловых носителей.

при этом одновременно возрастает число дырок и электронов;

2) электроны вырываются из заполненной зоны и забрасываются на свободные примесные уровни; при этом возрастает дырочная проводимость (рис. 1.28—2);

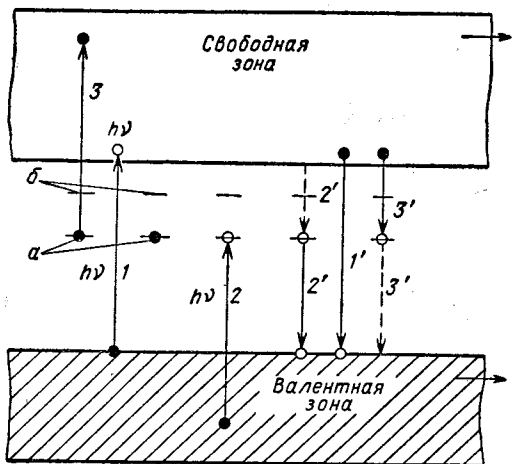


Рис. 1.28. Схема возможных путей генерации и рекомбинации носителей:

1' — прямая (межзонная) рекомбинация; 2' и 3' — рекомбинация через примесные центры; 2' — электрон с примесного центра «падает» в валентную зону, а на освободившееся место «падает» электрон из зоны проводимости; 3' — то же происходит в обратном порядке; а — энергетические уровни, соответствующие нормальным состояниям электрона на примесном центре; б — энергетические уровни, соответствующие возбужденным состояниям электрона на примесном центре.

3) электроны забрасываются с примесных уровней в свободную зону и увеличивается электронная проводимость (рис. 1.28—3).

Рассмотрим вначале качественно явление фотопроводимости на примере третьего случая.

ФЕНОМЕНОЛОГИЧЕСКОЕ ОПИСАНИЕ ФОТОПРОВОДИМОСТИ

Пусть интенсивность света такова, что в единицу времени с донорных атомов срывается под его действием g электронов. Тогда с момента начала освещения концентрация носи-

телей начнет расти по закону

$$\frac{d\Delta n}{dt} = g, \quad (1.160)$$

где

$$\Delta n = n - n_0$$

и n_0 — равновесная концентрация носителей.

Очевидно, что процесс роста концентрации не может продолжаться бесконечно. Действительно, по мере увеличения числа свободных носителей и числа свободных мест на примесных уровнях будет расти вероятность рекомбинации и через некоторое время второй процесс уравнивает первый. После этого полное число тепловых (g_0) и световых (g) забросов в 1 сек будет равно полному числу актов рекомбинаций и концентрация носителей будет стационарной ($\Delta n_{ст}$).

Обозначим время жизни фотоносителя, т. е. время, которое в среднем проводит носитель, созданный светом, в зоне проводимости, τ . Тогда число фотоэлектронов, рекомбинирующих в 1 сек *, будет $\Delta n/\tau$, и, следовательно, уравнение (1.160) с учетом рекомбинации должно быть переписано в виде

$$\frac{d\Delta n}{dt} = g - \frac{\Delta n}{\tau}; \quad (1.161)$$

решение дифференциального уравнения (1.161) имеет вид

$$\Delta n = \Delta n_{ст} (1 - e^{-\frac{t}{\tau}}), \quad \text{где } \Delta n_{ст} = g\tau, \quad (1.162)$$

или

$$n = n_0 + \Delta n_{ст} (1 - e^{-\frac{t}{\tau}}). \quad (1.163)**$$

Как видно из (1.162), Δn асимптотически приближается к своему стационарному значению $\Delta n_{ст}$. Поэтому величину τ в этом случае называют также временем релаксации для фотопроводимости.

После прекращения освещения генерация носителей прекратится и уравнение (1.161) должно быть переписано

*) Т. е. возвращающихся обратно на пустые донорные уровни.

***) Проведенный выше анализ справедлив лишь в том случае, когда τ не зависит от n (подробнее об этом см. в гл. 9).

В виде

$$\frac{d\Delta n}{dt} = -\frac{\Delta n}{\tau}; \quad (1.164)$$

решение уравнения (1.164) с начальным условием при $t=0$ и $\Delta n = \Delta n_{ст}$ имеет вид

$$\Delta n = \Delta n_{ст} e^{-\frac{t}{\tau}}$$

или

$$n = n_0 + \Delta n_{ст} e^{-\frac{t}{\tau}}, \quad (1.165)$$

т. е. после прекращения освещения концентрация носителей в зоне проводимости будет падать, постепенно возвращаясь к своему исходному (равновесному) значению.

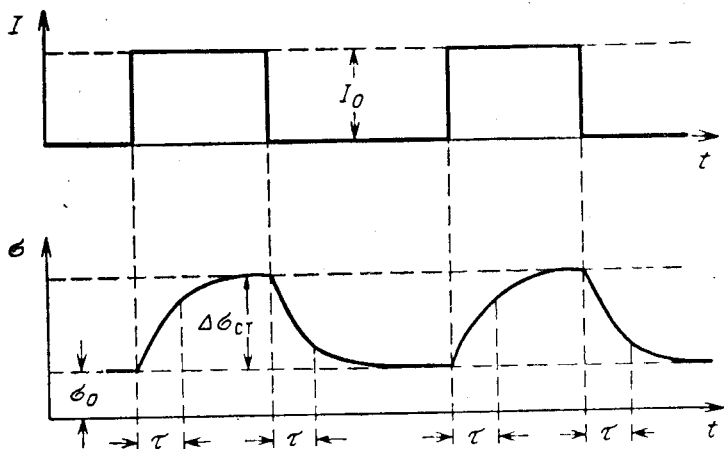


Рис. 1.29. Изменение фотопроводимости при освещении прямоугольными импульсами.

Таким образом, полная кривая изменения проводимости при освещении полупроводника прямоугольными световыми импульсами будет иметь вид, представленный на рис. 1.29, согласно (1.163) и (1.165).

Стационарная фотопроводимость выражается формулой

$$\Delta \sigma_{ст} = e \Delta n_{ст} \mu = e u g \tau. \quad (1.166)$$

Рекомбинация носителей может происходить двумя различными путями (см. рис. 1.28): прямая (межзонная) рекомбинация (I') и рекомбинация через примесные центры ($2'$ и $3'$). В случае $2'$ вначале электрон с примесного центра падает в заполненную зону (этот процесс называется захватом дырки), а затем на освободившийся на примесном центре уровень падает электрон из зоны проводимости; в случае $3'$ сначала захватывается электрон, а затем — дырка.

При рекомбинации носителей, так же как и при генерации, должны соблюдаться законы сохранения энергии и импульса. При тепловой генерации электроны получают энергию и импульс от фонона или другого свободного электрона, имеющего достаточную кинетическую энергию (этот механизм генерации называется ударным), при световой — от фотона. В ряде случаев существенную роль играют комбинированные процессы, в которых электрон получает энергию в основном от фотона, а импульс — в основном от фонона (см. гл. 9). При рекомбинации носителей есть также ряд возможностей перераспределения энергии.

Излучательная рекомбинация — когда энергия и импульс, освободившиеся при рекомбинации электрона и дырки, излучаются в виде света. Однако и расчет и опыт показывают, что вероятность такого процесса относительно велика в сравнительно редких случаях — в достаточно чистых кристаллах и даже при этих условиях далеко не во всех кристаллах. Излучательная рекомбинация проявляется в люминесценции кристаллов; с ней также неразрывно связано действие лазеров.

Безызлучательная рекомбинация — процесс, аналогичный рассмотренному выше (электрон непосредственно «падает» в заполненную зону), с той лишь разницей, что освободившаяся при этом энергия выделяется в виде тепла (фононов). Однако вероятность и этого процесса также обычно мала. Дело в том, что энергия одного фонона не превышает нескольких сотых электрон-вольта. Если бы при ширине запрещенной зоны порядка электрон-вольта вся энергия выделилась бы в виде тепла, то это означало бы, что одновременно «родилось» несколько десятков фононов. Такой процесс почти так же маловероятен, как столкновение в одной точке нескольких десятков частиц. Теории безызлучательных переходов посвящена монография [39].

Прямые переходы начинают играть заметную роль в веществах с узкой запрещенной зоной.

В большинстве случаев решающую роль играет уже упоминавшаяся выше рекомбинация через центры захвата. Это означает, что освобожденный светом (или теплом) электрон сначала захватывается примесным центром, а затем уже падает в заполненную зону. В принятой терминологии это явление описывается следующим образом: сначала примесным центром захватывается электрон, а затем — дырка. Разумеется, эти процессы могут происходить и в обратном порядке.

По сравнению с непосредственной рекомбинацией рекомбинация через примесные центры обычно более вероятна по ряду причин:

1. В этом случае энергия рекомбинации разбивается на две части, и в каждом элементарном акте она меньше, чем в прямых переходах, следовательно, при этом одновременно образуется меньшее число фононов. Поэтому для рекомбинации наиболее эффективны центры, уровни которых расположены вблизи середины запрещенной зоны.

2. Примесный центр в первый момент может присвоить себе частично или полностью энергию или импульс электрона (или дырки).

Две стадии этого процесса — захват электрона и дырки — могут иметь совершенно различную вероятность. Если примесный центр расположен ближе к краю свободной зоны, то в соответствии со сказанным выше захват электрона будет во много раз более вероятен, чем захват дырки, и он (электрон) будет очень долго «ожидать» своего партнера и может даже случиться, что при очередной тепловой флюктуации будет заброшен обратно в зону и лишь после многократного повторения этих процессов, наконец, свалится в заполненную зону.

В тех случаях, когда обратный заброс более вероятен, чем рекомбинация, соответствующие уровни называются уровнями прилипания для электрона (или дырки). Вероятность захвата носителя зависит также от заряда примесного атома. Для нейтральной примеси при прочих равных условиях эти вероятности одинаковы, для положительного иона более вероятен захват электрона, для отрицательного — дырки.

Так как зарядовое состояние и степень заполнения примесных центров зависит от положения уровня Ферми,

а следовательно, и от температуры, то эти факторы также существенно влияют на время жизни дырок и электронов.

При больших интенсивностях света степень заполнения различных уровней, а следовательно, и вероятность рекомбинации носителей начинают зависеть от числа поглощенных квантов, т. е. от интенсивности и времени.

В глубоко легированных полупроводниках, а также при высоком уровне освещенности решающую роль начинает играть так называемая ударная рекомбинация. При ударной рекомбинации освободившаяся энергия передается второму (свободному) электрону.

КВАНТОВЫЙ ВЫХОД

Квантовым выходом называется отношение числа фотоэлектронов g к общему числу поглощенных квантов N :

$$n = \frac{g}{N}. \quad (1.167)$$

Выше было рассмотрено несколько путей создания (генерации) фотоносителей. Все механизмы поглощения света, приводящие к появлению свободных электронов, называются фотоактивными и если бы других механизмов поглощения света не было, то квантовый выход всегда был бы равен единице *).

Однако существует целый ряд возможностей поглощения света, не приводящих (по крайней мере непосредственно) к образованию фотоносителей: поглощение света, связанное с возбуждением экситонов (см. ниже), примесных атомов, ионных колебаний, и поглощение свободными электронами. Наличие этих процессов приводит к тому, что в большинстве случаев квантовый выход много меньше единицы. Есть, впрочем, случаи, когда каждый квант, обладая большой энергией, освобождает не один, а несколько электронов; при этих условиях квантовый выход может быть и больше единицы. Это имеет место при облучении полупроводникового материала ультрафиолетовым светом или еще более жесткими излучениями, т. е. тогда, когда энергия кванта достаточно велика, чтобы не только вырвать электрон из заполненной зоны, но и придать ему такую кинетическую энергию, чтобы при последующих столкновениях вызвать еще один или несколько актов ионизаций.

*) Или больше единицы в тех случаях, когда энергия кванта достаточна, чтобы создать несколько фотоэлектронов.

Спектральное распределение фоточувствительности теснейшим образом связано со спектром поглощения полупроводника, который схематически представлен на рис. 1.30, а (подробнее о спектрах поглощения см. гл. 9). Как видно из рисунка, весь спектр можно разбить на две части: так

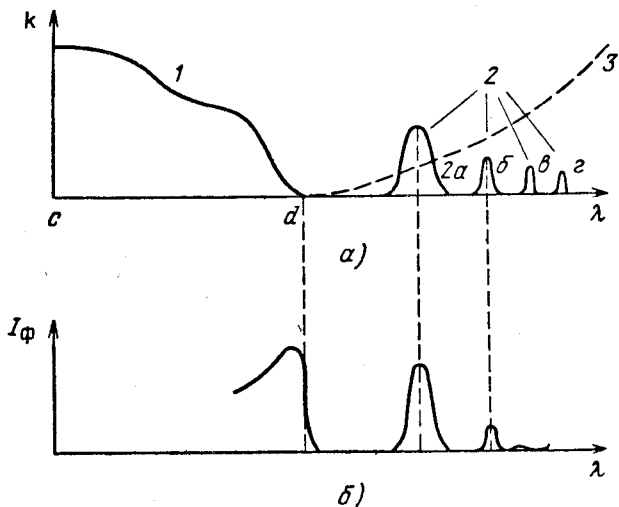


Рис. 1.30. Схема спектра поглощения полупроводника (а) и спектральное распределение фоточувствительности (б):

1 — полоса собственного поглощения (межзонные подходы); 2 — поглощение, обусловленное ионизацией примесных уровней (2 б, в, з — поглощение, обусловленное возбуждением примесных уровней); 3 — поглощение свободным электроном.

называемую полосу собственного поглощения ($c - d$), непосредственно связанного с межзонными переходами, где коэффициент поглощения достигает очень больших значений $K = 10^4 \div 10^6 \text{ см}^{-1}$ *) и область прозрачности (правее d), в которой поглощение сравнительно мало и обусловлено наличием примесей и свободных носителей.

*) Так как интенсивность света убывает по закону $I = I_0 e^{-\mu x}$, то это значит, что свет практически полностью поглощается на глубине $10^{-4} - 10^{-6} \text{ см}$.

При низких температурах и при исследованиях спектрометрами с большой разрешающей силой иногда обнаруживается сложная структура края полосы собственного поглощения, т. е. здесь появляется ряд более или менее ярко выраженных дискретных линий. В полупроводниках нестехиометрического состава, а также при наличии примесей и дефектов в полосе прозрачности появляются дополнительные линии примесного поглощения (рис. 1.30, *a*, кривые 2), имеющие более или менее ярко выраженную куполообразную форму.

Наконец, в полупроводниках с большой проводимостью и тонкая структура края полосы поглощения, и примесные линии исчезают и в то же время появляется большой фон — заметное поглощение, растущее в сторону длинных волн (рис. 1.30, *a*, кривая 3).

Межзонные переходы. В чистом полупроводнике поглощение света и фотопроводимость связаны с перебросом электрона из заполненной зоны в свободную. Красная граница полосы определяется наименьшей энергией, необходимой для такого переброса, т. е. на первый взгляд должна определяться соотношением

$$h\nu_0 = \Delta\mathcal{E}_0. \quad (1.168)$$

Однако соотношение это не вполне точно. Дело в том, что при поглощении света должен соблюдаться не только закон сохранения энергии, но и закон сохранения импульса — электрон, «проглотивший» фотон, получает в наследство не только его энергию $h\nu$, но и импульс $p_\phi = h\nu/c$ (где c — скорость света). Следовательно, если электрон освобождается с самого верхнего края заполненной зоны, где его скорость v_1 равнялась нулю, то в верхней зоне он должен иметь такую скорость v_2 , чтобы выполнялось соотношение

$$\frac{h\nu_0}{c} = mv_2. \quad (1.169)$$

Условия (1.168) и (1.169) не могут быть выполнены одновременно; поэтому соотношение (1.168) должно быть заменено более точным:

$$h\nu_0 = \Delta\mathcal{E}_0 + \frac{mv_2^2}{2}. \quad (1.170)$$

Совместное решение уравнений (1.169) и (1.170) дает правильное значение граничной частоты ν_0 полосы погло-

щения. Однако если учесть, что импульс фотона $p_{\text{ф}} = hv/c$ относительно мал (так как скорость света c очень велика), то оказывается, что уравнения (1.168) и (1.170) дают почти одинаковое значение v_0 .

В более общем случае, когда электрон вырывается светом не с верхнего края валентной зоны, а с произвольного уровня, на котором он имел скорость v_1 , законы сохранения энергии и импульса приобретают вид

$$hv = \Delta \mathcal{E}_0 + \frac{mv_2^2}{2} - \frac{mv_1^2}{2} \quad (1.171)$$

и

$$\frac{hv}{c} = |mv_2 - mv_1|.$$

Из сказанного выше следует, что $|v_2| \approx |v_1|$. В общем случае вырывание светом электрона может происходить не только из валентной зоны, но и из более глубоких и не только в первую пустую зону, но и в более высокие. Совокупность всех возможных переходов и дает полосу фундаментального (собственного) поглощения.

Число электронов на различных уровнях в заполненных зонах и свободных состояний (в свободных зонах), а также вероятность соответствующих переходов различны. Поэтому значение k колеблется внутри полосы поглощения. Тепловое размытие и смещение края полосы поглощения обусловлено изменениями ширины запрещенной зоны, связанными с тепловыми колебаниями и тепловым расширением кристаллической решетки.

Казалось бы, что область fotocувствительности должна распространяться на всю полосу собственного поглощения. В действительности это не так: фотоэффект обычно наблюдается только при поглощении света вблизи красной границы полосы поглощения (см. рис. 1.30, б), которая поэтому также называется красной границей фотоэффекта. Объясняется это следующим.

При $hv \gg \Delta \mathcal{E}_0$ коэффициент поглощения очень велик и практически весь свет поглощается в очень тонком поверхностном слое полупроводника. При этом концентрация фотоэлектронов (и освобожденных ими примесных центров и дырок) оказывается очень велика, что соответственно увеличивает вероятность рекомбинации и уменьшает время жизни носителей. Кроме того, в поверхностном слое

всегда имеется большое число примесей и дефектов, облегчающих рекомбинацию и затрудняющих дрейф свободных электронов (т. е. снижающих их подвижность практически до нуля). В силу этих причин свет с частотой $\nu \gg \Delta \mathcal{E}_0/h$ поглощается в тонком поверхностном слое и обычно не вызывает заметного фотоэффекта. Таким образом, красная граница фотоэффекта определяется возможностью генерации носителей ($h\nu > \Delta \mathcal{E}_0$), а синяя — условиями движения и рекомбинации. Чем совершеннее поверхность кристалла и сам кристалл (т. е. чем меньше центров захвата и прилипания носителей и чем больше подвижность) и чем ниже температура, тем дальше распространяется фоточувствительность в сторону коротких волн.

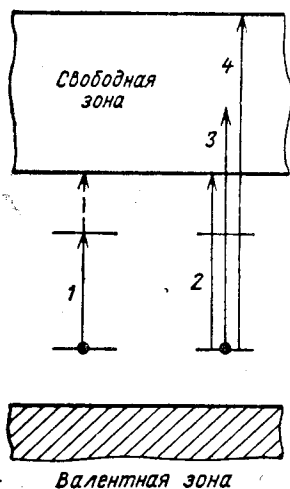


Рис. 1.31. Энергетический спектр примесного центра.

Примесное поглощение. На рис. 1.31 изображен схематически энергетический спектр донорного примесного уровня вблизи нижнего края зоны проводимости и стрелками показаны возможные переходы *). Таким образом, спектр поглощения примесного атома должен состоять из одной широкой линии (рис. 1.30, кривая 2а) и серии узких (кривые б, в, г), что действительно наблюдается на опыте.

Широкая линия соответствует полной ионизации атома и созданию фотоэлектрона, узкие линии —

возбуждению атома, который может быть потом ионизован тепловым движением или еще одним квантом или вернуться в нормальное состояние. Ширина линий, соответствующих возбужденным состояниям, возрастает с ростом температуры за счет теплового движения атомов.

Законы сохранения энергии и импульса в случае примесного поглощения имеют более сложную форму (так как в нем фигурирует «третье тело» — примесный атом), и здесь мы на них останавливаться не будем.

*) Среди этих переходов могут быть переходы, запрещенные правилами отбора [19].

Экситонное поглощение. Процессы, аналогичные рассмотренным выше, могут происходить и на основных (а не примесных) атомах решетки. При этом ионизация атома основной решетки означает переброс электрона из заполненной зоны в свободную и образование свободных электрона и дырки; этот процесс уже подробно рассматривался выше. Но наряду с этим возможно и другое возбужденное состояние основного атома, при котором электрон не отрывается от него, а лишь переходит на один из незаполненных уровней. Такое состояние получило название экситонного (от английского слова *excitation* — возбуждение). Отличие экситонного состояния от возбужденного состояния примесного атома заключается в том, что энергия кристалла не зависит (в силу трансляционной симметрии, т. е. эквивалентности одинаковых точек разных ячеек) от того, какой из атомов возбужден, поэтому возбужденное состояние может «путешествовать» по всему кристаллу *).

В результате этого образуются зоны экситонных состояний. Существенно, что движение экситонов не создает электрического тока, так как перемещается не электрон, а лишь возбужденное состояние атома. Экситон можно представить также как совместное движение электрона и дырки, связанных кулоновскими силами и вращающихся вокруг общего центра тяжести; кроме этого, такая пара может поступательно перемещаться по всему кристаллу.

Из закона сохранения импульса следует, что под действием света могут образовываться только неподвижные экситоны, поэтому экситонным возбуждениям соответствуют тонкие линии вблизи края полосы поглощения, с ростом температуры эти линии размываются тепловым движением. На образование экситона — возбуждение атома — требуется меньшая энергия, чем на его ионизацию, поэтому экситонные линии лежат в спектре поглощения справа от основной полосы, т. е. в области меньших частот и энергий.

Роль экситона в фотоэффекте, точно так же, как возбужденного состояния примесного атома, может быть различна. За время своей жизни он может быть разорван тепловым движением или еще одним квантом и таким образом дать

*) Грубо схематически этот процесс перемещения экситона можно представить следующим образом: экситон на данном атоме «захлопывается», т. е. электрон переходит в нормальное состояние, выделившаяся при этом энергия передается соседнему атому, на котором образуется экситон, и т. д.

два носителя — электрон и дырку. Экситон может при своем движении наткнуться на примесный атом и, «захлопнувшись», при этом ионизовать его, таким образом появится один свободный электрон.

Энергия рекомбинации экситона может быть также затрачена на заброс электрона из валентной зоны на акцепторный уровень, при этом образуется одна дырка. Энергия рекомбинации также может быть передана и тепловым колебаниям решетки или свободным электронам, при этом она будет растратена с точки зрения фотоэффекта зря, т. е. приведет к уменьшению квантового выхода.

Поглощение света свободными электронами. Как уже упоминалось, в сильно легированных полупроводниках возникает дополнительное поглощение, обусловленное свободными электронами. При этом следует учитывать, что совершенно свободные электроны не могут поглотить энергию; под действием электромагнитного поля они совершают синхронные с полем колебания и лишь вносят в силу этого соответствующий вклад в коэффициент преломления. Однако, как мы знаем, электрон, движущийся или колеблющийся под действием электрического поля, в твердом теле непрерывно испытывает столкновения. Если частота этих столкновений будет больше частоты электромагнитных колебаний, то значительная часть энергии будет переходить в тепло. Этим и объясняется поглощение света свободными носителями и рост его в сторону низких частот. Поглощение света свободными электронами не является фотоактивным процессом и приводит поэтому к уменьшению квантового выхода.

Поглощение света оптическими колебаниями решетки. В сложных кристаллах, т. е. в тех, в которых в одной элементарной ячейке содержится несколько атомов, наряду с акустическими колебаниями решетки имеются и оптические. При оптических колебаниях соседние атомы колеблются со значительным сдвигом по фазе по отношению друг к другу, это приводит к возникновению переменного дипольного момента, а следовательно, к испусканию света.

Если свет той же частоты падает на кристалл, то он возбуждает эти колебания и поглощается. Таким образом, в инфракрасной области спектра появляется ряд полос поглощения, по своему виду аналогичных примесному поглощению. Поглощение света оптическими колебаниями также приводит к уменьшению квантового выхода фотоэффекта.

СТРОЕНИЕ КРИСТАЛЛОВ

2.1. НЕКОТОРЫЕ ВОПРОСЫ КВАНТОВОЙ ТЕОРИИ

Ряд вопросов физики полупроводников можно изложить качественно, основываясь на представлениях классической теории, приписав электрону только корпускулярные свойства, а свету — только волновые. Однако существует большой круг явлений, неразрывно связанных с двойственной природой материи, говорить о которых невозможно, не пользуясь некоторым минимумом понятий квантовой механики. Постараемся изложить эти вопросы в возможно более краткой и простой форме.

ВОЛНЫ И ЧАСТИЦЫ

В XIX веке господствовали взгляды о волновой природе света, на основе которых удалось вывести все законы геометрической оптики, объяснить явления интерференции и дифракции и целый ряд других известных к тому времени закономерностей в поведении электромагнитного излучения.

Однако ряд принципиальных вопросов был необъясним. Так, согласно волновой теории равновесное для любой температуры электромагнитное излучение представляло собой систему с бесконечной энергией, так как в классической статистике на каждую степень свободы приходится одна и та же средняя энергия теплового движения, а электро-

Магнитное излучение представляет собой систему с бесконечным числом степеней свободы *).

Система с бесконечной энергией при любой температуре, отличной от нуля, должна обладать бесконечной теплоемкостью; следовательно, не только сама такая система, но и любая другая, равновесная с ней, должна находиться всегда при температуре абсолютного нуля.

Для объяснения этого парадокса один из величайших ученых, Планк, в 1900 г. предположил, что энергия \mathcal{E} волны с частотой ν может принимать не любые значения, а лишь кратные определенной порции (кванту) **):

$$\mathcal{E}_n = nh\nu, \quad (2.1)$$

где n — произвольное целое число,

h — универсальная постоянная (названная в дальнейшем постоянной Планка).

Этого предположения оказалось достаточно, чтобы ликвидировать так называемую „ультрафиолетовую катастрофу“ и открыть новую эру в физике — эру к в а н т о в о й физики.

На рис. 1.12 представлен энергетический спектр (т. е. совокупность возможных значений энергии) такого квантованного осциллятора. В соответствии со сказанным выше он состоит из серии линий, отстоящих друг от друга на равном расстоянии: $\mathcal{E}_0 = h\nu$ (энергия одного кванта).

Представим себе такой осциллятор в равновесии со стенками ящика или газом, средняя энергия молекул которого порядка kT (на чертеже обозначена пунктиром ***)). Если бы все молекулы обладали такой энергией, недостаточной для перевода осцилляторов в первое возбужденное состояние, то электромагнитное излучение этой частоты вообще

*) Представим себе излучение в ящике размером $l_x \times l_y \times l_z$, стенки которого имеют температуру T и находятся в термодинамическом равновесии с излучением. Тогда в направлении x в нем будут представлены волны $\lambda_x = 2l_x/n$, где $n = 1, 2, \dots, \infty$, и то же самое для y и для z . Каждая волна должна иметь среднюю энергию порядка kT , следовательно, полная энергия излучения должна быть бесконечна.

**) Впоследствии, в 1911 г., аналогичные представления о квантовании энергии колеблющихся атомов были предложены Эйнштейном для объяснения температурной зависимости теплоемкости твердого тела.

***) Представления о том, что энергия колебаний твердого тела также квантуется (которые появились позже) не изменили ни качественно, ни количественно дальнейших выводов этого параграфа.

не возбуждалось бы и его вклад в теплоемкость равнялся бы нулю. Но в действительности энергия атомов лишь в среднем порядка kT и всегда имеется небольшое число атомов, энергии которых много больше средней; эти атомы «изредка» будут возбуждать электромагнитные колебания.

Кроме того, как мы уже упоминали, в равновесном (для данной температуры и данного резервуара-ящика) электромагнитном излучении всегда имеется набор волн и каждой волне соответствует своя частота:

$$\nu = \frac{c}{\lambda}, \quad (2.2)$$

где c — скорость света.

На рис. 1.12, *в* представлен полный спектр излучения в ящике для всего набора частот. Как видно из рисунка, в спектре имеются и очень низкие частоты, которые будут возбуждаться и при очень низких температурах. Но для нас теперь важно отметить одно — развитые выше качественные представления снимают вопрос о бесконечности энергии.

Действительно, хотя электромагнитное излучение по-прежнему остается системой с бесконечным числом степеней свободы, но энергия его будет конечна, так как частоты с энергией $\epsilon \gg kT$ почти не будут возбуждаться. Эти соображения нетрудно оформить количественно.

Подсчитаем для этого среднюю энергию $\bar{\epsilon}$ осциллятора с частотой ν . По определению

$$\bar{\epsilon} = \frac{\sum \omega(\epsilon) \epsilon}{\sum \omega(\epsilon)} \quad (2.3)$$

(где $\omega(\epsilon)$ — вероятность того, что осциллятор приобретает энергию ϵ) будем считать так же, как в классической статистике, что $\omega(\epsilon)$ определяется выражением

$$\omega(\epsilon) = e^{-\frac{\epsilon}{kT}} = e^{-\alpha\epsilon},$$

где $\alpha = 1/kT$.

Полагая в соответствии с (2.1), что энергия ϵ может принимать значения

$$\epsilon = nh\nu$$

(где n — любое целое число), получим

$$\bar{\epsilon} = \frac{\sum_0^{\infty} e^{-\alpha nh\nu} nh\nu}{\sum_0^{\infty} e^{-\alpha nh\nu}} = \frac{\partial}{\partial \alpha} \ln \sum e^{-\alpha nh\nu}. \quad (2.4)$$

Но выражение, стоящее в (2.4) под знаком логарифма, есть сумма членов бесконечной геометрической прогрессии с знаменателем $e^{-\alpha h\nu}$, таким образом,

$$\sum_0^{\infty} e^{-\alpha n h\nu} = 1 + e^{-(\alpha h\nu)} + [e^{-(\alpha h\nu)}]^2 + \dots = \frac{1}{1 - e^{-\alpha h\nu}}, \quad (2.5)$$

и, дифференцируя по α логарифм выражения (2.5), получим окончательно

$$\bar{\epsilon}(\nu) = \frac{h\nu}{e^{h\nu/kT} - 1}. \quad (2.6)$$

Для того чтобы получить полную энергию равновесного излучения \mathcal{E} при данной температуре, надо просуммировать (2.6) по всем частотам; для достаточно большой полости суммирование можно заменить интегрированием, так как частоты расположены очень близко одна к другой, и нижний предел положить равным нулю (так как при $l \rightarrow \infty$, $\lambda \rightarrow \infty$ и $\nu = c/\lambda \rightarrow 0$)

$$\mathcal{E} = \int_{\nu=0}^{\nu=\infty} g(\nu) \bar{\epsilon}(\nu) d\nu,$$

где $g(\nu)$ — число колебаний в интервале частот ν , $\nu + d\nu$; можно показать, что для непрерывной среды

$$g(\nu) = \frac{8\pi\nu^2}{c^3} V,$$

(V — объем полости). Таким образом, получаем окончательно при $V=1$

$$\mathcal{E} = \int_0^{\infty} \frac{8\pi h\nu^2}{c^3} \frac{d\nu}{e^{h\nu/kT} - 1}. \quad (2.7)$$

Формула (2.7) и выражает знаменитое распределение Планка, из которого вытекают все основные энергетические закономерности электромагнитного излучения (и закон смещения Вина, и закон Стефана — Больцмана).

В дальнейшем, в 1907—1908 гг., для того чтобы объяснить законы фотоэффекта, второй величайший физик, Альберт Эйнштейн, предположил, что свет может испускаться и поглощаться также только в виде квантов с энергией $h\nu$.

Представления о квантах, введенные вначале в значительной степени формально, оказались чрезвычайно плодотворными. Оказалось, что поглощение или испускание кванта света означает рождение или исчезновение особой

частицы — фотона, обладающей энергией $\epsilon_\nu = h\nu$, скоростью, равной скорости света, и количеством движения (импульсом):

$$|\mathbf{p}| = \left| \frac{h\nu}{c} \right| = \left| \frac{h}{\lambda} \right| = |h\mathbf{k}|,$$

где $|\mathbf{k}| = 1/\lambda$ — волновой вектор, абсолютная величина которого равна числу волн на единицу длины (точно так же, как частота — числу колебаний в единицу времени); при этом целый ряд явлений можно с равным успехом объяснить и на волновом языке и на корпускулярном, а ряд других может быть объяснен только корпускулярной природой света.

С другой стороны, все явления, связанные с интерференцией, в частности дифракция, по-прежнему могли быть объяснены только на основе волновой теории.

Таким образом, физики впервые столкнулись с двойственной природой материи — в одних условиях излучение проявляло свои дискретные (корпускулярные) свойства, в других — волновые. Тот же, кто встречается с этими фактами (т. е. двойственной природой света) впервые, не должен этому особенно удивляться, ведь мы привыкаем к тому, что одни и те же предметы с разных точек зрения, в разных проекциях и в разных условиях выглядят совершенно по-разному: при одних обстоятельствах проявляются одни их свойства, при других — другие. Приблизительно так же было и со светом. Вначале человечество, имевшее больше дела с частицами, чем с волнами, и свету приписало корпускулярный характер и в этой концепции хорошо укладывались сравнительно скудные опытные факты, известные к тому времени. Потом были открыты явления интерференции и дифракции, которые потребовали признания волновой картины, но и она, как мы уже упоминали, оказалась безупречной, и в XX веке восторжествовала квантовая механика, включающая в себя и фотоны и волны.

Остается уточнить, как уживаются друг с другом эти понятия. Мы уже установили, что скорости волны и фотона одинаковы, что частота волны ν определяет энергию одного фотона $\epsilon_\nu = h\nu$; с другой стороны, полная электромагнитная энергия поля \mathcal{E} , как известно, измеряется квадратом амплитуды волны электромагнитного поля A_ν :

$$\mathcal{E}_\nu = \frac{A_\nu^2}{4\pi}.$$

Следовательно, число фотонов в 1 см^3

$$n = \frac{\mathcal{E}_\nu}{h\nu} = \frac{A_\nu^2}{4\pi h\nu}, \quad (2.8)$$

или, иными словами, вероятность нахождения фотона в любой точке поля пропорциональна квадрату амплитуды волны в этой точке.

В дальнейшем по аналогии с фотонами (носителями световой энергии) были введены фононы — носители звуковой энергии и вообще энергии упругих волн. Все, что было сказано выше о фотонах, было полностью перенесено на фононы: энергия фонона $\mathcal{E}_\nu = h\nu$, где ν — частота упругой волны, импульс $q = h\nu/\omega_\phi$, где ω_ϕ — скорость звуковой волны, и число фононов (или вероятность найти фонон в данной точке) пропорционально квадрату амплитуды упругой волны. Как уже упоминалось в первой главе, оказалось, что целый ряд свойств твердого тела (взаимодействия колебаний решетки с электронами, фотонами, всевозможными дефектами и друг с другом) можно объяснить и на «фононном языке» и на «волновом языке», а ряд явлений, в частности теплоемкость, при низких температурах — только с помощью фононов. Но, с другой стороны, и упругим волнам свойственны явления дифракции и интерференции, и, следовательно, им также пришлось приписать двойственную природу. Таким образом, свет и звук оказались в преимущественном положении перед весовыми частицами, которые оставались частицами в обычном смысле этого слова. Но такое положение сохранялось недолго. Целый ряд необъяснимых явлений требовал пересмотра применимости законов классической механики и по отношению к обычным частицам.

На первом месте следует поставить непонятную с классической точки зрения устойчивость атомов и молекул. С точки зрения классической механики электрон, колеблющийся или вращающийся в поле ядра, должен непрерывно излучать электромагнитную энергию и, теряя, таким образом, скорость, упасть, в конце концов, на ядро. Частота этого излучения должна состоять из основного тона и ряда кратных частот — обертонов, что также не соответствовало опытным фактам. Наконец, необъяснимо было то, что движение электронов внутри атома не увеличивает значения теплоемкости.

Для того чтобы объяснить эти факты, надо было распространить квантовые законы на теорию строения атома. Первый шаг в этом направлении был сделан в 1913 г. Нильсом Бором. Квантовая теория атома Бора была построена на следующих предположениях.

Каждый атом имеет ряд стационарных состояний, энергия которых определена и дискретна: находясь в любом из этих состояний, атом не излучает энергии. Орбиты электронов в этих состояниях определяются двумя условиями:

классической устойчивостью — по отношению к круговым орбитам это означает равенство центробежной и центростремительной сил

$$\frac{mv^2}{r_n} = \frac{ze^2}{r_n^2}, \quad (2.9)$$

где ze — заряд ядра и r_n — радиус орбиты, и

квантовой устойчивостью — момент количества движения для стабильной орбиты может принимать не любые значения, а лишь кратные постоянной Планка \hbar ; для круговой орбиты это означает:

$$mvr_n = n\hbar, \quad (2.10)$$

где n — любое целое число (номер орбиты).

Формулы (2.9) и (2.10) позволяют вычислить радиусы стабильных (стационарных) орбит

$$r_n = n^2 \frac{\hbar}{me^2} \frac{1}{z};$$

потенциальную энергию электрона на них

$$U_n = -\frac{e^2}{r_n} z = -\frac{me^4 z^2}{\hbar^2 n^2}; \quad (2.11)$$

кинетическую энергию

$$T_n = \frac{1}{2} mv^2 = \frac{me^4 z^2}{2\hbar^2 n^2} \quad (2.12)$$

и полную энергию

$$\mathcal{E}_n = T_n + U_n = -\frac{me^4 z^2}{2\hbar^2 n^2} = -\frac{A}{n^2}, \quad (2.13)$$

где $A = \frac{me^4 z^2}{2\hbar^2}$.

Приведенный выше вывод энергетических термов атома водорода можно с небольшими изменениями применить для определе-

ния энергии ионизации примесных атомов в полупроводнике; для этого в соотношениях (2.9) и (2.10) надо заменить массу свободного электрона эффективной массой и учесть в (2.9), что в среде с диэлектрической постоянной κ кулоновская сила уменьшается в κ раз; следовательно, (2.11а) примет вид

$$\frac{mv^2}{r_n} = \frac{ze^2}{\kappa r_n^2}; \quad (2.9a)$$

повторяя те же выкладки, получим для $z = 1$

$$U_n = -\frac{me^4}{\hbar^2 n^2 \kappa^2} \quad (2.11a)$$

и

$$\mathcal{E}_n = -\frac{me^4}{2\hbar^2 n^2 \kappa^2}. \quad (2.13a)$$

Соотношение (2.13а) сравнивалось рядом исследователей с энергией ионизации примесных уровней в германии и кремнии. Оказалось, что оно удовлетворительно согласуется с опытом для мелких уровней (т. е. донорных уровней, расположенных близко ко дну зоны проводимости, и акцепторных уровней, расположенных недалеко от верхнего края валентной зоны), создаваемых в германии и кремнии элементами пятой и третьей групп.

Для глубоких уровней, создаваемых другими элементами, энергии активации не имеют ничего общего со значениями, получаемыми из (2.13а).

Это можно понять, если учесть, что радиус орбиты [согласно (2.11)] обратно пропорционален энергии ее ионизации. Для мелких уровней радиус орбиты очень велик; орбита охватывает много атомов матричного вещества; в этом случае применимы и понятия диэлектрической постоянной и эффективной массы.

Напротив, для глубоких уровней электрон (по крайней мере на части своей траектории) проходит очень близко к ядру примесного атома; при этом понятия эффективной массы и диэлектрической постоянной среды теряют всякий смысл.

Атом излучает (или поглощает) электромагнитную энергию в виде одного кванта при переходе электрона с одной орбиты на другую; при этом энергия \mathcal{E}_{nm} кванта равна разности энергий электрона на этих орбитах (т. е. в начальном \mathcal{E}_n и конечном \mathcal{E}_m состояниях):

$$\mathcal{E}_{nm} = h\nu_{nm} = \mathcal{E}_n - \mathcal{E}_m = A \left(\frac{1}{n^2} - \frac{1}{m^2} \right) \quad (2.14)$$

где ν_{nm} — частота излучения *). Эта формула блестяще подтвердилась для спектров испускания и поглощения атома водорода.

*) Выражение (2.14) можно также написать в виде $\mathcal{E}_{nm} = \hbar\omega_{nm}$, где ω_{nm} — круговая частота, равная $2\pi\nu_{nm}$.

В 1916 г. Вильсон и Зоммерфельд распространили теорию Бора на эллиптические орбиты.

Итак, в 1913 г. Бор своими постулатами открыл дверь квантовой механике в теорию строения атома. Можно сказать, что по отношению к атому он сделал то, что Планк и Эйнштейн сделали в 1900—1907 гг. для электромагнитного излучения. Следующий решающий шаг был сделан Луи де Бройлем в 1925 г. Для того чтобы дать физическое объяснение постулатам Бора, де Бройль предположил, что и электрон, так же как фотон, имеет двойственную, и корпускулярную и волновую, природу, и перенес на электрон соответствующие соотношения между импульсом и волновым вектором:

$$\mathbf{p} = \hbar \mathbf{k} \quad (2.15)$$

и энергией и частотой

$$\varepsilon = \hbar \nu, \quad (2.16)$$

введенные Эйнштейном для фотона в 1907 г.

Соотношение (2.15) придает условию стационарности орбиты (2.10) весьма наглядный вид. Действительно, заменяя в (2.10) импульс электрона через длину волны, получим

$$\frac{2\pi \hbar r}{\lambda} = n \hbar \quad (2.17)$$

или

$$2\pi r = n\lambda, \quad (2.18)$$

т. е. условие (2.10) означает, что на орбите должно укладываться целое число волн.

В таком виде условие (2.10) аналогично условиям устойчивости колебаний обычной классической системы в ограниченном пространстве (струны, мембраны и т. д.) и означает, что волна при многократном повторении своего пути должна оставаться в фазе сама с собой *).

Удивительно смелая идея приписать электрону волновые свойства тогда, когда не было для этого никаких экспе-

*) Если в ограниченной среде могут распространяться колебания, то в ней устанавливаются стоячие волны, форму которых можно найти из граничных условий. Так, например, для струны, закрепленной по концам, мы находим длины волн и частоты из того условия, что амплитуды всех колебаний на концах должны обращаться в нуль. Аналогично для «замкнутой струны» — орбиты — собственные длины волн должны находиться из условий цикличности (2.18).

риментальных оснований, увенчалась блестящим успехом. В 1927 г. Дэвидсон и Джермер подтвердили представления о волновой природе электронов прямыми опытами по их дифракции.

В первой половине XX века рядом величайших ученых была создана квантовая механика — новая теория, включившая в себя классическую физику и сумевшая объяснить волновые свойства обычных частиц и корпускулярный характер электромагнитного излучения. Здесь мы изложим лишь некоторые простейшие положения этой теории, необходимые для понимания дальнейшего.

Согласно этой теории все обычные частицы (электроны, протоны, атомы, молекулы и т. д.) также обладают двойственной природой — волновой и корпускулярной; их основное отличие от фотонов и фононов заключается в том, что их масса покоя отлична от нуля и что они могут двигаться с любой скоростью, не превышающей скорости света (в том числе и со скоростью, равной нулю).

Мы уже упоминали, что электрон (в дальнейшем мы для определенности будем говорить только об электронах, движущихся со скоростью v), описывается волной с волновым вектором $\mathbf{k} = mv/h$ и частотой $\nu = \epsilon/h$, где ϵ — энергия электрона.

В отличие от обычного волнового движения, которое описывается тригонометрическими функциями, волновая функция ψ , описывающая волны материи, должна быть записана в комплексной форме:

$$\psi = Ae^{2\pi i (kx - \nu t)} = Ae^{\frac{2\pi i}{h} (px - \epsilon t)} \quad (2.19)$$

Квадрат амплитуды волновой функции A^2 характеризует вероятность w нахождения частицы в данной точке пространства (для одномерного случая в данной точке оси x):

$$w \sim A^2 = \psi\psi^*,$$

где ψ^* — комплексно-сопряженная волновая функция:

$$\psi^* = Ae^{-2\pi i (kx - \nu t)}.$$

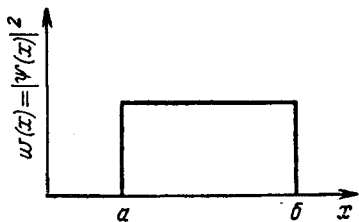
При этом возникает вопрос: если амплитуда A волны ψ постоянна, то это значит, что электрон может с равной вероятностью находиться в любой точке оси x ; как же описать электрон, находящийся на вполне определенном

отрезке оси, скажем $a < x < b$ (рис. 2.1). Можно выйти из положения, сделав переменной амплитуду: при $x < a$ и $x > b$, $A = 0$, а $a < x < b$

$$A = \frac{1}{\sqrt{ab}} \text{ и } \int_a^b \omega(x) dx = \int_a^b A^2 dx = 1,$$

но такое решение нельзя считать окончательным. Как известно, представленный на рис. 2.1 волновой пакет уже

Рис. 2.1. График квадрата амплитуды волновой функции для одномерного движения. Электрон локализован на отрезке ab .



не является монохроматической волной; амплитуда его также раскладывается в интеграл Фурье:

$$A(x) = \int_{k=0}^{\infty} A(k) \sin(kx + \varphi_k) dk.$$

При этом $A(k)$ близко к нулю при всех $k \gg 1/ab$ и принимает наибольшие значения ($\sim \sqrt{ab}$) при $0 < k < 1/ab$.

Таким образом, задав координаты электрона в пределах отрезка $\Delta x = a$, мы уже не можем описать его монохроматической волной, а должны представить в виде волнового пакета, волновой вектор которого принимает значения в пределах $\Delta k \approx 1/a$. Чем больше мы ограничиваем координаты электрона, тем более неопределенным становится его волновое число

$$\Delta k \Delta x \geq 1 \quad (2.20)$$

и согласно соотношению (2.15)

$$\Delta p_x \Delta x \geq h \quad (2.21)$$

или, так как $p_x = mv_x$, то

$$\Delta mv_x \Delta x \geq h. \quad (2.22)$$

В случае трехмерного движения можно вывести совершенно аналогичные соотношения для других проекций

импульса и координат:

$$\Delta p_y \Delta y \geq h \quad (2.21a)$$

и

$$\Delta p_z \Delta z \geq h. \quad (2.21б)$$

Можно также показать, что между временем τ пребывания электрона в некотором состоянии и энергией этого состояния ε имеется аналогичное соотношение:

$$\Delta \tau \Delta \varepsilon \geq h, \quad (2.21в)$$

т. е. чем точнее зафиксировано время пребывания электрона в данном состоянии, тем более неопределенной становится его энергия.

Соотношения (2.21), (2.21а, б, в) носят название принципа неопределенности Гейзенберга и являются важнейшими положениями квантовой механики.

ФАЗОВАЯ ПЛОСКОСТЬ И ФАЗОВОЕ ПРОСТРАНСТВО

Для того чтобы дать соотношению (2.21) геометрическую интерпретацию, отложим на рис. 2.2 по оси абсцисс координату электрона (x), а по оси ординат — его импульс (p_x).

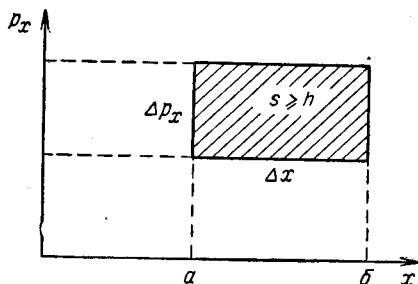


Рис. 2.2. Фазовая плоскость для одномерного движения электрона.

Плоскость с такими координатами носит название фазовой плоскости. В классической физике каждый электрон (и любая другая частица) имеют вполне определенную координату и импульс. Следовательно, состояние частицы характеризуется определенной точкой на фазовой плоскости. Более того, координаты частицы в фазовой плоскости в данный момент определяют всю ее дальнейшую судьбу.

В самом деле, зная импульс частицы, т. е. ее скорость, мы можем точно предсказать изменение ее координаты: $dx = v_x dt = (p_x/m) dt$.

С другой стороны, мы знаем и силы, действующие на частицу (так как силы могут зависеть только от координат и скорости, которые точно известны), и, следовательно, можем предсказать изменение ее импульса: $dp_x = F dt$.

Таким образом, положение частицы на фазовой плоскости в данный момент определяет ее положение в следующий момент и, следовательно, всю ее траекторию. Если частица движется не в одномерном, а трехмерном пространстве, то ее состояние будет характеризоваться тремя координатами и тремя проекциями импульса. Следовательно, в этом случае мы должны вместо фазовой плоскости ввести шестимерное фазовое пространство; в остальном же все сказанное выше остается в силе. Положение точки в фазовом пространстве определяет ее прошлое и будущее, так как через каждую точку проходит только одна фазовая траектория и ни одна траектория не пересекается с другой *).

Совершенно по-иному обстоит дело в квантовой механике; начнем опять с разбора одномерного случая. Как мы только что убедились, в классической механике состояние частицы характеризуется для одномерного движения точкой на фазовой плоскости. Соотношение (2.21) показывает, что в квантовой механике состояние частицы характеризуется некоторой площадью на фазовой плоскости: $S \geq h$.

Можно представить, что вся фазовая плоскость разбита на ячейки площадью h и свойства элементарных частиц таковы, что они не могут уместиться на меньшей площади, чем площадь одной ячейки. При этом форма таких ячеек зависит от характера движения частиц, но размер всегда

*) Если мы рассматриваем не одну частицу, а систему из N независимых частиц, то ее состояние определяется заданием $3N$ координат и $3N$ импульсов; следовательно, чтобы характеризовать такую систему одной точкой, нужно ввести понятие о фазовом пространстве, имеющем $6N$ измерений. Объем фазовой ячейки в таком гиперпространстве равен h^{3N} . Так как величина h очень мала, то для системы, состоящей из очень большого числа частиц, элементарная ячейка становится ничтожно малой. Поэтому такие макроскопические системы хорошо описываются классической механикой.

равен h^*). Эти представления позволяют по-другому сформулировать принцип Паули: в одной ячейке фазового пространства может поместиться не более двух электронов с противоположно направленными спинами.

Таким образом, если в классической механике состояние частицы характеризуется точкой, то в квантовой согласно (2.22) характеризуется прямоугольником площадью $S = \Delta x \Delta p \geq h$, при этом со временем эта неопределенность будет увеличиваться, так как

$$\Delta x_t = \Delta x_{t=0} + \frac{\Delta p}{m} t \quad (2.23)$$

и

$$\Delta p_t = \Delta p_{t=0} + \Delta F t, \quad (2.24)$$

где ΔF — неопределенность в силе, возникающей из-за неопределенности в координате.

Таким образом вместо определенной (классической) траектории в квантовой механике мы имеем пучок расходящихся траекторий.

Строгий анализ показывает, что соотношения, которые мы пояснили на весьма частном примере, в действительности носят весьма общий характер и полностью сохраняются для трехмерного пространства (и шестимерного фазового пространства).

В этом случае, как мы уже упоминали,

$$\begin{aligned} \Delta x \Delta p_x &\geq h, \\ \Delta y \Delta p_y &\geq h, \\ \Delta z \Delta p_z &\geq h \end{aligned} \quad (2.25)$$

и, следовательно, перемножая, получаем

$$\Delta V = \Delta v \Delta \Omega \geq h^3, \quad (2.26)$$

где Δv — объем в обычном пространстве;

$\Delta \Omega$ — объем в пространстве импульсов и

ΔV — объем в шестимерном фазовом пространстве.

Таким образом, обобщенные координаты частицы (включая сюда и проекции импульса) могут быть определены не точнее положения частицы в элементарной ячейке фазового пространства объемом h^3 , причем с течением вре-

*) При движении частицы в трехмерном пространстве фазовое пространство имеет шесть измерений и объем одной ячейки равен h^3 (см. ниже).

мени этот начальный объем еще более «расплывается». Теория также показывает, что в зависимости от конкретных условий существования частицы форма элементарной ячейки фазового пространства может варьироваться в весьма широких пределах, но объем ее всегда остается равным h^3 . Может возникнуть естественный вопрос: почему же не «расплываются» все предметы, существующие вокруг нас и мы сами?

В действительности дело обстоит следующим образом. Для макроскопических предметов неопределенность, возникающая в результате соотношения (2.23), весьма незначительна. В самом деле, представим себе пылинку весом 10^{-3} г и представим, что ее начальная координата Δx_0 определена с точностью до 10^{-3} см. В этом случае минимальная неопределенность в начальной скорости будет

$$\Delta v_x = \frac{h}{m\Delta x} \sim 10^{-20}.$$

Это значит, что со временем неопределенность в координате будет расти по закону

$$\Delta x = \Delta x_0 + \Delta v t = 10^{-3} + 10^{-20}t,$$

т. е. Δx удвоится приблизительно через 10^{12} лет! Очевидно, что за такой срок и с самой пылинкой и во всей вселенной произойдут более существенные перемены. Этот пример, таким образом, иллюстрирует тот факт, что принцип неопределенности не отражается заметным образом на характере движения макроскопических тел.

Теоретический анализ показывает, как мы уже упоминали, что между энергией системы \mathcal{E} , находящейся в каком-то определенном состоянии, и временем, в течение которого частица находится в этом состоянии $\Delta\tau$, существует соотношение неопределенности, аналогичное (2.21):

$$\Delta\mathcal{E}\Delta\tau \geq h, \quad (2.27)$$

т. е. чем точнее зафиксировано время существования данной системы, тем неопределенней становится его энергия. Можно точно так же, как это было сделано только что для соотношения (2.21), показать, что (2.27) не сказывается существенно на поведении макроскопических объектов.

К совершенно иным выводам приводит применение этих соотношений к микромиру, например, к поведению электрона в пределах атома. Действительно, полагая $m \approx 10^{-27}$ г

(масса электрона) и $\Delta x = 6 \cdot 10^{-8}$ см (атомные размеры), получаем неопределенность в скорости:

$$\Delta v_x = \frac{h}{m\Delta x} \approx 10^8 \frac{\text{см}}{\text{сек}}.$$

С другой стороны, простые расчеты показывают, что сама скорость валентных электронов в атоме того же порядка (10^8 см/сек); это значит, что в данном случае понятие об орбите теряет всякий смысл и необходимо ввести новые представления и новую теорию, которая описывала бы поведение электрона в атоме*). Эту новую теорию, являющуюся дальнейшим развитием представлений Бора и де Бройля, создали в 1926—1927 гг. Шредингер, Гейзенберг, Борн и др.

УРАВНЕНИЕ ШРЕДИНГЕРА

В предыдущем параграфе мы рассмотрели движение электрона в свободном пространстве; в этом случае он (т. е. электрон) описывается плоской волной или пакетом из плоских волн. Не менее важными являются случаи, когда частица имеет отрицательную потенциальную энергию, ограничивающую ее движение в каком-то объеме (т. е. рассматриваемая нами частица связана силами притяжения с какой-либо другой частицей или системой частиц и ее кинетическая энергия недостаточна, чтобы преодолеть это притяжение). Сюда относится движение электронов в атоме, молекуле и кристалле, колебания атомов в молекуле и кристалле и др. Мы должны найти волну (волновую функцию) ψ , описывающую электрон (или другую частицу) в этих случаях.

В классической механике движение частицы или системы частиц определялось уравнениями движения (написанными в форме второго закона Ньютона или уравнениями Лагранжа, Гамильтона или в форме вариационного принципа) и начальными условиями. Очевидно, что и в квантовой механике должны существовать дифференциальные (или

*) Можно показать, что для «сильно» возбужденных состояний, т. е. с большим n в формуле (2.13), классические представления опять приобретают смысл.

интегральные) уравнения, которые бы описывали движение частицы.*)

Нетрудно убедиться, что плоская волна (2.19) является решением дифференциального уравнения

$$\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} = -\mathcal{E} \psi \quad (2.28)$$

или для объемного случая

$$\frac{\hbar^2}{2m} \Delta \psi = -\mathcal{E} \psi, \quad (2.29)$$

где

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Уравнения (2.28) и (2.29) описывают свободный электрон; необходимо их обобщить таким образом, чтобы в них входила потенциальная энергия U .

Нетрудно убедиться, что левая часть (2.29) представляет собой кинетическую энергию электрона. Действительно, согласно (2.19)

$$\frac{\hbar^2}{2m} \Delta \psi = -\frac{p^2}{2m} \psi = -T \psi \quad (2.29a)$$

и, следовательно, (2.29) можно рассматривать как волновую формулировку закона сохранения энергии; в случае наличия потенциальной энергии U естественно его переписать в форме $T \psi = (\mathcal{E} - U) \psi$ или

$$\frac{\hbar^2}{2m} \Delta \psi = -(\mathcal{E} - U) \psi,$$

или

$$-\frac{\hbar^2}{2m} \Delta \psi + U \psi = \mathcal{E} \psi. \quad (2.30)$$

Уравнение (2.30) и есть уравнение Шредингера для стационарных состояний. Разумеется, приведенные выше рассуждения ни в какой мере не могут претендовать на обоснование или, тем более, доказательство этого уравнения, и Шредингер вывел его на основе несоизмеримо более глу-

*) Мы начнем с рассмотрения стационарных (не зависящих от времени) состояний. В этом случае: 1) время не должно входить явно в уравнение для определения ψ и 2) начальные условия должны быть заменены какими-то другими условиями: однозначностью волновой функции ψ , ее конечностью, убыванием при $x \rightarrow \infty$ и т. д.

бокого анализа существовавшей ситуации. С другой стороны, следует заметить, что, как большинство вновь открываемых законов природы, оно не могло быть строго получено из старых представлений, а требовало для своего вывода элемента догадки.

Задавая конкретный вид потенциальной энергии и дополнительные условия, соответствующие данной задаче, можно получить конкретные выражения волновой функции для любой физической задачи:

- 1) для электрона в атоме водорода *) $U = -e^2/r$;
- 2) для частицы, связанной с положением x_0 равновесия упругой силой

$$F = -f(x - x_0) \quad \text{и} \quad U = -\frac{1}{2}f(x - x_0)^2; \quad (2.31)$$

частица, описываемая потенциальной энергией такого вида, называется линейным осциллятором. Такой вид U приблизительно соответствует внутримолекулярным колебаниям атомов и т. д.

Остановимся коротко еще на одном из основных положений квантовой механики — вероятности переходов из одного состояния в другое. Как нетрудно убедиться, волновая функция (2.19), описывающая движение свободного электрона, удовлетворяет также следующему дифференциальному уравнению:

$$\frac{\hbar^2}{2m} \Delta \psi = \frac{\hbar}{i} \frac{\partial \psi}{\partial t}, \quad (2.32)$$

в которое уже явно входит время; будем исходить из него для обобщения уравнения Шредингера на нестационарные процессы. Заменяя в (2.30) энергию \mathcal{E} оператором $(\hbar/i)(\partial/\partial t)$, получим

$$\mathcal{E} \psi = \frac{\hbar}{i} \frac{\partial \psi}{\partial t} \quad (2.33)$$

[в справедливости этого равенства можно убедиться дифференцированием (2.19) по времени] и оставив все другие

*) В других атомах, даже атоме гелия, вид потенциальной энергии значительно более сложен, так как она должна включать взаимодействия с другими электронами; задача решается лишь приближенно.

члены без изменения, получим уравнение (2.30), обобщенное для нестационарных процессов. При этом оно приобретает следующий вид:

$$-\frac{\hbar^2}{2m} \Delta\psi + U\psi = \frac{\hbar}{i} \frac{\partial\psi}{\partial t}. \quad (2.34)$$

В этом виде следует писать и решать уравнение Шредингера в том случае, когда потенциальная энергия электрона зависит от времени. Уравнение (2.34) также может быть переписано в виде

$$H\psi = \frac{\hbar}{i} \frac{\partial\psi}{\partial t}, \quad (2.34a)$$

где H — оператор Гамильтона (см. стр. 132).

В ряде случаев и здесь удастся упростить задачу, прибегнув к приближенным методам, а именно: в тех случаях, когда потенциальную энергию можно разбить на два члена $U = U_0 + V(t)$, где $V(t)$ — часть, зависящая от времени, много меньше чем U_0 , не зависящее от времени, решают вначале уравнение (2.34) в виде (2.30), подставив вместо U символ U_0 , и находят таким образом стационарные состояния систем $\psi_1, \psi_2, \dots, \psi_i, \dots, \psi_k$, а затем рассматривают $V(t)$ как возмущение, делающее эти состояния нестационарными, т. е. обуславливающее переходы между ними.

Можно показать, что вероятность таких переходов ω определяется квадратом так называемого матричного элемента возмущения:

$$\omega_{ik} \sim \left| \int \psi_i V \psi_k \, d\tau \right|^2. \quad (2.35)^*$$

НЕКОТОРЫЕ СВОЙСТВА УРАВНЕНИЯ ШРЕДИНГЕРА И ВОЛНОВЫХ ФУНКЦИЙ

Уравнение (2.30) имеет для каждого вида $U(\mathbf{r})$ целый ряд решений $[\psi_i(\mathbf{r})]$, каждому из которых соответствует вполне определенная энергия \mathcal{E}_i ; но одному и тому же значению энергии \mathcal{E}_i очень часто соответствует целый ряд различных решений $\psi_i^1, \psi_i^2, \dots, \psi_i^n$; в этом случае мы говорим, что это энергетическое состояние n -кратно вырождено.

*) В этом и других случаях, когда буквой V обозначается возмущение, для обозначения объема используется буква v .

При этом линейная комбинация из этих функций

$$\psi_i = \sum_{k=1}^n c^k \psi_i^k \quad (2.36)^*$$

(где c^k — произвольные постоянные коэффициенты) будет также описывать состояние системы, соответствующее той же энергии — это следует из линейности уравнения Шредингера. Чем более симметричный вид имеет $U(r)$, тем в большей степени вырождено состояние системы. Мы уже столкнулись с этим в первой главе при описании атома водорода: состояния с различным эксцентриситетом и различной ориентацией орбит соответствовали одной и той же энергии. Если к потенциальной энергии $U(r)$ добавляется член, обладающий меньшей симметрией (например, в случае атома — внешнее поле), то степень вырождения уменьшается или может совсем исчезнуть. Этим объясняются эффект Зеемана — расщепление спектральных линий в магнитном поле и эффект Штарка — в электрическом.

Уравнение (2.30) может быть переписано также в виде

$$H\psi = \mathcal{E}\psi, \quad (2.38)$$

где дифференциальный оператор

$$H = -\frac{\hbar^2}{2\pi m} \Delta + U(r) = T + U \quad (2.39)$$

называется оператором Гамильтона.

Таким образом, можно сказать, что совокупность дозволённых значений энергии $\mathcal{E}_1, \dots, \mathcal{E}_i$, которые могут частично или целиком образовывать непрерывный ряд, есть собственные значения оператора Гамильтона [для данного конкретного вида $U(r)$] и соответствующие им функции — суть собственные функции оператора Гамильтона.

Оператор Гамильтона является так называемым эрмитовым оператором (см., например, В. И. Смирнов. Курс высшей математики, т. II). Собственные функции любого не зависящего от времени эрмитова оператора ψ_1, \dots, ψ_n

*) При этом мы всегда должны заботиться о том, чтобы соблюдалось условие нормировки

$$\int \psi_i \psi_i^* dv = \int w(r) dv = 1. \quad (2.37)$$

Это условие следует из того очевидного факта, что электрон часто должен «где-то» находиться, и поэтому, интегрируя его вероятность нахождения в данной точке по всему пространству, мы должны получить единицу; этому соотношению должны удовлетворять коэффициенты c^k .

образуют ортогональную и полную систему функций. Это значит, что

$$\int_{i \neq k} \psi_i^* \psi_k \, dv = 0. \quad (2.40)^*$$

Это условие соответствует условию ортогональности векторов в обычном пространстве. Свойство полноты означает, что любая пространственная функция φ и в частности функция данного конкретного или другого оператора Гамильтона, может быть представлена в виде суммы

$$\varphi_k = \sum a_{ki} \psi_i; \quad (2.41)$$

в случае, если φ_k есть одна из собственных функций, $\varphi = \psi_k$, то $a_k = 1$, а все остальные коэффициенты равны нулю. Для того чтобы получить коэффициенты a_{ki} , в общем случае умножим и левую и правую части (2.41) на ψ_i^* и проинтегрируем; тогда согласно (2.40) и (2.42)

$$a_{ki} = \int \varphi_k \psi_i^* \, dv. \quad (2.43)$$

Разложение в ряд Фурье есть частный случай разложения (2.41). Действительно, все гармоники ряда Фурье есть не что иное, как совокупность собственных функций Гамильтона для случая $U = 0$, — в этом случае уравнение Шредингера превращается в обычное волновое уравнение (2.28), из которого мы и исходили.

Сопоставление соотношений (2.40), (2.41) и (2.42) допускает и другое, более широкое толкование. Действительно, можно представить, что полная, ортогональная система функций ψ_1, \dots, ψ_n представляет собой системы ортогональных векторов в n -мерном пространстве (где n большей частью бесконечно) и что формула (2.41) представляет собой разложение вектора φ по нормированным векторам этой системы; коэффициенты a_{ki} являются обобщенными косинусами углов между векторами φ_k и ψ_i , *т. е. задается*

*) Кроме этого, как упоминалось выше, условие нормировки требует, чтобы

$$\int \psi_i \psi_i^* \, dv = 1, \quad (2.42)$$

поэтому (2.40) и (2.42) мы можем объяснить в одном равенстве:

$$\int \psi_i \psi_k^* \, dv = \delta_{ik},$$

где $\delta_{ik} = 0$ при $i \neq k$ и $\delta_{ik} = 1$ при $i = k$.

С этой точки зрения сами волновые функции, выраженные в виде функций координат, являются косинусами перехода из обычного пространства в пространство ψ -функций; квадрат косинуса, т. е. $\psi_n(r) \psi_n^*(r)$, характеризует вероятность того, что, находясь в данном ψ_n -м состоянии, частицы имеют координату r .

Соотношение ортогональности (2.40) означает, что если частица имеет энергию \mathcal{E}_k и находится в k -м состоянии, то она не может одновременно находиться в i -м состоянии, и наоборот; иными словами, это означает, что оба эти состояния при отсутствии дополнительных воздействий (возмущений) стационарны.

Если же появляются какие-то дополнительные поля, а следовательно, и соответствующие дополнительные члены (в потенциальной энергии), то они могут вызвать два следствия: 1) изменения в стационарных значениях энергии и соответствующих им волновых функциях; это имеет место, если возмущение не зависит от времени, например в эффектах Зеемана и Штарка, и 2) в том случае, когда возмущение зависит от времени — при его учете стационарные состояния перестают быть таковыми, появляется вероятность переходов между ними. Наиболее характерными примерами такого возмущения является воздействие на систему электромагнитного излучения и столкновений.

НЕКОТОРЫЕ СООТНОШЕНИЯ ТЕОРИИ ВОЗМУЩЕНИЙ

Возмущение, не зависящее от времени. Уравнение Шредингера решается точно для небольшого числа случаев, для сравнительно простых видов потенциальной энергии $U_0(r)$ и гамильтониана

$$H_0 = T + U_0(r), \quad (2.44)$$

где

$$T = -\frac{\hbar^2}{8\pi^2m} \Delta = -\frac{\hbar^2}{8\pi^2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \quad (2.45)$$

— так называемый оператор кинетической энергии *).

*) Нетрудно убедиться, что применение этого оператора к волновой функции свободного электрона

$$\Psi = A \exp \left[\frac{2\pi i}{\hbar} (p_x x + p_y y + p_z z) \right] \quad (2.47a)$$

В тех случаях, когда вид потенциальной энергии U , а следовательно, и гамильтониана более сложен, задачу приходится решать приближенно. Это удастся сделать сравнительно просто, когда потенциальную энергию U можно разбить на две части:

$$U = U_0 + V, \quad (2.46)$$

где U_0 — вид потенциальной энергии, допускающий точное решение;

V — небольшая добавка, которую называют возмущением.

В соответствии с (2.40)

$$H = H_0 + V. \quad (2.47)$$

Будем считать, что невозмущенное уравнение Шредингера

$$H_0 \psi^0 = \mathcal{E}^0 \psi^0 \quad (2.48)$$

имеет известные волновые функции $\psi_1^0, \psi_2^0, \dots$ и соответствующие значения энергии $\mathcal{E}_1^0, \mathcal{E}_2^0, \dots$. Будем вначале считать, что эти состояния не вырождены, т. е. каждому значению энергии соответствует одна волновая функция.

Наша задача состоит в том, чтобы найти решения ψ_1, ψ_2, \dots и соответствующие значения энергии $\mathcal{E}_1, \mathcal{E}_2, \dots$ для возмущенной задачи:

$$(H_0 + V) \psi = \mathcal{E} \psi. \quad (2.49)$$

При этом, так как возмущение невелико: $V \ll H$, можно ожидать, что изменения волновых функций и энергетических уровней будут также невелики, т. е. $\psi_m \approx \psi_m^0$ и $\mathcal{E}_m \approx \mathcal{E}_m^0$.

Для нахождения решений уравнения (2.49) разложим его решения по невозмущенным функциям ψ^0 :

$$\psi = \sum a_m \psi_m^0; \quad (2.50)$$

таким образом, наша задача сводится к нахождению коэффициентов a_m .

равносильно ее умножению на кинетическую энергию; операторы $p_x = (h/2\pi i) (\partial/\partial x)$, $p_y = (h/2\pi i) (\partial/\partial y)$ и $p_z = (h/2\pi i) (\partial/\partial z)$ называются операторами проекций импульса; применение их к этой же волновой функции эквивалентно умножению на соответствующую проекцию импульса. Подробнее о значении операторов в квантовой механике см. [6].

Подставив (2.50) в (2.49) и учтя (2.48), получим

$$\sum_m a_m (\mathcal{E}_m^0 + V) \psi_m^0 = \sum_m a_m \mathcal{E} \psi_m^0. \quad (2.51)$$

Умножив левую и правую части (2.51) на ψ_k^{0*} и проинтегрировав, получим

очень легко $(\mathcal{E} - \mathcal{E}_k^0) a_k = \sum V_{km} a_m, \quad (2.52)$

смысл
где введено обозначение

$$V_{km} = \int \psi_k^{0*} V \psi_m^0 dv, \neq \int V \psi_k^{0*} \psi_m^0 dv \quad (2.53)$$

величина V_{km} называется матричным элементом возмущения. *)

Будем решать уравнение (2.52), в котором неизвестными являются коэффициенты a_{km} и значения энергии \mathcal{E}_m , методом итераций (последовательных приближений), т. е. будем считать в соответствии с (2.47) и (2.50), что

$$\mathcal{E}_m = \mathcal{E}_m^0 + \mathcal{E}_m^1 + \mathcal{E}_m^2 + \dots$$

и

$$a_m = a_m^0 + a_m^1 + \dots, \quad (2.54)$$

где величины \mathcal{E}_m^1 и a_m^1 того же порядка малости, что и возмущение; \mathcal{E}_m^2 и a_m^2 — величины второго порядка малости и т. д.

Чтобы из (2.46) получить поправку первого порядка к n -му уровню энергии \mathcal{E}_n , положим (в нулевом приближении) $a_n = 1$, а все остальные коэффициенты равными нулю, тогда

$$\mathcal{E}_n - \mathcal{E}_n^0 = V_n' = V_{nn} = \int \psi_n^{0*} V \psi_n^0 dv = \int V |\psi_n^0|^2 dv. \quad (2.55)$$

Как во втором

Таким образом, поправка к энергии n -го состояния равна энергии возмущения, усредненной по всему пространству в соответствии с плотностью распределения вероятности электрона в пространстве в этом состоянии: $\omega_n = |\psi_n^0|^2$. Полагая теперь $k \neq n$ и сохраняя в левой и правой частях

*) При $V = \text{const}$ (т. е. в отсутствие возмущения) интеграл (2.53) в силу ортогональности векторов ψ_k и ψ_m равен 0. Таким образом, V_{km} характеризует, насколько возмущение «перекосило» по отношению друг к другу эти два вектора, т. е. степень их неортогональности.

(2.46) члены первого порядка малости, получим

$$a_k^1 = \frac{V_{kn}}{\mathcal{E}_k^0 - \mathcal{E}_n^0}. \quad (2.56)$$

Выражение (2.49) дает более точный критерий применимости теории возмущений:

$$V_{kn} \ll \mathcal{E}_k^0 - \mathcal{E}_n^0, \quad (2.57)$$

т. е. матричные элементы возмущения должны быть много меньше, чем расстояния между соответствующими уровнями. Согласно (2.50) и (2.56)

$$\psi_m = \psi_n^0 + \sum'_m \frac{V_{mn}}{\mathcal{E}_n^{(0)} - \mathcal{E}_m^0} \psi_m^0. \quad (2.58)$$

Для того чтобы найти поправку второго порядка к энергии, сохраняем в (2.52) члены второго порядка малости; с учетом (2.49) получаем

$$\mathcal{E}_n^{(2)} a_n^0 = \sum_m V_{nm} a_m^2, \quad (2.59)$$

откуда согласно (2.56)

$$\mathcal{E}_n^{(2)} = \sum'_m \frac{(V_{mn})^2}{\mathcal{E}_n^{(0)} - \mathcal{E}_m^{(0)}}. \quad (2.60)$$

Рассмотрим теперь кратко случай, когда невозмущенное состояние имеет вырожденные собственные функции ψ_n^0 , $\psi_{n'}^0$, $\psi_{n''}^0$, ..., соответствующие одному и тому же значению энергии $\mathcal{E}_n^{(0)}$. Как мы уже упоминали, любая комбинация этих функций также будет решением этого уравнения, соответствующим тому же уровню энергии.

В этом случае действие возмущения приводит в первую очередь к двум следствиям:

- 1) снимается полностью (или частично) вырождение,
- 2) стационарным решением уравнения Шредингера становятся не любые линейные комбинации вырожденных функций $\psi_{n'}$, $\psi_{n''}$, $\psi_{n'''} \dots$, а такие, изменения которых под действием возмущения малы.

Подставляя в (2.52) последовательно n , n' , n'' и т. д. и ограничиваясь первым приближением в энергии $\mathcal{E}_n =$

*) Штрих при знаке суммы (Σ') означает, что суммирование проводится по всем m , кроме $m=n$.

— $\mathcal{E}_n^0 + \mathcal{E}_n^1$ и нулевым в коэффициентах разложения, получим

$$\mathcal{E}_n^{(n)} a_n^{(0)} = \sum V_{nn'} a_n^{(0)} \quad (2.61)$$

или

$$\sum_{n'} (V_{nn'} - \mathcal{E}^{(1)} \delta_{nn'}) a_n^{(0)} = 0. \quad (2.62)$$

Мы имеем систему однородных уравнений; условием совместимости их является равенство нулю определителя:

$$|V_{nn'} - \mathcal{E}^{(1)} \delta_{nn'}| = 0. \quad (2.63)$$

Таким образом, раскрыв (2.63), мы получаем уравнение k -й степени (где k — кратность вырождения исходного уровня) относительно $\mathcal{E}^{(1)}$, которое имеет k корней; если все корни разные, это значит, что возмущение сняло вырождение полностью, если есть среди них одинаковые, значит частично.

Подставляя эти значения энергии в (2.62), получаем значения коэффициентов $a_n^0, a_{n'}^0$ и т. д.

Нестационарная теория возмущений. Мы упоминали выше, что при наличии вырождения возмущение приводит к тому, что стационарными состояниями становятся не любые комбинации вырожденных состояний, а лишь такие, на которые возмущение действует слабо.

Если же исходное состояние системы не удовлетворяет этому требованию, то под действием возмущения начнется «перемешивание», т. е. переходы в другие состояния с той же энергией.

Если же возмущение периодически с частотой ω зависит от времени, то под действием его возможны переходы между состояниями с различной энергией, удовлетворяющие резонансному условию Бора:

$$\hbar\omega = \mathcal{E}_m - \mathcal{E}_n. \quad (2.64)$$

В обоих упомянутых выше случаях возмущение является причиной переходов; чтобы получить вероятность таких переходов, мы должны исходить из уравнения Шредингера в виде (2.34а):

$$H\psi = i\hbar \frac{\partial \psi}{\partial t}. \quad (2.65)$$

Будем, так же, как выше, считать, что оператор Гамильтона можно разбить на две части:

$$H = H_0 + V, \quad (2.66)$$

где H_0 — оператор, не зависящий от времени, для которого стационарное уравнение Шредингера

$$H_0 \psi^0 = E_0 \psi^0 \quad (2.67)$$

имеет точные решения: $\psi^0 \dots \psi_k^0 \dots$; V — небольшая добавка, которая может зависеть от времени.

Будем искать опять-таки решение уравнения (2.65) в виде разложения по невозмущенным функциям:

$$\Psi = S a_h \Psi_h^0, \quad (2.68)$$

где символ S означает суммирование по дискретному и интегрирование по непрерывному спектру и Ψ_h^0 включает в себя временной множитель

$$\Psi_h^0 = \psi_k^0 \exp \left(-\frac{i \mathcal{E}_k}{\hbar} t \right) \quad (2.69)$$

и $a_h(t)$ — коэффициенты, зависящие от времени. Подставляя (2.68) в (2.65) и учитывая, что Ψ_h^0 являются решениями невозмущенного уравнения, получим после умножения на Ψ_m^* и интегрирования по пространству

$$\frac{i \hbar da_m}{dt} = S V_{mk}(t) dk, \quad (2.70)$$

где

$$V_{mk}(t) = \int \Psi_m^* V \Psi_k^0 dv. \quad (2.71)$$

Соотношение (2.62) является точным и исходным для большого раздела квантовой механики, так называемой теории переходов. Мы же ограничимся здесь тем, что рассмотрим несколько частных случаев, которые понадобятся нам в дальнейшем.

Воспользуемся, как и выше, методами теории возмущений, т. е. представим коэффициенты $a_m(t)$ в виде разложения по порядку малости

$$a_m = a_m^0 + a_m^{(1)} + a_m^{(2)} + \dots \quad (2.72)$$

Подставляя (2.72) в (2.70), считая при этом матричные коэффициенты возмущения величинами первого по-

рядка малости и приравнивая в левой и правой частях (2.70) члены одинакового порядка малости, получим

$$\frac{da_m^{(0)}}{dt} = 0; \quad \frac{da_m^{(s+1)}}{dt} = (i\hbar)^{-1} SV_{mk} a_k^{(s)} e^{i\omega_{mk}t}, \quad (2.73)$$

где

$$\omega_{mk} = \frac{\mathcal{E}_m - \mathcal{E}_k}{\hbar} \text{ и } s = 1, 2, \dots \quad (2.74)$$

Интегрируя последовательно уравнения (2.73), можно в принципе получить решение с любой степенью точности; в подавляющем большинстве случаев оказывается достаточным первое или второе приближение. Мы здесь ограничимся рассмотрением приближения первого порядка.

Постоянное или «почти» постоянное возмущение. Первое из уравнений (2.73) показывает, что коэффициенты нулевого порядка не зависят от времени — их значения должны быть заданы начальными условиями.

Предположим, что в начальный момент система находилась в одном определенном состоянии: $a_k^0 = \delta_{km}$, если k принадлежит к дискретному спектру, или $\delta_{(k-m)}$, если к непрерывному. Интегрируя уравнения первого порядка, получим

$$a_k^1 = (i\hbar)^{-1} \int_{-\infty}^t \omega_{km} e^{i\omega_{km}t} dt. \quad (2.75)$$

Если возмущение действует в течение ограниченного промежутка времени τ , удовлетворяющего условию $\tau \gg \gg 1/\omega_{km}$, то его можно разложить в ряд Фурье; при этом соотношение (2.75) показывает, что коэффициенты $a_k^1(t)$ при $k = m$ будут пропорциональны компоненте разложения ряда Фурье, для которой выполняется условие Бора,*), остальные будут быстро осциллировать и в среднем по времени дадут величину, близкую к нулю.

Выражение (2.75) легко интегрируется, когда возмущение в течение этого времени постоянно; в этом случае получаем

$$a_k^1(t) = -\frac{V_{km}}{\hbar} \frac{e^{i\omega_{km}\tau} - 1}{\omega_{km}^2} \quad \text{при } t > \tau. \quad (2.76)$$

*) При этом, если нулевая гармоника в ряду Фурье (т. е. постоянная составляющая возмущения) $V^0 \neq 0$, то главный максимум будет соответствовать изоэнергетическим переходам (см. ниже).

Вероятность перехода ω_{km} из m -го начального состояния за время действия возмущения (t) в конечное k -е состояние будет пропорциональна $|a_k^1|^2$, таким образом, согласно (2.76) получаем

$$\omega_{mk} = \frac{4 (V_{km})^2 \sin^2 (\omega_{km} \tau / 2)}{\hbar^2 \omega_{km}^2}. \quad (2.77)$$

Как мы уже упоминали, главный максимум зависимости функции (2.69) от ω_{km} соответствует $\omega_{km} = 0$; высота его пропорциональна τ^2 , а его полная площадь — τ . Резонанс может быть соблюден не точно, если либо начальное, либо конечное состояние принадлежит к непрерывному спектру *). Таким образом, в этом случае вероятность перехода пропорциональна времени действия возмущения и мы можем вычислить вероятность перехода в единицу времени. Полагая для определенности, что конечное состояние размыто, причем характеризуется плотностью состояний $\rho(k)$, получим для вероятности перехода за одну секунду [поделив (2.77) на τ и проинтегрировав в пределах первого максимума]

$$\omega = t^{-1} \int |a_k^1(t)|^2 \rho(k) d\mathcal{E}_k. \quad (2.78)$$

Если $\rho(k)$ является слабо меняющейся функцией, то оно тоже может быть вынесено из-под интеграла и окончательно получаем

$$\omega = \frac{2\pi}{\hbar} \rho(k) |V_{km}|^2. \quad (2.79)$$

вер-ть перехода в ед. времени.

Гармоническое возмущение. Предположим, что возмущение также действует в течение ограниченного промежутка времени τ , но подчиняется гармоническому закону $V(t) = V_0 \sin \omega t$; тогда согласно (2.76) *можно это сделать*

$$a_k^1(t) = \frac{V_{km}^0}{2i\hbar} \left[\frac{e^{i(\omega_{km} + \omega)t} - 1}{\omega_{km} + \omega} - \frac{e^{i(\omega_{km} - \omega)t} - 1}{\omega_{km} - \omega} \right]. \quad (2.80)$$

Выражение (2.80) показывает, что $a_k^1(t)$ будет достигать больших значений (даже при том условии, что V_{km} — маленькая величина), если знаменатель одного из выраже-

*) То, что ширина максимума должна быть обратно пропорциональна τ , следует также из принципа неопределенности $\Delta \mathcal{E} \tau \geq \hbar$. Нетрудно убедиться, что полученные соотношения удовлетворяют этому условию.

ний в скобках близок к нулю. Таким образом, в этом случае условие $\mathcal{E}_k = \mathcal{E}_m$, полученное для постоянного возмущения, заменяется условием Бора:

$$\mathcal{E}_k \approx \mathcal{E}_m \pm h\omega, \quad (2.81)$$

где знак плюс соответствует поглощению кванта энергии, а знак минус — вынужденному (стимулированному) испусканию.

Впервые на возможность испускания лучистой энергии под влиянием поля указал Эйнштейн. Рассмотрев равновесие атомов твердого тела с излучением черного тела Эйнштейн показал также, что вероятность такого индуцированного излучения должна быть равна вероятности поглощения, связанного с переходами между теми же уровнями. На индуцированном излучении основано действие генераторов когерентного излучения (см. гл. 9).

Уравнение Шредингера для системы частиц. Выше мы написали уравнение Шредингера для одной частицы, находящейся в стационарных (2.34) и в нестационарных условиях (2.39). Эти уравнения легко написать (но несоизмеримо трудней решить) для системы частиц. В этом случае оператор кинетической энергии частицы $T = -\hbar^2 \Delta / 2m$ (где Δ — оператор Лапласа, $\Delta = \partial^2 / \partial x^2 + \partial^2 / \partial y^2 + \partial^2 / \partial z^2$) должен быть заменен оператором кинетической энергии системы частиц:

$$T = - \sum_i \frac{\hbar^2}{2m_i} \Delta_i = - \sum_i \frac{\hbar^2}{2m_i} \left(\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right), \quad (2.82)$$

где суммирование производится по всем частицам и вместо потенциальной энергии одной частицы в (2.34) должна быть подставлена потенциальная энергия всей системы частиц. Таким образом, и (2.38), и (2.34а) внешне сохраняют тот же вид:

$$H\Psi = \mathcal{E}\Psi \quad (2.83)$$

и

$$H\Psi = i\hbar \frac{\partial \Psi}{\partial t}, \quad (2.84)$$

но со следующими существенными отличиями. В данном случае H — оператор Гамильтона системы:

$$H = - \sum_i \frac{\hbar^2}{2m_i} \Delta_i + U,$$

где U — потенциальная энергия всей системы;

\mathcal{E} — полная энергия системы и

Ψ — волновая функция координат всех частиц;

$\Psi = \Psi(x_1, y_1, z_1, x_2, y_2, z_2, \dots)$.

В том случае, когда потенциальную энергию системы можно представить в виде суммы потенциальных энергий отдельных частиц:

$$U = \sum U_i, \quad (2.85)$$

т. е. когда частицы не взаимодействуют друг с другом или их взаимодействием в первом приближении можно пренебречь, или взаимодействие можно учесть приближенно введением эквивалентного (самосогласованного) поля (см. гл. 4), уравнение (2.83) может быть удовлетворено простым решением Ψ , представляющим произведение волновых функций отдельных частиц:

$$\Psi = \prod_i \psi_i(x_i, y_i, z_i), \quad (2.86)$$

при этом уравнение (2.83) распадается на i независимых уравнений для каждой отдельной частицы.

ХИМИЧЕСКАЯ СВЯЗЬ

Ионная связь. Мы уже упоминали в предыдущей главе, что энергетически наиболее устойчивыми являются электронные структуры благородных газов, и атомы, вступая в химическое взаимодействие, стремятся приблизить свою структуру к ближайшему благородному газу. Поэтому атомы элементов, стоящих в левой части таблицы Менделеева, стремятся отдать свои «лишние» (по сравнению с ближайшим благородным газом) электроны, а атомы, стоящие в правой части таблицы, — дополнить число электронов до ближайшего благородного газа. В результате этого образуются положительно и отрицательно заряженные ионы, имеющие замкнутые электронные оболочки и связанные друг с другом кулоновскими силами.

Можно сказать, что если каждому атому, входящему в соединение, можно приписать вполне определенную структуру, не зависящую от наличия других атомов и при этом атомы имеют отличный от нуля электрический заряд и электростатические силы дают основной вклад в энергию образования соединения, то мы имеем дело с ионной связью.

Ковалентная связь. Вторым способом, которым атом может, вступая в соединение, приблизить свою электронную структуру к структуре благородного газа,— это спаривание электронных орбит с соседними. Так, например, каждый атом углерода в кристалле алмаза находится в центре тетраэдра, в вершинах которого находятся такие же атомы; спаривая свои четыре электрона с четырьмя соседними, каждый атом дополняет число своих электронов до восьми и таким образом образует оболочку, в известном смысле подобную атому гелия. При этом коренное отличие заключается в том, что эту структуру ни в какой мере нельзя считать замкнутой и не зависящей от соседей: все атомы переплетаются здесь в одну гигантскую молекулу.

Металлическая связь обычно образуется в тех случаях, когда число валентных электронов настолько мало, что ни при какой координации (даже самой плотной упаковке) не могут образоваться замкнутые оболочки. В этом случае атомы, как правило, стремятся обеспечить себе наибольшее число соседей в первой и второй координационных зонах.

Здесь следует заметить, что между ковалентными и ионными соединениями и ковалентными соединениями и металлами не существует резкой границы. Действительно, представим себе, что один из атомов, образующих бинарное соединение, отдал второму некоторое количество электронов и таким образом образовалась ионная молекула. Но при этом электронная оболочка отрицательного иона будет деформироваться под действием поля положительного иона и центр тяжести ее (оболочки) будет смещаться по направлению к последнему. Если при этом поляризуемость отрицательного иона велика и его оболочка деформируется настолько, что захватит положительный ион, связь уже будет носить частично ковалентный характер. Точно так же ковалентная связь в большинстве случаев бывает частично ионной по следующим причинам:

1) если у одного из элементов, образующих ковалентное соединение, электронное сродство больше, то центр тяжести электронного заряда будет смещен по направлению к этому элементу и поэтому связь приобретает частично ионный характер;

2) если для образования ковалентных связей один из атомов передал некоторое число электронов второму *).

Наконец, даже если связываются два совершенно одинаковых атома (например, в молекуле водорода H_2), то всегда возникает вопрос, какая связь энергетически выгодней: ковалентная, когда оба электрона движутся по орбите, охватывающей оба атома, или ионная, когда оба электрона находятся вместе на одном из атомов (попеременно), а второй атом (положительный ион) притягивается кулоновскими силами. Если энергии этих состояний, ковалентного и ионного, сравнимы, то даже в этом случае связь будет носить частично ионный характер.

Чем тяжелее атомы, образующие ковалентный кристалл, тем меньше их потенциал ионизации и больше поляризуемость и тем менее направленный, более размытый характер носят ковалентные связи. Можно сказать, что при переходе к более тяжелым атомам происходит «металлизация» ковалентных связей.

В металлах валентные электроны коллективизируются и получают возможность перемещаться свободно по всему кристаллу, но и в этом случае свободные электроны также не распределены совершенно равномерно по всему объему, а движутся преимущественно по «мостикам», соединяющим соседние атомы, и с этой точки зрения между металлической и двумя упомянутыми выше типами связи также нет резкой грани.

Однако коренное отличие металлов заключается в том, что в этом случае мы имеем дело с нецеликом заполненными оболочками-мостиками; поэтому электроны могут беспрепятственно перескакивать с одного мостика на другой под действием электрического поля, чем, грубо говоря, и обусловлена металлическая проводимость. Незаполненностью орбит в металлах и размытостью связей обусловлены и другие их особенности — большие координационные числа, механические, оптические и другие свойства.

Дисперсионные силы связи. Наряду с рассмотренными выше типами кристаллов и характером связей существует ряд кристаллов, построенных из замкнутых нейтральных молекул; внутримолекулярные силы в них носят в большинстве случаев ковалентный характер (сюда, в частно-

*) Такую связь иногда химики называют донорно-акцепторной или координационной.

сти, относятся кристаллы благородных газов, азота и кислорода, построенные соответственно из молекул N_2 и O_2 , кристаллы органических веществ и др.), но иногда и частично ионный (NO , HCl и др.). В силу замкнутости и нейтральности таких молекул ковалентное и ионное взаимодействие между ними отсутствует, и здесь на первый план выступает более слабое дипольное взаимодействие. При этом следует различать два случая: а) взаимодействие полярных молекул с готовыми диполями (так называемое ориентационное взаимодействие) и б) взаимодействие симметричных молекул, в которых дипольные моменты возникают за счет взаимной поляризации (силы Ван-дер-Ваальса). Энергия взаимодействия готовых диполей убывает как куб расстояния, ван-дер-ваальсова убывает пропорционально шестой степени. Таким образом, первый тип связи является более дальнедействующим.

На первый взгляд может показаться, что первый тип связи (ориентационный) является более прочным, однако это не так. Объясняется это большой зависимостью координационного взаимодействия от направления: полярные молекулы, помещенные в решетку, не могут быть все «благоприятно» ориентированы по отношению друг к другу. Взаимная же поляризация может носить, во-первых, синхронный во времени, а во-вторых, более сложный (мультипольный) характер и быть одновременно благоприятной (с точки зрения минимума свободной энергии) для большого числа молекул, находящихся в сфере взаимного влияния.

Дипольное взаимодействие по сравнению с упомянутыми выше является более слабым; однако нетрудно убедиться, что между ними и другими типами связи также нет резкой грани; его частично ионный характер очевиден. С другой стороны, можно представить, как взаимная поляризация молекул по мере своего усиления переходит в ковалентную связь.

Следует также отметить, что точный подсчет энергии связи и ионных и ковалентных кристаллов требует учета дополнительной энергии, возникающей за счет поляриционного (ван-дер-ваальсового) взаимодействия.

Характер связи в молекуле и кристалле определяется, в первую очередь, отношением элементов, из которых они (т. е. молекулы и кристаллы) построены, к электронам, т. е. потенциалом ионизации и электронным средством.

Качественные соображения и расчет в соответствии с опытом показывают, что металлическую связь образуют элементы с небольшим потенциалом ионизации. Ионная связь образуется, когда в соединение вступают элементы, один из которых имеет большое электронное сродство, а другой — низкий потенциал ионизации; ковалентную связь образуют элементы с достаточно большим и не слишком отличающимся друг от друга электронным сродством.

Наконец, ван-дер-ваальсовы кристаллы образуют очень прочные молекулы и атомы благородных газов. Различным типам связи соответствует также различная координация атомов и в соответствии с этим различные типы структуры кристаллов. В валентных кристаллах электроны распределены между соседними атомами на валентных связях. Число связей, которое может образовываться в кристалле, определяется правилом

$$N = 8 - \bar{n}, \quad (2.87)$$

где \bar{n} — среднее число валентных электронов на атом.

Таким образом, в то время как в металлах, ионных и ван-дер-ваальсовых кристаллах преобладают структуры, обеспечивающие большое число ближайших соседей, координационное число в валентных кристаллах должно определяться формулой (2.87). Для теллура и селена это число (N) равно 2, для висмута и сурьмы — 3, для элементов четвертой группы и соединений типа $A^{III}B^V$ $n = N = 4$, что в точности соответствует структуре кристаллов этих элементов и соединений.

2.2. ГЕОМЕТРИЯ КРИСТАЛЛИЧЕСКОЙ РЕШЕТКИ

ПРЕОБРАЗОВАНИЯ АСИММЕТРИИ ТВЕРДОГО ТЕЛА

Понятие о симметрии вытекает из обобщения понятия о движении твердого тела (или фигуры). Определим движение тела как такую операцию, при которой расстояния между любыми двумя точками твердого тела остаются неизменными. При таком определении движения в него входят *обычные физические перемещения*:

— повороты вокруг точки и оси;

— трансляции, т. е. поступательное перемещение, при котором все точки тела описывают одинаковые траектории;

— винтовое движение, представляющее собой комбинацию вращения вокруг оси (не обязательно проходящей внутри тела) и перемещения вдоль оси.

Можно показать, что любое перемещение твердого тела можно получить путем винтового движения, которое,

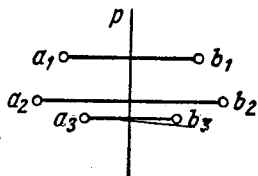


Рис. 2.3. Зеркальное отражение в плоскости P.

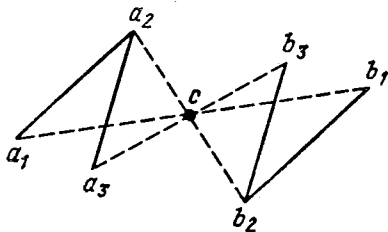


Рис. 2.4. Инверсия в точке c.

таким образом, является наиболее общим видом физического движения тела.

При данном выше определении движения в него включаются также *другие виды преобразований*:

— зеркальное отражение в плоскости, при помощи которого можно превратить правую фигуру (например, руку) в левую (рис. 2.3);

— обращение или инверсия (отражение в точке), т. е. такое преобразование, при котором все координаты данной точки относительно центра инверсии, выбранного за начало координат, меняют знак (рис. 2.4): $x_i = -x$, $y_i = -y$ и $z_i = -z$, т. е. $r_i = -r^*$).

Наиболее общий вид движения — комбинация отражения и инверсии со всеми видами физического движения. Простейшим видом обобщенного движения является отражение; покажем, что все другие виды движения можно получить путем комбинации последовательных отражений в ряде зеркал;

— трансляция фигуры на вектор \mathbf{a} эквивалентна двум последовательным отражениям в двух параллельных зеркалах, перпендикулярных вектору \mathbf{a} и расположенных на расстоянии $\mathbf{a}/2$;

*) Операция зеркального отражения эквивалентна отражению в обычном плоском зеркале, операция инверсии — действительному изображению тела, помещенного на двойном фокусном расстоянии выпуклой линзы.

— поворот на угол α эквивалентен последовательному отражению в двух зеркалах, пересекающихся по оси поворота и расположенных под углом $\alpha/2$;

— операция инверсии эквивалентна последовательному отражению в трех взаимно перпендикулярных зеркалах, пересекающихся в центре инверсии.

Введенных выше понятий достаточно, чтобы перейти непосредственно к теме настоящего раздела — к определению преобразования симметрии.

Преобразованиями симметрии называются такие движения твердого тела или фигуры (трансляция, повороты, отражения и инверсия), в результате которых тело (или геометрическая фигура) совмещается с самим собой.

Каждое из таких движений называется элементом симметрии данного тела. Совокупность всех элементов симметрии данного тела называется группой симметрии. Группы симметрии, содержащие операции отражения, поворота и инверсии, называются точечными, так как при них остается на месте по крайней мере одна точка тела или фигуры. Точечные группы симметрии характерны для конечных фигур атомов, молекул и многогранников и бесконечных анизотропных сплошных сред.

Группы симметрии, включающие в себя также и трансляции, называются соответственно линейными, плоскими и пространственными; эти группы характерны для периодических структур.

Как следует из сказанного выше, последовательное применение двух преобразований симметрии дает третье преобразование. Так, например, трансляция вдоль оси и поворот дают винтовое движение, два отражения в зависимости от расположения плоскостей дают поворот или трансляцию, трансляции размножают точечные элементы симметрии и т. д. Поэтому законченной группой симметрии данного тела (в дальнейшем мы будем говорить кристалла) является полная совокупность элементов симметрии, которая получается путем применения комбинации любых его элементов.

КРИСТАЛЛИЧЕСКИЕ КЛАССЫ

Все кристаллы, внешняя форма которых одинакова, т. е. которые могут быть совмещены одинаковыми поворотами, отражениями и инверсиями, образуют один кристаллический класс. Как следует из сказанного выше, эти

кристаллы должны иметь точечную симметрию одного и того же вида, т. е. входить в одну точечную группу симметрии.

Введем обозначения для элементов симметрии: плоскость симметрии обозначим буквой P или m ; ось симметрии — буквой Z или g . Например, Z^4 или g^4 будет означать, что кристалл обладает осью симметрии четвертого порядка, т. е. может быть совмещен с самим собой путем поворота на 90° .

Зеркально поворотной осью будем называть одновременный поворот вокруг оси и отражение от перпендикулярной к ней плоскости; например, Z_6^3 означает ось симметрии третьего порядка и зеркально поворотную ось шестого порядка. Центр инверсии будем обозначать буквой c . На первый взгляд может показаться, что число элементов симметрии, например поворотных осей (Z^1, Z^2, \dots, Z^n) и их комбинаций, может быть бесконечно, поэтому должно быть бесконечно и число кристаллических классов. В действительности это неверно: число точечных групп симметрии и равное ему число кристаллических классов всего 32.

Вывод всех видов симметрии был сделан Гесселем в 1836 г., но остался не замеченным в мировой литературе, и затем, независимо, русским ученым А. В. Гадолиным в 1876 г.; этот вывод приводится во всех учебниках кристаллографии, мы не будем здесь повторять его, а приведем лишь несколько положений, иллюстрирующих конечность числа групп симметрии.

1. В кристалле могут присутствовать только оси симметрии Z^1, Z^2, Z^3, Z^4 и Z^6 (это следует из периодической структуры кристалла), плоскость можно заполнить только правильными треугольниками, четырехугольниками и шестиугольниками. Таким образом, возникает первый ряд — 5 видов симметрии, содержащих только одну ось.

2. Плоскость симметрии может быть добавлена либо перпендикулярно к оси, либо параллельно (если она будет направлена под углом, это приведет к появлению в результате отражения от нее новых осей, а мы этот случай пока отбрасываем). Плоскость, параллельная оси, дает следующий ряд видов симметрии:

$$Z^1P = P, Z^2Pc; Z^3Pc, Z^4P \text{ и } Z^6Pc *).$$

*) Следует обратить внимание, что любая из этих пяти групп симметрии является замкнутой, т. е. любое последовательное применение преобразований, входящих в группу, дает одно из уже имеющихся преобразований.

3. Прибавление плоскости, перпендикулярной оси, дает еще один ряд видов (или групп) симметрии:

$$Z^1P = P, Z^22P, Z^33P, Z^44P, Z^66P$$

(в этом случае каждая плоскость размножается соответственно симметрии оси в 2, 3, 4 и 6 плоскостей).

4. Можно вносить одновременно параллельную и перпендикулярную оси плоскости — это дает ряд, содержащий пять групп.

5. Шестой ряд образует три вида кристаллов, содержащих только зеркально-поворотные оси: Z_2^1 , Z_4^3 и Z_6^3 .

6. Анализ случая кристаллов, содержащих более одной поворотной оси, представляет наибольшие трудности; мы здесь лишь укажем, что требование, чтобы в результате своего сочетания элементы симметрии не размножались бесконечно (и не приводили таким образом к шаровой симметрии), приводит к ограниченному числу таких групп — всего 13.

Таким образом, общее число точечных групп оказывается 32.

ПРОСТРАНСТВЕННЫЕ ГРУППЫ СИММЕТРИИ

Вывод всех пространственных групп, которые получаются путем присоединения к точечным группам еще одной группы элементов симметрии — подгруппы трансляций по трем осям — значительно сложнее; он был сделан впервые лишь в 1890 г. русским ученым Е. С. Федоровым, который получил общее число пространственных групп равным 230. 230 федоровских групп описывают симметрию скалярного поля, величина которого в каждой точке описывается одним положительным числом (например, плотностью распределения электронного заряда или массы и т. д.).

А. В. Шубников обобщил теорию Федорова на черные поля, в которых изменяющаяся величина может менять знак. Дальнейшее обобщение этих представлений приводит к цветовым полям, в которых каждая точка может иметь l цветов, что соответствует, например, l -дискретным ориентациям вектора, расположенного в данной точке. Таким образом, теория симметрии обобщается на векторные поля. Теория А. В. Шубникова привела к введению новых понятий антиравенства (т. е. равенства с точностью до знака) и антисимметрии. Советские ученые вывели все точечные

и пространственные группы с обобщением на изменение цвета. При этом наряду с 32 обычными точечными группами было получено еще 58 черно-белых групп. Общее число пространственных (цветовых) групп возросло до 1651. Классификация симметрии векторных полей находит свое применение в теории магнитных и диэлектрических структур.

ЭЛЕМЕНТАРНАЯ ЯЧЕЙКА

Любую линейную периодическую структуру можно получить путем применения к исходной точке одной элементарной трансляции. Такая цепочка будет обладать, кроме трансляционной, бесконечным множеством элементов симметрии, так как ее ось является осью симметрии по отношению к повороту на любой угол, т. е. осью симметрии бесконечного порядка.

Симметрия цепочки не уменьшится, если мы в каждую из точек поместим сферу или диск, перпендикулярный оси. Если же в каждую из точек одномерной решетки поместить правильный n -угольник, перпендикулярный оси, то ось симметрии уже будет иметь порядок n . Если мы сделаем ось цепочки осью симметрии второго порядка, то наша одномерная решетка может обладать еще значительным числом элементов симметрии. Это иллюстрируется плоскими орнаментами, приведенными на рис. 2.5, которые образуют 5 точечных групп симметрии (и 7 пространственных).

Приведенные примеры иллюстрируют тот факт, что даже в простейшем случае одномерной решетки распределение вещества в ее элементарной ячейке определяет ее симметрию.

Аналогичные представления можно распространить на плоскую и объемную решетки. Плоская решетка определяется двумя элементарными трансляциями, на основе которых можно построить параллелограмм, который называется элементарной ячейкой.

Как видно из рис. 2.6, выбор элементарных трансляций не однозначен, поэтому элементарные трансляции, которые также называются базисными векторами решетки, выбираются всегда так, чтобы элементарная ячейка обладала максимальным числом элементов симметрии и таким образом представляла точечную группу симметрии всей решетки. На рис. 2.6 элементарная ячейка обведена сплошной

линией. Трехмерная решетка определяется тремя некопланарными трансляциями; как и в случае плоской решетки, они выбираются таким образом, чтобы элементарная ячейка

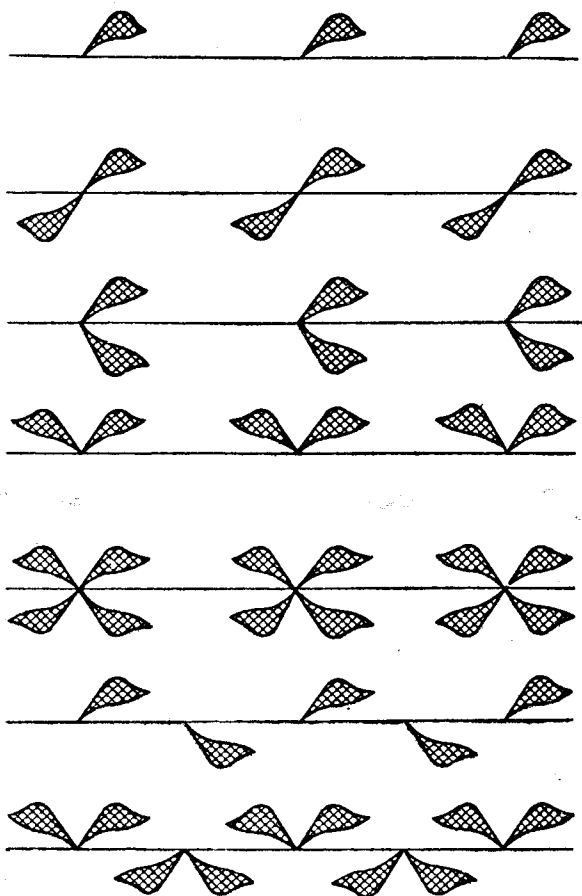


Рис. 2.5. Плоские орнаменты, иллюстрирующие 5 точечных и 7 пространственных групп линейной решетки.

обладала максимальной симметрией. В зависимости от свойств и взаимного расположения атомов решетки подразделяются на простые и сложные. Простой или примитивной называется решетка, в которой все атомы эквива-

лентны (т. е. одинаковы и одинаково расположены). На рис. 2.7, *а* представлена простая линейная решетка. На рис. 2.7, *б* и *в* представлены сложные линейные решетки.

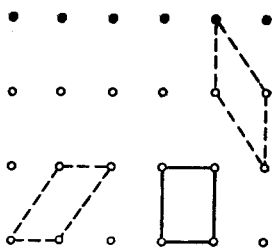


Рис. 2.6. Иллюстрация неоднозначности выбора элементарных трансляций в плоской решетке.

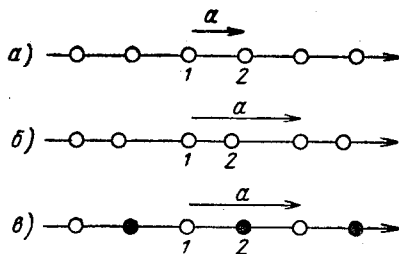


Рис. 2.7. Линейные решетки: *а* — простая; *б*, *в* — сложные.

ки — в них положение атомов 1 и 2 неэквивалентно. На рис. 2.8 представлены простая (*а*) и сложная (*б*) плоские решетки.

Из сказанного следует, что элементарная ячейка простой решетки содержит всего один атом. В приведенных

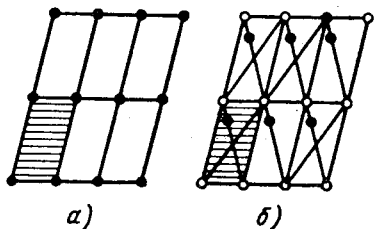


Рис. 2.8. Плоские решетки: *а* — простая; *б* — сложная (в обоих случаях заштрихована элементарная ячейка).

выше двух примерах сложной решетки элементарная ячейка содержала всего два атома. В более общем случае сложная решетка может содержать любое число атомов в элементарной ячейке и при этом она будет определяться элементарными трансляциями плюс координатами всех атомов в элементарной ячейке. Сложную решетку можно представить в виде совокупности двух или большего числа простых подрешеток, смещенных друг относительно друга, и задать при этом относительные смещения подрешеток.

Выше мы упоминали, что базисные векторы и элементарная ячейка выбираются таким образом, чтобы они отражали точечную группу симметрии кристалла; по внешнему виду (ограничению), т. е. по взаимному расположению и соотношению базисных векторов, кристаллы объединяются по сингониям.

Сингонией, таким образом, называется совокупность групп (или, что то же самое, видов) симметрии, т. е. классов кристаллов, обладающих сходными элементарными ячейками. Всего существует семь сингоний (систем), определение которых приведено ниже:*)

- 1) кубическая $a = b = c$, $\alpha = \beta = \gamma = 90^\circ$;
- 2) тетрагональная $a = b \neq c$, $\alpha = \beta = \gamma = 90^\circ$;
- 3) ромбическая $a \neq b \neq c$, $\alpha = \beta = \gamma = 90^\circ$;
- 4) моноклинная $a \neq b \neq c$, $\alpha = \gamma = 90^\circ \neq \beta$;
- 5) тригональная $a = b = c$, $\alpha = \beta = \gamma \neq 90^\circ$;
- 6) триклинная (ромбоэдрическая) $a \neq b \neq c$, $\alpha \neq \beta \neq \gamma \neq 90^\circ$;
- 7) гексагональная $a = b \neq c$, $\alpha = 120^\circ$, $\beta = \gamma = 90^\circ$.

Согласно сказанному выше сингония определяется внешним видом (формой) кристалла или, если можно так выразиться, внешним видом элементарной ячейки, но совершенно не зависит от расположения атомов внутри элементарной ячейки. Поэтому в каждой сингонии объединяется целый ряд групп точечной симметрии. С другой стороны, так как тип сингонии никак не связан с распределением атомов по элементарной ячейке, то он не отражает полностью трансляционную симметрию кристалла. Расположением атомов внутри элементарной ячейки можно изменить трансляционную симметрию кристалла. Поэтому каждый из типов сингонии объединяет в себе также ряд кристаллов с различной трансляционной симметрией. Простейшими представителями каждой сингонии являются примитивные решетки с одним атомом в каждой элементарной ячейке.

Наряду с семью типами простых решеток, соответствующих семи сингониям, путем центрировки граней и объема элементарных ячеек можно образовать еще 7 решеток, которые также могут быть получены путем элементарных трансляций, но уже других, если пожертвовать при этом точечной симметрией.

*) Здесь a, b, c —базисные векторы, а α, β, γ —углы между ними.

Таким образом, если классифицировать решетки по подгруппам трансляционной симметрии (независимо от подгруппы вращений), то всего имеется 14 типов различных

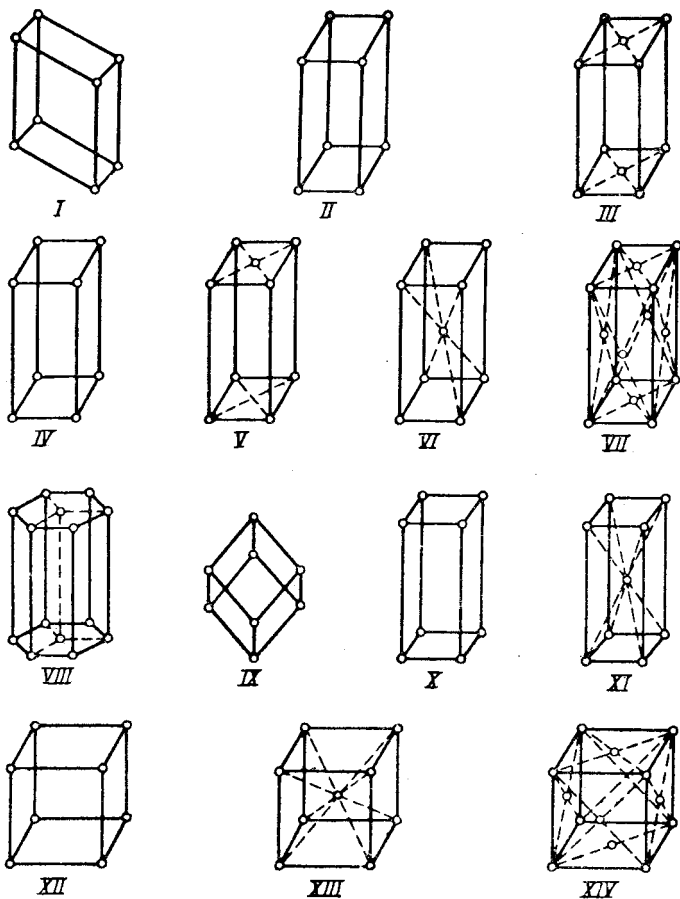


Рис. 2.9. Решетки Бравэ.

решеток (рис. 2.9), называемых решетками Бравэ (по имени кристаллографа, который вывел эти типы решеток в середине прошлого века).

Следует еще раз подчеркнуть, что классификация кристаллов по сингониям не учитывает периодической струк-

туры кристалла и в равной мере может быть отнесена к анизотропным средам. Здесь размеры элементарных трансляций имеют лишь относительное значение: все они могут быть бесконечно малыми или бесконечно большими. Их взаимное отношение и расположение характеризуют лишь анизотропию пространства, поэтому по сингониям (так же как по точечным группам) можно классифицировать либо конечные ограниченные кристаллы, либо бесконечный анизотропный континуум.

Решетки Бравэ учитывают дискретную структуру вещества, т. е. трансляционную симметрию, поэтому некоторые из сингоний (систем) разбиваются на несколько решеток в зависимости от расположения атомов в элементарной ячейке. Так, в кубическую систему помимо простой кубической решетки включается объемно-центрированная, содержащая два атома в элементарной ячейке, и грани-центрированная, содержащая в элементарной ячейке четыре атома.

Как мы уже упоминали выше, все решетки Бравэ являются формально простыми, т. е. могут быть получены путем трех элементарных трансляций и, следовательно, элементарная ячейка может быть выбрана таким образом, чтобы она содержала только один атом. Однако такой выбор элементарной ячейки будет неправильным, так как эта ячейка не будет содержать всей точечной группы симметрии кристалла.

МАТЕМАТИЧЕСКОЕ ВЫРАЖЕНИЕ СИММЕТРИИ РЕШЕТКИ

В соответствии со сказанным выше основным свойством кристаллической решетки является ее трансляционная симметрия, характеризуемая тремя базисными векторами: \mathbf{a}_1 , \mathbf{a}_2 и \mathbf{a}_3 . Это значит, что при смещении всего кристалла на вектор

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

где n_1 , n_2 , n_3 — любые целые числа, она совместится с собой. Это должно также относиться и к любой физической величине, определяемой однозначно расположением атомов в объеме: электростатическому потенциалу, электронной плотности и др. Следовательно, эти величины, в частности электростатический потенциал V , также должны обладать

той же трансляционной симметрией, т. е.

$$V(\mathbf{r} + \mathbf{a}_n) = V(\mathbf{r}). \quad (2.88)$$

Это же свойство трансляционной симметрии можно выразить иначе, воспользовавшись преобразованием координат. Если мы перенесем начало координат в новую точку таким образом, что новые координаты \mathbf{R} будут связаны со старыми соотношением

$$\mathbf{r} = \mathbf{R} \pm \mathbf{a}_n, \quad (2.89)$$

и подставим \mathbf{r} , выраженное через \mathbf{R} согласно (2.89), в $V(\mathbf{r})$, то $V(\mathbf{R})$ должно быть тождественно с той же функцией новых переменных:

$$V(\mathbf{R}) \equiv V(\mathbf{r}). \quad (2.90)$$

Это же определение преобразованной симметрии можно распространить на группу вращений. Если в кристалле имеется ось симметрии n -го порядка, то поворот координат на угол $\alpha_n = 2\pi/n$ должен приводить к тождественному преобразованию всех соответствующих функций, т. е.

$$V(X, Y, Z) \equiv V(x_1, y_1, z_1), \quad (2.91)$$

если X, Y, Z и x_1, y_1, z_1 связаны соотношениями:

$$\begin{aligned} X &= x_1 \cos \alpha_n - y_1 \sin \alpha_n, \\ Y &= x_1 \sin \alpha_n + y_1 \cos \alpha_n, \\ Z &= z_1. \end{aligned} \quad (2.92)$$

ОБРАТНОЕ ПРОСТРАНСТВО И ОБРАТНАЯ РЕШЕТКА

Состояние классической частицы характеризуется тремя координатами x, y, z в обычном классическом пространстве. Состояние свободного электрона в квантовой механике, как мы уже упоминали, описывается плоской волной *):

$$\begin{aligned} \psi(\mathbf{r}) &= \sqrt{V} \bar{V} e^{2\pi i \mathbf{k} \cdot \mathbf{r}} = \\ &= \sqrt{L_1 L_2 L_3} \exp[2\pi i (k_x x + k_y y + k_z z)], \end{aligned} \quad (2.93)$$

*) Мы отбрасываем не интересующий нас здесь временный множитель.

где $V = L_1 L_2 L_3$ *) — объем прямоугольного ящика, в котором «заперт» электрон.

Нормирующий множитель перед экспонентой получается из того условия, что вероятность нахождения электрона во всем ящике должна быть равна единице:

$$\int_V \psi(\mathbf{r}) \psi^*(\mathbf{r}) dV = 1. \quad (2.94)$$

Таким образом, состояние электрона целиком характеризуется тремя величинами: k_x , k_y и k_z , которые имеют размерность обратной длины (так как показатель степени должен быть безразмерной величиной). Следовательно, если мы построим пространство с размерностью, обратной длине, то k_x , k_y и k_z будут определять координаты какой-то точки в этом (обратном) пространстве и координаты этой точки будут полностью определять состояние $\psi(\mathbf{r})$ электрона, в то время как координаты его в прямом пространстве неопределенны (с точностью до объема ящика).

Так как вероятность нахождения электрона на границах ящика должна обращаться в нуль:

$$|\psi|^2 = 0 \text{ при } x=0, L_1, y=0, L_2 \text{ и } z=0, L_3,$$

то максимальная длина волны электрона составляет: $\lambda_{\text{макс } x} = 2L_1$, $\lambda_{\text{макс } y} = 2L_2$ и $\lambda_{\text{макс } z} = 2L_3$; соответственно минимальные значения проекций волнового вектора будут: $k_{x \text{ мин}} = 1/2L_1$, $k_{y \text{ мин}} = 1/2L_2$, $k_{z \text{ мин}} = 1/2L_3$. Таким образом, обратное пространство мы можем разбить на ячейки размером $1/2L_1 \cdot 1/2L_2 \cdot 1/2L_3$, и состояние электрона — волны будет характеризоваться точкой в вершине одной из ячеек.

Все сказанное выше об электроне-волне будет относиться к любой периодической функции координат в кристалле. Например, любые колебания решетки можно разложить по трем координатным осям, причем длина волны колебаний λ может изменяться в пределах:

$$\begin{aligned} 2a_1 &\leq \lambda_x \leq 2L_1, \\ 2a_2 &\leq \lambda_y \leq 2L_2, \\ 2a_3 &\leq \lambda_z \leq 2L_3, \end{aligned} \quad (2.95)$$

*) Свободным мы называем электрон в том смысле, что внутри ящика его энергия постоянна и на него не действуют никакие силы.

где a_1 , a_2 и a_3 — постоянная решетки (длина базисных векторов элементарной ячейки) для этих направлений.

Следовательно, волновое число q колебаний решетки может изменяться в пределах:

$$\begin{aligned} \frac{1}{2a_1} &\geq q_x \geq \frac{1}{2L_1}, \\ \frac{1}{2a_2} &\geq q_y \geq \frac{1}{2L_2}, \\ \frac{1}{2a_3} &\geq q_z \geq \frac{1}{2L_3}. \end{aligned} \quad (2.96)$$

Приведенные выше соотношения показывают общие черты и различие электронных и упругих волн; общим у них будет то, что максимальная длина их одинакова (и поэтому размер элементарной ячейки обратного пространства в обоих случаях может быть выбран одним и тем же). Различие — в том, что если минимальная длина волны решетки ограничена выражением (2.95), т. е. атомистической структурой кристалла, и поэтому все возможные значения волнового числа уместятся в «довольно большом», но все же ограниченном ящике обратного пространства (размером $1/a_1 \cdot 1/a_2 \cdot 1/a_3$ *), то размеры волнового вектора электрона на первый взгляд ничем не ограничены.

В действительности же это различие формально может быть уничтожено; как мы увидим в дальнейшем, периодический потенциал кристалла накладывает свой отпечаток и на движение электрона; это, в частности, проявляется в том, что электрон описывается волной не с постоянной, а с модулированной амплитудой:

$$\psi = A(\mathbf{r}) e^{2\pi i \mathbf{k} \cdot \mathbf{r}}, \quad (2.97)$$

где $A(\mathbf{r})$ имеет периодичность, равную периоду решетки (или еще более мелкую), т. е.

$$A(\mathbf{r} + n\mathbf{a}) = A(\mathbf{r}), \quad (2.98)$$

где \mathbf{a} — любой вектор решетки.

Например, если волна направлена вдоль оси x , то

$$\psi = A(x) e^{2\pi i k_x x} \quad (2.99)$$

и

$$A(x + a_1) = A(x).$$

*) Эту часть обратного пространства мы будем называть основной (или приведенной).

Предположим теперь, что $k_x > 1/a_1$, т. е. выходит за пределы основной (приведенной) области. В этом случае мы можем разбить волновой вектор на две части: целую (в единицах $1/a_1 \rightarrow n/a_1$) и дробную $\rightarrow k'_x$:

$$k_x = \frac{n}{a_1} + k'_x, \quad (2.100)$$

где $k'_x < 1/a_1$ и, следовательно, уже уместается в основной области.

Подставляя (2.100) в (2.99), получаем

$$\psi = A(x) e^{\frac{i2\pi n}{a_1} x} e^{i2\pi k'_x x} = A'(x) e^{i2\pi k'_x x}, \quad (2.101)$$

где величина $A'(x)$ также удовлетворяет условию (2.98), т. е. ее можно рассматривать как периодически модулированную амплитуду, а k' уже ограничено теми же условиями, что и q .

Введение обратного пространства оказывается также чрезвычайно плодотворным для анализа самого периодического потенциала решетки, распределения электронной плотности и любой другой периодической функции в прямом пространстве. Действительно, любая такая величина, (а следовательно, и периодический потенциал V) может быть разложена в трехмерный ряд Фурье:

$$V(\mathbf{r}) = \sum V_{n_1 n_2 n_3} \exp \left[2\pi i \left(\frac{n_1}{a_1} x + \frac{n_2}{a_2} y + \frac{n_3}{a_3} z \right) \right]. \quad (2.102)$$

Совокупность чисел n_1 , n_2 и n_3 характеризует точку в обратном пространстве или вектор обратной решетки с компонентами:

$$g_x = \frac{n_1}{a_1}, \quad g_y = \frac{n_2}{a_2} \quad \text{и} \quad g_z = \frac{n_3}{a_3}, \quad (2.103)$$

и разложение (2.102) может быть представлено в виде

$$V(\mathbf{r}) = \sum_{\mathbf{g}} V(\mathbf{g}) e^{2\pi i(\mathbf{g}\mathbf{r})}. \quad (2.104)$$

Таким образом, любая периодическая функция в прямом пространстве может быть охарактеризована совокупностью амплитуд $V(\mathbf{g})$ трехмерного разложения ряда Фурье. Эти амплитуды характеризуют значения функции V в точках \mathbf{g} обратного пространства. Для того чтобы обобщить развитые выше представления на косоугольные системы координат и кристаллические решетки, базисные

векторы которых не ортогональны друг другу, надо базисные векторы обратной решетки определить следующим образом [9]:

$$\mathbf{b}_1 = \frac{[\mathbf{a}_2\mathbf{a}_3]}{a_1[\mathbf{a}_2\mathbf{a}_3]}, \quad \mathbf{b}_2 = \frac{[\mathbf{a}_3\mathbf{a}_1]}{a_1[\mathbf{a}_2\mathbf{a}_3]} \quad \text{и} \quad \mathbf{b}_3 = \frac{[\mathbf{a}_1\mathbf{a}_2]}{a_1[\mathbf{a}_2\mathbf{a}_3]}. \quad (2.105)$$

Если электрон локализован в определенной области $(\Delta x, \Delta y, \Delta z)$ прямого пространства, то он описывается уже не монохроматической волной, а волновым пакетом:

$$\psi(\mathbf{r}) = \int_{\mathbf{k}+\Delta\mathbf{k}}^{\mathbf{k}-\Delta\mathbf{k}} A(\mathbf{k}) e^{2\pi i \mathbf{k} \cdot \mathbf{r}} d k_x d k_y d k_z. \quad (2.106)$$

Как следует из теории интегралов Фурье, Δk_x и Δx , Δk_y и Δy , Δk_z и Δz связаны соотношениями:

$$\Delta k_x \Delta x \geq 1, \quad \Delta k_y \Delta y \geq 1, \quad \Delta k_z \Delta z \geq 1. \quad (2.107)$$

Таким образом, чем более определенной становится координата электрона в прямом пространстве, тем больше расплывается объем, определяющий его в обратном пространстве. Соотношение (2.107) непосредственно связано с принципом неопределенности. Напомним, что импульс электрона связан с его волновым числом соотношениями:

$$p_x = \hbar k_x, \quad p_y = \hbar k_y, \quad p_z = \hbar k_z. \quad (2.108)$$

Из (2.95) и (2.96) следует непосредственно соотношение неопределенности

$$\Delta\Omega \Delta v \geq \hbar^3, \quad (2.109)$$

где $\Delta\Omega = \Delta p_x \Delta p_y \Delta p_z$ и $\Delta v = \Delta x \Delta y \Delta z$.

Из (2.108) также следует, что пространство импульса эквивалентно обратному пространству с точностью до множителя \hbar .

МИЛЛЕРОВСКИЕ ИНДЕКСЫ

Если заданы три базисных вектора, то положение любого узла решетки характеризуется началом координат и радиусом-вектором

$$\mathbf{r}_{n_1 n_2 n_3} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3, \quad (2.110)$$

определяющим положение данного узла по отношению к началу координат. Теперь мы должны найти математическое выражение для положения кристаллических плоскостей.

Кристаллической плоскостью мы будем называть любую плоскость, проходящую через какие-нибудь три узла кристалла, поэтому ее в принципе и можно бы было задать положением этих трех узлов, однако это чрезвычайно громоздко и неудобно; с другой стороны, в ряде случаев, и в первую очередь при рентгеноструктурном анализе, существенно положение не отдельной плоскости, а целого семейства плоскостей, расположенных параллельно друг другу, так как все эти плоскости принимают одинаковое участие в дифракции рентгеновских лучей.

Если за оси координат выбрать продолжения базисных векторов, образующих элементарную ячейку, то каждую плоскость можно характеризовать отрезками, отсекаемыми этой плоскостью на осях. Если нас интересует не одна определенная плоскость, а совокупность параллельных плоскостей, то мы должны брать не абсолютные значения, а отношения этих отрезков, так как они остаются постоянными для всех параллельных плоскостей. При этом остается некоторый произвол, так как эти числа всегда можно изменить на общий множитель (или общий делитель), поэтому естественно выражать его через отношение трех наименьших целых чисел. Однако принято характеризовать систему плоскостей не этим отношением, а обратными числами, называемыми миллеровскими индексами. Объясняется это двумя причинами:

а) для плоскости, параллельной одной из осей и отсекающей, например, равные отрезки на двух других, отношение отрезков было бы $1 : 1 : \infty$, миллеровские же индексы будут $[1, 1, 0]$, а с нулем всегда удобнее иметь дело, чем с бесконечностью;

б) величины, обратные отрезкам, отсекаемым на осях, в прямом пространстве определяют точку в обратном пространстве; таким образом, система плоскостей в прямом пространстве характеризуется точкой (или соответствующим вектором) в обратном пространстве; это основная и не-соизмеримо более важная причина.

2.3. ДЕФЕКТЫ В КРИСТАЛЛАХ

КЛАССИФИКАЦИЯ ДЕФЕКТОВ

До сих пор мы рассматривали кристалл, элементы которого (атомы, ионы или молекулы) построены в идеальном периодическом порядке и лишь совершают колебания

около равновесных положений. В действительности таких кристаллов не существует.

В реальных кристаллах по ряду причин возникают разнообразные дефекты:

— точечные дефекты — пустые узлы, атомы в междуузлиях, нарушения в правильности чередования атомов основного вещества, а также примеси посторонних атомов, которые также в зависимости от своих размеров могут располагаться и в узлах, и в междуузлиях *). Если число точечных дефектов велико, то они начинают взаимодействовать друг с другом, т. е. образовывать более или менее крупные комплексы;

— линейные дефекты — дислокации;

— плоские дефекты — границы монокристалла, зерен, двойников;

— объемные дефекты — трещины, полости, включения другой фазы и др.

В настоящее время большинство исследований по физике твердого тела посвящено прямо или косвенно изучению дефектов в кристаллах.

Это легко понять, если учесть влияние дефектов практически на все свойства кристаллов и, в частности, полупроводников:

— тип и концентрацию носителей примесных полупроводников;

— подвижность носителей, особенно при низких температурах;

— процессы генерации, рекомбинации и прилипания носителей;

— теплопроводность кристаллической решетки, в особенности при низких температурах;

— гальвано- и термомагнитные явления;

— процессы диффузии и ионной проводимости;

— механические свойства;

— каталитические свойства;

— работу выхода;

— взаимодействие с излучением (фотопроводимость, люминесценция, стимулированное излучение, резонансные явления — циклотронный, парамагнитный и ядерный резонансы).

*) Называя все эти дефекты точечными, мы абстрагируемся от их собственных размеров и деформаций, которые они создают в кристалле.

Теории и экспериментальному исследованию дефектов посвящено много оригинальных работ и ряд фундаментальных монографий.

В приведенном выше перечне структурно-чувствительных свойств полупроводников нас интересуют первые пять, на которые наибольшее влияние оказывают точечные дефекты, концентрация которых может колебаться в весьма больших пределах. Так, концентрация примесных атомов в очень чистых кристаллах германия и кремния не превышает 10^{12} — 10^{13} $см^3$, однако такая чистота достигается очень дорогой ценой и необходима лишь для ограниченного числа применений.

Для большинства приборов, применяемых в полупроводниковой электронике и энергетике, используются легированные кристаллы с концентрацией специально введенных примесей от 10^{14} до 10^{18} . В лазерах, туннельных диодах и термоэлементах применяются вырожденные полупроводники, в которых концентрация примесей достигает 10^{19} — 10^{21} $см^3$. В некоторых случаях в качестве матрицы применяются твердые растворы, в которых концентрации компонент соизмеримы и понятие «примесь» теряет свой смысл. Что касается «собственных» дефектов, т. е. дефектов атомов основного вещества, находящихся в междуузлиях и пустых узлах, то каждой температуре соответствует определенная равновесная концентрация и тех и других, причем для образования атома в междуузлии требуется в большинстве случаев больше энергии, чем для создания пустого узла, поэтому первых обычно меньше, чем вторых.

Если кристаллизация происходит путем застывания расплава, то в первый момент кристалл будет содержать очень большое число дефектов, соответствующее температуре плавления; при последующем охлаждении кристалл продолжает «благоустроиваться» и число сохранившихся в конечном счете дефектов зависит от скорости охлаждения (или если специально применяется отжиг, то от его температуры и длительности). Однако скорости всех процессов в кристалле, связанных с образованием, исчезновением и диффузией дефектов, экспоненциально зависят от температуры. Поэтому начиная с некоторой температуры, составляющей для различных веществ 0,6—0,8 от температуры плавления, эти процессы начинают протекать настолько медленно, что отжиг при более низких темпера-

гурах становится малоэффективным и число дефектов, соответствующее этой температуре, замораживается.

Обычно число пустых узлов в 1 см^3 кристалла колеблется в пределах от 10^{15} — 10^{19} , а число собственных атомов в междуузлиях на 1—2 порядка меньше. Наряду с этим существуют так называемые дефектные структуры, в которых число пустых и занятых узлов одного и того же порядка.

Атом любого вещества в вакууме характеризуется валентностью и потенциалами ионизации, т. е. работой, необходимой для отрыва первого, второго и т. д. электронов, а также электронным сродством, т. е. энергией, выделяющейся при захвате электрона.

Аналогичные понятия можно ввести и для примесных атомов в полупроводнике; при этом, однако, следует учитывать следующие факторы:

1. Валентность примесного атома, находящегося в узле, будет определяться разностью между числом его валентных электронов и числом электронов атома, который он заместил. Если это число больше, чем число электронов атома основного вещества, то такой атом будет вести себя как донор, в противоположном случае — как акцептор.

2. Для примесного атома, находящегося в междуузлии, определяющим будет его электроотрицательность: атомы щелочных металлов (например, литий) ведут себя всегда как доноры, такие атомы, как кислород и фтор, — как акцепторы.

3. Наряду с этим ряд примесных атомов проявляет «амфотерный» характер, т. е. их можно характеризовать несколькими величинами:

— «потенциалами ионизации», т. е. энергиями активации (энергиями, необходимыми для перевода в валентную зону);

— «электронным сродством», т. е. энергиями, необходимыми для переброса электрона из валентной зоны на примесный атом.

Зарядовое состояние такого атома будет зависеть от конкретной обстановки, т. е. от положения уровня химического потенциала.

ДЕФЕКТЫ ПО ФРЕНКЕЛЮ И ПО ШОТТКИ

По мере повышения температуры кристалла возрастают амплитуды тепловых колебаний ионов (или атомов). Часть из них (мы в дальнейшем для определенности будем гово-

ритель об ионах) приобретает при этом настолько большую энергию, что ионы срываются из равновесных положений и переходят в междуузлия. Таким образом, при температуре, отличной от нуля, всегда в кристалле существует некоторое число пустых узлов и ионов в междуузлиях, причем в равновесии число и тех и других должно экспоненциально расти с температурой. Аналогичный процесс может происходить и вблизи поверхности кристалла: один атом может из приповерхностного слоя перескочить на поверхность, другой атом — занять его место, третий атом — занять место второго атома и т. д. Таким образом, в объеме кристалла с его поверхности продиффундирует пустой узел. Этот процесс приведет, следовательно, к образованию в объеме кристалла одних только пустых узлов.

Первый механизм образования дефектов — переход из узла в междуузлие — называется механизмом образования дефектов по Френкелю (и соответствующие дефекты — дефектами по Френкелю); второй — механизмом образования по Шоттки и дефекты — дефектами по Шоттки *).

Попытаемся теперь количественно оценить равновесное число дефектов по Френкелю, при этом, как и в случае определения концентраций свободных электронов, возможны два подхода — кинетический и статистический.

Кинетический подход. Пусть число ионов данного сорта в единице объема N , из них n перешли в междуузлие, причем энергия, необходимая для такого перехода, будет \mathcal{E}_Φ . Тогда вероятность ω_d того, что еще один ион перейдет в междуузлие, будет пропорциональна числу ионов, еще сидящих в узлах, т. е. $(N - n)$, и вероятности того, что один из этих ионов приобретает необходимую для такого перехода энергию, т. е. экспоненциальному множителю $e^{-\mathcal{E}_\Phi/kT}$. Таким образом,

$$\omega_d = \alpha (N - n) e^{-\frac{\mathcal{E}_\Phi}{kT}}, \quad (2.111)$$

где α — неизвестный коэффициент.

* Следует отметить, что впервые существование и тех и других дефектов предсказал Я. И. Френкель. Можно в принципе представить образование и так называемых антишотткиевских дефектов: атом с поверхности внедряется в междуузлие и уходит в объем, при этих условиях в кристалле будут существовать атомы в междуузлиях и не будет пустых узлов, однако число таких дефектов, как правило, мало.

Вероятность обратного процесса «рекомбинации», т. е. возвращения иона на пустой узел ω_p , будет пропорциональна числу атомов в междуузлиях n и числу пустых узлов, также n , точнее, тому, что атом в междуузлии «наткнется» на пустой узел, т. е. n/N :

$$\omega_p = \beta n \frac{n}{N}, \quad \text{а не exp.} \quad (2.112)$$

где β — постоянный неизвестный коэффициент *).

Таким образом,

$$\frac{dn}{dt} = \alpha (N - n) e^{-\frac{\mathcal{E}_\Phi}{kT}} - \beta n^2. \quad (2.113)$$

Равновесие наступит тогда, когда вероятности обоих процессов будут равны и $dn/dt = 0$. Полагая $n \ll N$, получим

$$n \approx \sqrt{\frac{\alpha}{\beta}} N e^{-\frac{\mathcal{E}_\Phi}{2kT}}. \quad (2.114)$$

Достоинство этого вывода в том, что он позволяет проследить кинетику процесса, например, при изменении температуры; недостаток — в том, что он содержит отношение двух неизвестных коэффициентов α/β .

Перейдем теперь к статистическому методу, который не обладает ни этими недостатками, ни этими достоинствами.

Статистический подход. Пусть n и N сохраняют прежние значения и N_1 — общее число междуузлий, в которых могут разместиться n перешедших в них ионов. Тогда эти ионы могут разместиться следующим числом способов:

$$\omega_1 = \frac{N_1!}{(N_1 - n)! n!} = C_{N_1}^n \quad (2.115)$$

Точно так же вычислим число способов, которым можно выбрать n вакантных узлов из общего числа N :

$$\omega = \frac{N!}{(N - n)! n!} = C_N^n \quad (2.116)$$

*) Мы могли бы также написать $\omega_p = \beta n^2$, включив $1/N$ в коэффициент β .

Таким образом, соответствующее увеличение энтропии*) данного состояния будет (по сравнению с упорядоченным)

$$S = k \ln (\omega_1 \omega). \quad (2.117)$$

Увеличение внутренней энергии кристалла (по сравнению с состоянием полного порядка) составит

$$\Delta \mathcal{E} = n \mathcal{E}_\phi. \quad (2.118)$$

Если предположить (в первом приближении), что объем кристалла не зависит от числа дефектов, и пренебречь также влиянием дефектов на спектр колебаний (т. е. теплосодержание кристалла), то условием равновесия будет минимум свободной энергии $\partial F / \partial n = 0$, где

$$F = \mathcal{E} - TS. \quad (2.119)$$

Воспользуемся формулой Стирлинга для приближенного выражения логарифмов факториалов в формуле (2.116) для энтропии:

$$\ln (n!) \approx \left(n + \frac{1}{2} \right) \ln n - n + \ln \sqrt{2\pi} \quad (2.120)$$

или при $n \gg 1$

$$\ln (n!) \approx n (\ln n - 1). \quad (2.121)$$

Следовательно, согласно (2.117) и (2.121)

$$S = k [N \ln N - (N - n) \ln (N - n) - n \ln n] + \\ + k [N_1 \ln N_1 - (N_1 - n) \ln (N_1 - n) - n \ln n]. \quad (2.122)$$

Воспользовавшись (2.119) и (2.122), получаем

$$\frac{\partial F}{\partial n} = \mathcal{E}_\phi - kT [\ln (N - n) + \ln (N_1 - n) - 2 \ln n] = 0, \quad (2.123)$$

откуда

$$\frac{n^2}{(N - n)(N_1 - n)} = e^{-\frac{\mathcal{E}_\phi}{kT}}, \quad (2.124)$$

и, полагая по-прежнему $n \gg N_1$ и $n \gg N$, получим выражение для n без неизвестных коэффициентов

$$n \approx \sqrt{NN_1} e^{-\frac{\mathcal{E}_\phi}{2kT}}. \quad (2.125)$$

*) Коротко основные понятия термодинамики и статистики охарактеризованы в § 1 гл. 4.

Подсчитаем теперь число дефектов по Шоттки n_1 , обозначим через \mathcal{E}_{III} энергию, необходимую для перевода атома из объема на поверхность *).

Проведя рассуждения, совершенно аналогичные приведенным ранее, получим

$$F = n_1 \mathcal{E}_{III} - k \ln \frac{N!}{(N-n_1)! n_1!}, \quad (2.126)$$

откуда из условий $\partial F / \partial n_1 = 0$ получаем

$$\frac{n_1}{N-n_1} = e^{-\frac{\mathcal{E}_{III}}{kT}}. \quad (2.127)$$

Как видно из (2.127), в этом случае множитель $1/2$ в числителе экспоненты отсутствует. Объясняется это тем, что в первом случае рекомбинация пустого узла и атома в междуузлии носила характер бимолекулярной реакции, а во втором — мономолекулярной **).

Так, однако, может происходить дело только в атомных кристаллах, в полярных же (при сохранении стехиометрического состава) из условия нейтральности следует, что на поверхность ионы могут выходить только парами; таким образом, число дефектов по Шоттки в обеих подрешетках должно быть одинаково. Обозначив энергию, необходимую для образования пары вакансий, \mathcal{E}_p , число отрицательных ионов N' , повторив проведенные выше рассуждения, получим

$$n_1 = \sqrt{NN'} e^{-\frac{\mathcal{E}_p}{2kT}}. \quad (2.128)$$

Согласно (2.125) и (2.128) отношение числа парных дефектов по Шоттки к числу дефектов по Френкелю

$$\frac{n_1}{n} \approx e^{-\frac{\mathcal{E}_p - \mathcal{E}_f}{2kT}}, \quad (2.129)$$

*) Атом на поверхности может занять разные положения. Мы, однако, должны взять среднее значение энергии на поверхности. Нетрудно убедиться, что это будет энергия атома, примкнувшего к «ступеньке». При этих условиях $\mathcal{E}_{III} = 1/2 \mathcal{E}$, где \mathcal{E} — энергия, необходимая для полного отрыва атома от кристалла.

**) Для исчезновения дырки никакой «встречи» не надо — достаточно пустому узлу подойти к поверхности, а число атомов на поверхности не зависит от n_1 .

\mathcal{E}_p и \mathcal{E}_f имеют порядок одного электрон-вольта. Две величины порядка 1 эв могут лишь случайно (очень редко) совпасть, обычно же их разность будет составлять несколько десятых электрон-вольта, и так как $kT \approx 0,03$ эв, то

$$\frac{n_1}{n} = e^{\pm 10}. \quad (2.130)$$

Эта оценка показывает, что мало вероятно, что число дефектов обоих сортов будет одного и того же порядка. Обычно один из типов явно преобладает над другим.

Более строгие оценки показывают, что в большинстве случаев это имеет место; в частности в щелочно-галогидных кристаллах преобладают дефекты по Шоттки. Напротив, в галогенидах серебра небольшие ионы серебра легко забираются в междуузлия; в этом случае преобладают дефекты по Френкелю, а некоторые авторы считают, что даже антишотткиевские дефекты. Результаты опытов показывают, что ионная проводимость здесь осуществляется, главным образом, путем диффузии ионов серебра по междуузлиям.

ИОННАЯ ПРОВОДИМОСТЬ

Как впервые показал Я. И. Френкель, идеальный кристалл не может обладать ионной проводимостью. Действительно, в идеальном кристалле для прохождения тока должны одновременно перемещаться целые цепочки ионов, содержащие примерно 10^{22} атомов — процесс, вероятность которого близка к нулю. Точно так же нельзя предположить, что под действием поля ионы срываются со своих мест. Энергия связи иона в узле порядка 1 эв, следовательно, для того чтобы на расстоянии от узла до междуузлия, примерно равном 10^{-8} см, поле могло совершить такую работу, его напряженность должна быть

$$E = \frac{V}{l} \approx 10^8 \frac{\text{в}}{\text{см}}.$$

В действительности ионная проводимость начинает проявляться уже в полях 1—10 в/см, таким образом, и этот механизм проводимости исключается. Как показал Я. И. Френкель, ионная проводимость так же, как и диффузия ионов, может осуществляться только следующими способами:

— перемещением ионов по междуузлиям;

— последовательным эстафетным перемещением пустых мест, эквивалентным перемещению заряда противоположного знака в обратном направлении; диффузия может также осуществляться последовательным циклическим обменом местами соседних ионов.

Ионную проводимость можно представить точно в таком же виде, как электронную:

$$\sigma = \sum_i e_i n_i u_i, \quad (2.131)$$

где n_i — число дефектов i -го сорта;

e_i — их эффективный заряд;

u_i — подвижность.

Число дефектов и его температурную зависимость мы определили в предыдущем параграфе, нам остается вычислить подвижность.

Вычислим сначала коэффициент диффузии, так как и подвижность и диффузия определяются одной и той же вероятностью перескоков, но вычисление коэффициента диффузии несколько проще. Для определенности мы будем говорить о движении ионов, хотя все рассуждения можно было с тем же правом отнести к вакантным узлам.

Представим кристалл, в котором по какой-либо причине имеется градиент концентрации атомов в междуузлиях. Выделим две междуузельные плоскости и обозначим концентрацию атомов в них соответственно n и $n' = n + a (dn/dx)$, где a — постоянная решетки. Пусть ν_0 — число колебаний атома в междуузлии в 1 сек, \mathcal{E}_a — энергия активации, необходимая для преодоления потенциального барьера между междуузлиями, тогда вероятность перескока будет

$$\nu = \nu_0 e^{-\frac{\mathcal{E}_a}{kT}}.$$

Вероятность перескока в заданном направлении (x) будет

$$\nu_x = \frac{\nu_0}{6} e^{-\frac{\mathcal{E}_a}{kT}}, \quad (2.132)$$

так как всего имеется шесть эквивалентных направлений: $+x$, $-x$, $+y$, $-y$, $+z$, $-z$ (более строгий расчет, связанный с интегрированием по углам, приводит к тому же результату). Скорость диффузии данного иона будет равна

произведению числа перескоков за одну секунду v_x на расстояние, проходимое при одном перескоке, т. е. a ,

$$v_x = av_x = \frac{av_0}{6} e^{-\frac{\mathcal{E}_a}{kT}}. \quad (2.133)$$

Таким образом, диффузионный поток будет равен разности общего числа перескоков из первой плоскости во вторую и обратно:

$$Q = nav_x - \left(n + a \frac{dn}{dx}\right) av_x = -\frac{v_0 a^2}{6} e^{-\frac{\mathcal{E}_a}{kT}} \frac{dn}{dx}.$$

Следовательно, коэффициент диффузии

$$D = \frac{a^2 v_0}{6} e^{-\frac{\mathcal{E}_a}{kT}}. \quad (2.134)$$

Перейдем теперь к вычислению подвижности. В этом случае высота потенциальных горбов для перескоков слева направо и справа налево будет неодинакова:

$$\mathcal{E}_1 = \mathcal{E}_a + \frac{Eea}{2}$$

*E - напряж-ть поля
e - заряд электрона
a - расстояние между атомами*

и

$$\mathcal{E}_2 = \mathcal{E}_a - \frac{Eea}{2};$$

соответственно будут различаться и вероятности перескоков:

$$v_1 = \frac{v_0}{6} e^{-\frac{\mathcal{E}_a + \frac{Eea}{2}}{kT}} \approx v_p \left(1 - \frac{Eea}{2kT}\right) \quad (2.135)$$

и

$$v_2 = \frac{v_0}{6} e^{-\frac{\mathcal{E}_a - \frac{Eea}{2}}{kT}} \approx v_p \left(1 + \frac{Eea}{2kT}\right), \quad (2.136)$$

где v_p — равновесное число перескоков вдоль оси при данной температуре.

Скорость дрейфа в электрическом поле будет равна разности числа перескоков слева направо и в обратном направлении, умноженной на расстояние, проходимое при одном перескоке:

$$v_x = (v_2 - v_1) a = \frac{ea^2}{6kT} E v_0 e^{-\frac{\mathcal{E}_a}{kT}}, \quad (2.137)$$

и подвижность будет равна

$$u = \frac{v_x}{E} = \frac{ea^2\nu_0}{6kT} e^{-\frac{\mathcal{E}_a}{kT}}. \quad (2.138)$$

Сравнение (2.134) и (2.138) показывает, что подвижность и коэффициент диффузии связаны соотношением Эйнштейна:

$$u = \frac{eD}{kT}. \quad (2.139)$$

Согласно (2.125) и (2.138) электропроводность, обусловленная данным видом дефектов,

$$\sigma = enu = \frac{e^2 a^2}{kT} \sqrt{NN'} \exp\left(-\frac{\frac{1}{2}\mathcal{E}_p + \mathcal{E}_a}{kT}\right). \quad (2.140)$$

Если существует несколько типов дефектов, то надо пользоваться обобщенной формулой (2.131), однако, как мы уже упоминали (см. стр. 171), очень маловероятно, чтобы концентрации дефектов разного сорта были бы соизмеримы (в данном кристалле при данной температуре); то же самое относится и к парциальным электропроводностям.

2.4. ТЕПЛОВЫЕ КОЛЕБАНИЯ КРИСТАЛЛОВ

Равновесное расстояние r_0 между атомами в кристалле определяется равенством сил притяжения и сил отталкивания. При удалении атомов на расстояние, большее чем r_0 , силы притяжения начинают преобладать над силами отталкивания, при сближении имеет место обратное соотношение. Однако при $r \rightarrow \infty$ силы притяжения стремятся к нулю, а при $r \rightarrow 0$ силы отталкивания неограниченно возрастают, иными словами, зависимость силы и энергии взаимодействия атомов от расстояния между ними носит несимметричный характер и имеет приблизительно вид, представленный на рис. 2.10.

Энергию кристалла можно представить в виде суммы $U = U' + U''$, где U' — энергия сил отталкивания и U'' — энергия сил притяжения; разумеется, $|U''| > |U'|$ и $U'' < 0$. Силы отталкивания, преобладающие при сближении атомов на расстояние $r < r_0$, возникают главным образом в результате перекрытия электронных оболочек. При этом отталкивание возникает по двум причинам: непо-

средственно из-за кулоновского отталкивания электронов и вследствие роста кинетической энергии электронов при сближении атомов (чем больше плотность электронов, тем больше их кинетическая энергия, как это следует из принципа неопределенности Гейзенберга и принципа Паули *).

Теоретическое рассмотрение не дает простого выражения для энергии этого взаимодействия, и поэтому обычно его описывают (аппроксимируют) либо степенной зависимостью

$$U' = \frac{b}{r^n}, \quad (2.141)$$

либо экспоненциальной

$$U' = \lambda e^{-\frac{r}{\rho}}. \quad (2.142)$$

Результаты теоретических расчетов говорят в пользу экспоненциальной зависимости; преимущество обратной степенной зависимости в большей простоте. Константы, входящие в выражения (2.141) и (2.142), подбирают с помощью эмпирических данных. Опыт показал, что показатель степени в (2.141) можно варьировать в довольно широких пределах (9—12), подгоняя при этом соответственно коэффициент в (2.141). Так как силы отталкивания очень быстро исчезают при увеличении расстояния, абсолютное значение энергии сцепления кристалла U определяется главным

*) Чем меньше линейные размеры области Δl , занимаемой электроном Δx , тем больше его импульс ($p_{\text{мин}} \geq h/\Delta x$) и кинетическая энергия $T_{\text{мин}} = p_{\text{мин}}^2/2m$. Таким образом, чем ближе проходит орбита электрона к положительно заряженному ядру, тем меньше его потенциальная энергия $U = -e^2/2r$, но больше кинетическая. Из условия минимума полной энергии $\mathcal{E} = T + U$ можно найти наименьшее состояние частицы для любой формы потенциальной ямы (отсюда, в частности, следует, что энергия осциллятора при абсолютном нуле не может равняться нулю, а равна энергии нулевых колебаний). Минимальную энергию любой частицы можно также приблизительно оценить из того условия, что в потенциальной яме должна укладываться, по крайней мере, одна полуволна. В этом первом состоянии могут поместиться только два электрона с противоположно направленными спинами. Для последующих электронов в потенциальной яме должны поместиться уже две полуволны, и соответственно их полная энергия будет больше. Поэтому, чем больше плотность электронов в обычном объеме, тем больше объем, занимаемый ими в пространстве импульсов и больше кинетическая энергия.

образом природой сил, дающих основной вклад в энергию притяжения U'' *). Так, например, в ионных кристаллах и в металлах $U'' \approx 1/r$, в ван-дер-ваальсовых $U'' \approx 1/r^6$, в ковалентных — силы притяжения имеют своей природой квантовомеханический характер и не выражаются простой формулой; единственное, что можно сказать: они очень быстро убывают с расстоянием (но все же медленнее, чем силы отталкивания).

Если под влиянием какого-то воздействия атом выведен из своего равновесного положения, то он совершает колебательное движение, и так как все атомы твердого тела связаны друг с другом упругими силами, то колебания любого атома передаются соседнему и, таким образом, по всему телу во всевозможных направлениях распространяются упругие волны. Эти колебания отличаются друг от друга не только направлением распространения, но и длиной волны, сдвигом фазы соседних атомов и поляризацией. Наиболее коротковолновые колебания имеют длину, равную удвоенному расстоянию между соседними атомами, наиболее длинные — длину волны, равную удвоенной длине кристалла L .

Кроме этого, колебания кристаллической решетки могут отличаться своей поляризацией, т. е. направлением колебаний; для каждого направления распространения существуют две поперечно поляризованные волны и одна продольно поляризованная **).

Таким образом, каждая волна характеризуется следующими параметрами:

1) длиной λ и направлением распространения; в теории колебаний обычно вводят понятие волнового вектора \mathbf{q} , равного по абсолютной величине числу волн, укладываемых на единице длины ($\mathbf{q} = 1/\lambda$), а по направлению совпадающего с направлением распространения волны;

2) направлением поляризации \mathbf{p}_i ($i = 1, 2, 3$); при этом $\mathbf{p}_1 \parallel \mathbf{q}$ (продольные колебания) и $\mathbf{p}_{2,3} \perp \mathbf{q}$ (поперечные колебания);

3) частотой, которая в общем случае зависит от длины волны, направления ее распространения и направления

*) Коэффициент упругости, напротив, определяется силами отталкивания, так как вблизи равновесного положения они значительно более сильно зависят от расстояния, чем силы притяжения.

**) Для самых коротких волн направление колебаний может быть и не строго параллельным и не строго перпендикулярным направлению распространения.

поляризации:

$$v = v(q, p);$$

4) скоростью распространения.

Каждая волна характеризуется двумя скоростями: фазовой и групповой (которые в отдельных случаях могут совпадать). Если мы имеем монохроматическую волну с постоянной в пространстве амплитудой

$$\psi = A \sin 2\pi (qx - vt), \quad (2.143)$$

то скорость перемещения в пространстве точки с заданной фазой

$$qx - vt = \varphi$$

определяется равенством

$$x = \frac{\varphi}{q} + \frac{v}{q} t,$$

т. е. фазовая скорость

$$v_{\varphi} = \frac{v}{q} = v\lambda. \quad (2.144)$$

Если же мы имеем не монохроматическую волну, а импульс (волновый пакет), то он распространяется в пространстве с так называемой групповой скоростью

$$v_{гп} = \frac{\partial v}{\partial q}. \quad (2.145)$$

В теории волнового движения решающую роль играет дисперсия, т. е. зависимость частоты и скорости распространения волны от ее длины (или, что то же самое, волнового вектора). Для того чтобы найти эти зависимости, мы должны написать и решить уравнения движения для колеблющихся атомов.

Аналитическое выражение энергии взаимодействия атомов, представленной графически на рис. 2.10, является слишком сложным для решения этой задачи (даже если учитывать только взаимодействие ближайших соседей), поэтому обычно идут на радикальное упрощение: считая колебания вблизи точки равновесия малыми, зависимость энергии U от расстояния между атомами считают квадратичной, а силы — линейной.

При этих условиях уравнение движения для n -го атома линейной цепочки принимает вид

$$M \frac{\partial^2 u_n}{\partial x^2} = f(u_n - u_{n-1}) - f(u_n - u_{n+1}), \quad (2.146)$$

где u — смещение атомов из положения равновесия;
 M — масса;
 f — коэффициент упругой силы.

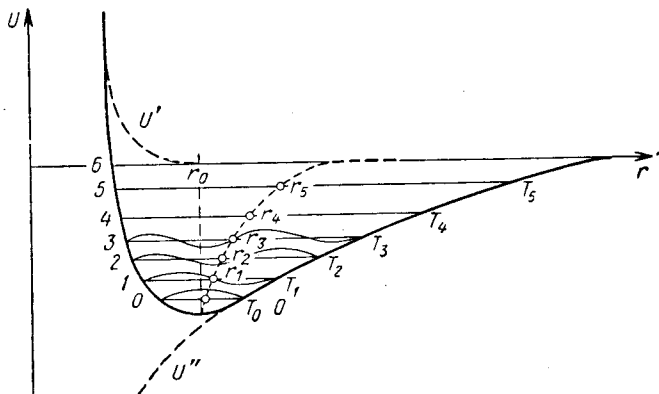


Рис. 2.10. Зависимость энергии взаимодействия атомов от расстояния r между ними и различные колебательные уровни молекулы.

Нижний уровень соответствует нулевым колебаниям; r_0 — равновесное расстояние при абсолютном нуле; r_1, r_2, r_3, r_4 и r_5 — равновесные расстояния при соответствующих температурах. Уровень 6 соответствует диссоциации молекулы.

Нетрудно убедиться, что общее решение уравнения (2.146) можно получить в виде суммы бегущих волн:

$$u_n = \sum_i A_i \sin 2\pi (v_i t - q_i x_n + \varphi_i), \quad (2.146a)$$

где $x_n = an$ — координата n -го атома в цепочке;
 a — постоянная решетки, т. е. расстояние между соседними атомами.

Подставляя (2.146a) в (2.146), после простых преобразований находим закон дисперсии:

$$v_i = \frac{1}{2\pi} \sqrt{\frac{f}{M}} \sin \pi a q_i. \quad (2.147)$$

На рис. 2.11, а представлена зависимость v от q , построенная согласно (2.147). На рис. 2.11, б и в построены аналогичные зависимости для $\omega_{\phi} = v/q$ и $\omega_{гр} = \partial v/\partial q$.

Как видно из этих рисунков, для обеих скоростей имеет место дисперсия, т. е. обе они зависят от длины волны, причем для самых коротких волн групповая скорость стремится к нулю, а для самых длинных обе скорости обращаются в постоянную величину $\omega_{гр} = \omega_{\phi} = a\sqrt{f/M}$, равную скорости распространения звука в кристалле *).

Интересно посмотреть, какому участку графика рис. 2.11 соответствуют звуковые волны. Как известно, частота звуковых колебаний колеблется от $\nu_{мин} \approx 10^2$ гц до $\nu_{макс} \approx 10^4$ гц; скорость звука $c \approx 10^5$ см/сек; следовательно, $\lambda_{мин} = c/\nu_{макс} = 10$ см и $\lambda_{макс} = 10^3$ см и $q_{макс}$ звук $= 1/\lambda_{мин} = 1/10 = 10^{-1}$, в то время как для всего спектра $q_{макс} = 1/2a \approx 10^8$ см $^{-1}$ и $\nu_{макс} \approx 10^{13}$ гц. Таким образом, звуковые волны представляют совершенно ничтожную часть спектра колебаний твердого тела. Нетрудно убедиться в том, что если бы твердое тело представляло собой непрерывную среду, то дифференциальное уравнение (2.146) перешло бы в волновое уравнение

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad (2.148)$$

*) Групповая скорость играет большую роль в явлениях переноса (в частности, теплопроводности), так как она определяет скорость переноса энергии (см. ниже).

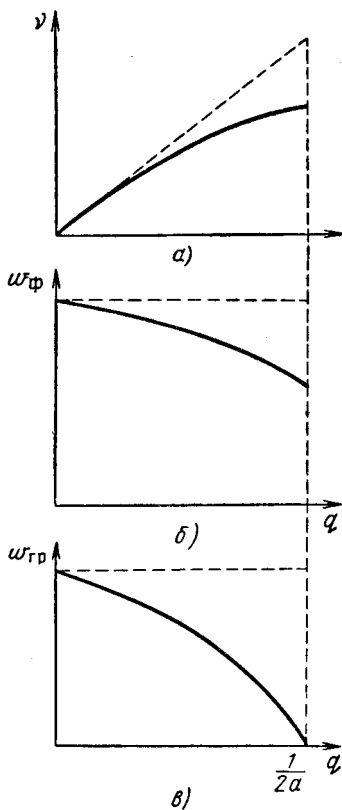


Рис. 2.11. Дисперсионные кривые для одноатомной цепочки. Зависимость от волнового вектора частоты (а), фазовой (б) и групповой (в) скорости. Пунктиром показаны соответствующие зависимости для струны.

Дисперсионная кривая выродилась бы в прямую $v = \omega q$, и групповая и фазовая скорости были бы постоянны и равны скорости звука. Это следует также и из качественных соображений: для длинных волн ($\lambda \gg a$) несущественно, «размазана» ли масса равномерно или сосредоточена в отдельных точках, для коротких же ($\lambda \approx 2a$) сосредоточение массы в точках максимального отклонения увеличивает инерционность системы (т. е. уменьшает частоту и скорость распространения волны).

Рассмотрим теперь второй случай, имеющий большое теоретическое и практическое значение, — кристалл (точнее цепочку), состоящий из двух сортов атомов с различной массой, соответственно M и m . В этом случае мы должны написать отдельно уравнения движения для атома каждого сорта (четных и нечетных):

$$M \frac{\partial u_{2n}}{\partial x^2} = f(u_{2n-1} - u_{2n}) - f(u_{2n+1} - u_{2n}), \quad (2.149)$$

$$m \frac{\partial u_{2n+1}}{\partial x^2} = f(u_{2n} - u_{2n+1}) - f(u_{2n+2} - u_{2n+1}). \quad (2.149a)$$

Решения этих уравнений можно искать в виде колебаний с одинаковой частотой и длиной волны, но с разной амплитудой (для удобства вычислений мы запишем их в комплексной форме):

$$\begin{aligned} u_{2n} &= A_M e^{2\pi i(qa2n - vt)}, \\ u_{2n+1} &= A_m e^{2\pi i[qa(2n+1) - vt]}. \end{aligned} \quad (2.150)$$

Подставляя (2.150) в (2.149), получим два уравнения, связывающие частоту с волновым вектором и амплитуды A_M и A_m :

$$\begin{aligned} \left(4\pi v^2 - \frac{f}{M}\right) A_M + \left(\frac{2f}{M} \cos 2\pi aq\right) A_m &= 0, \\ \left(\frac{2f}{m} \cos 2\pi aq\right) A_M + \left(4\pi v^2 - \frac{f}{m}\right) A_m &= 0. \end{aligned} \quad (2.151)$$

Решая совместно систему (2.151) двух однородных уравнений, находим

$$\frac{A_M}{A_m} = \frac{\cos 2\pi aq}{1 - \frac{4\pi M}{f} v^2} = \frac{1 - \frac{4\pi m}{f} v^2}{\cos 2\pi aq}, \quad (2.152)$$

откуда

$$v^2 = \frac{1}{2} v_0^2 \left(1 \mp \sqrt{1 - \gamma^2 \sin^2 \frac{2\pi a q}{2}} \right), \quad (2.153)$$

где $v_0 = \frac{1}{2\pi} \sqrt{\frac{f}{\bar{M}}}$, $\gamma = \frac{2mM}{(M+m)^2}$ и \bar{M} — «средняя масса», равная $\bar{M} = 2mM/(m+M)$.

Выражение (2.153) показывает, что в двухатомной цепочке в отличие от одноатомной у дисперсионной кривой есть две ветви. Можно показать, что и в данном случае волновой вектор может приобретать значения

$$q_i = \frac{i}{2Na}, \quad \text{где } 1 < i < N.$$

На рис. 2.12 представлены дисперсионные кривые, построенные согласно (2.153). Колебания, описываемые верх-

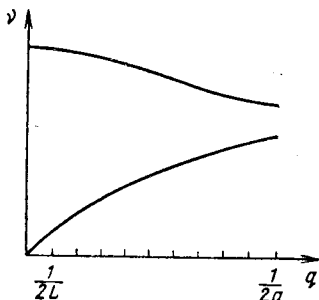


Рис. 2.12. Дисперсионные кривые, иллюстрирующие зависимость частоты от волнового вектора для двухатомной цепочки.

ней ветвью, называются оптическими, а нижней — акустическими, т. е. мы сталкиваемся с парадоксальным на первый взгляд фактом: каждой длине волны соответствуют два колебания с существенно различающимися частотами. Нижняя кривая соответствует знаку «—» в уравнении (2.153), верхняя — знаку «+». В действительности этот странный на первый взгляд факт имеет простое физическое объяснение. Наиболее наглядно это видно на примере самых коротких волн $\lambda_{\min} = 2a$ (см. рис. 2.13, а); как показано на рисунке, в одном случае колеблются более легкие атомы, а более тяжелые остаются в покое, а во втором — наоборот, движутся более тяжелые, а покоятся легкие. Так как длина волны и возвращающаяся в равновесие сила в обоих случаях одинаковы, то ясно, что в первом случае частота колебаний будет больше, чем во втором.

Случай очень длинных волн ($\lambda \gg a$) проиллюстрирован рис. 2.13, б и в. Колебание, представленное на рис. 2.13, б, ничем не отличается от случая одноатомной решетки. Ясно, что в данном случае частота будет определяться средней массой атомов, и это же следует из формулы (2.153); если взять перед радикалом знак $(-)$ и $\lambda \rightarrow \infty$, то $\nu \rightarrow 0$.

Совершенно по-другому будет обстоять дело в случае колебаний, представленных на рис. 2.13, в и г*). Здесь мы имеем дело по существу с внутримолекулярными колебаниями, с той лишь разницей, что сила,

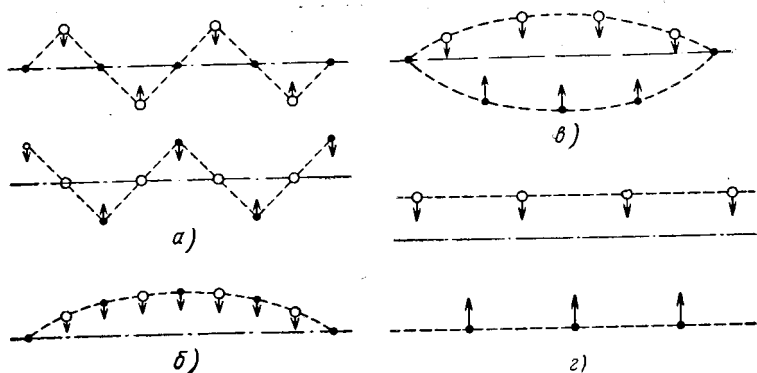


Рис. 2.13. Колебания двухатомной цепочки:

а) $\lambda = 2a$, б) $\lambda \gg a$ (акустические колебания), в) $\lambda \gg a$ (оптические колебания), г) $\lambda = \infty$ (оптические колебания).

возвращающая каждый атом в положение равновесия, будет вдвое больше. Поэтому и частота в этом случае отличается от частоты молекулярных колебаний множителем $\sqrt{2}$. Таким образом, в цепочке, состоящей из двух сортов атомов, дисперсионная кривая (т. е. зависимость частоты от волнового числа) разбивается на две ветви; как мы уже упоминали, первая из них та, для которой при $q \rightarrow 0$ значение $\nu \rightarrow 0$, называется акустической (так как начало этой ветви включает частоты звуковых колебаний), вторая ветвь называется оптической (так как для этой ветви при $q \rightarrow 0$ значение $\nu \rightarrow 10^{13}$, что соответствует частотам электромагнитных колебаний инфракрасной области спектра).

*) На рис. 2.13, г все молекулы колеблются в одной и той же фазе, следовательно, $\lambda = \infty$.

Оптическая ветвь колебаний возникает не только в результате различия масс атомов (случай, рассмотренный выше), но и тогда, когда массы всех атомов одинаковы, а расстояния между молекулами и внутри молекулы различны *), а следовательно, отличны и силы, возникающие при межмолекулярных и внутримолекулярных колебаниях. Нетрудно убедиться, что и математическая теория дает здесь два значения частоты для каждой длины волны (за счет двух коэффициентов упругой связи, f_1 и f_2).

В заключение этого параграфа остановимся коротко на колебаниях трехмерной кристаллической решетки. Математическая теория в этом случае весьма громоздка, поэтому ограничимся лишь качественными соображениями. В кристалле с одним атомом в элементарной ячейке существуют только акустические колебания. При этом каждому заданному по величине и направлению волновому вектору соответствуют три колебания: одно продольное и два поперечных. При той же амплитуде и длине волны возвращающая упругая сила при продольных колебаниях больше, чем при поперечных, поэтому частота и скорость поперечных колебаний соответственно меньше **).

В анизотропных средах частота и скорость распространения волны могут зависеть не только от величины волнового вектора, но и от его направления. Поэтому уже для *плоской* решетки зависимость частоты от волнового вектора представляется в виде сложных поверхностей, которые для разных направлений могут пересекаться и касаться. Для того чтобы описать геометрически дисперсионные зависимости для объемной решетки $v = f(\mathbf{q}) = f(q_x, q_y, q_z)$, надо представить дисперсионную кривую в четырехмерном пространстве.

В более сложных кристаллах, в элементарной ячейке которых содержится m атомов, есть также три ветви аку-

*) Это бывает, например, когда молекулы связаны между собой ван-дер-ваальсовыми силами, а атомы внутри молекулы — ковалентной связью, а также в сложных решетках Ge и Si и др.

***) Это связано с тем, что продольные колебания вызывают сжатие и растяжение, а поперечные — только сдвиг. В жидкостях для не слишком высоких частот модуль сдвига равен нулю и поперечные колебания вообще не возникают.

Мы уже упоминали, что для очень коротких волн грань между продольными и поперечными колебаниями стирается, так как вектор поляризации, как правило, составляет некоторый угол с направлением распространения волны, отличный от нуля и 90° .

стических колебаний, для которых все атомы одной элементарной ячейки колеблются вместе как целое, и $3(m-1)$ оптические ветви, соответствующие внутримолекулярным «внутриячеечным» колебаниям.

2.5. ТЕПЛОЕМКОСТЬ

Мы уже упоминали (см. гл. 1), что первое качественное объяснение температурной зависимости теплоемкости дал в 1911 г. А. Эйнштейн. Для этого он предположил, что кристалл можно представить в виде совокупности N осцилляторов, колеблющихся независимо с частотой $\nu = \frac{1}{2\pi} \sqrt{f/m}$, и что энергия этих осцилляторов может принимать не любые значения, а лишь кратные $h\nu$: $\epsilon = nh\nu$, где h — постоянная Планка и n — любое целое число.

Эйнштейн оформил изложенные выше соображения количественно и вывел формулу для средней энергии осциллятора. Мы приведем здесь эти вычисления, так как они чрезвычайно просты и имеют наглядный физический смысл.

Действительно, предположим, что в соответствии с распределением Больцмана, вероятность $\omega(\epsilon)$ того, что колеблющаяся частица имеет энергию $\epsilon = nh\nu$, пропорциональна $e^{-\epsilon/kT}$, тогда средняя энергия

$$\bar{\epsilon} = \frac{\sum_{n=0}^{\infty} nh\nu e^{-\frac{nh\nu}{kT}}}{\sum_{n=0}^{\infty} e^{-\frac{nh\nu}{kT}}} \quad (2.154)$$

или

$$\bar{\epsilon} = - \frac{d}{d\left(\frac{1}{kT}\right)} \ln \sum_{n=0}^{\infty} e^{-\frac{nh\nu}{kT}}. \quad (2.155)$$

Выражение под знаком логарифма есть сумма членов бесконечной геометрической прогрессии со знаменателем $a = e^{-h\nu/kT}$.

Таким образом,

$$\sum_{n=0}^{\infty} e^{-\frac{nh\nu}{kT}} = \frac{1}{1 - e^{-\frac{h\nu}{kT}}}. \quad (2.156)$$

После логарифмирования и дифференцирования по $1/kT$ получаем окончательно

$$\bar{\epsilon} = \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1}. \quad (2.157)$$

Нетрудно из (2.157) получить предельные значения энергии при очень высоких и очень низких температурах. Действительно, при $kT \ll h\nu$ единицей в знаменателе (2.157) можно пренебречь и

$$\bar{\epsilon}_v \approx e^{\frac{-h\nu}{kT}} h\nu;$$

при $kT \gg h\nu$, разлагая экспоненту в знаменателе в ряд и ограничиваясь в разложении двумя первыми членами, получим $\bar{\epsilon} \approx kT$ и, следовательно, энергия грамм-атома твердого тела $\mathcal{E} = 3N\bar{\epsilon} = 3RT$ и теплоемкость $C = d\mathcal{E}/dT = 3R$ в полном согласии с законом Дюлонга и Пти. Согласно (2.157) энергия грамм-атома твердого тела при любой температуре будет

$$\mathcal{E} = \frac{3N h\nu}{e^{\frac{h\nu}{kT}} - 1} \quad (2.158)$$

и теплоемкость

$$C = \frac{d\mathcal{E}}{dT} = \frac{3N(h\nu)^2}{\left(e^{\frac{h\nu}{kT}} - 1\right)^2} \frac{1}{kT^2}. \quad (2.159)$$

На рис. 1.11 пунктиром построена теоретическая температурная зависимость теплоемкости согласно (2.159); как видно из рисунка, при высоких температурах обе кривые хорошо совпадают, а при низких кривая теоретической теплоемкости убывает более круто, чем это следует из опыта.

Согласно опыту при $T \rightarrow 0$ теплоемкость полупроводников и диэлектриков убывает, как T^3 , в то время как согласно (1.70) она должна убывать экспоненциально. Объяснение этого противоречия принадлежит Дебаю (1912 г.). Дебай обратил внимание на то, что твердое тело нельзя рассматривать как совокупность $3N$ осцилляторов, колеблющихся с одной и той же частотой $\nu = \frac{1}{2\pi} \sqrt{f/m}$, так как

в твердом теле нет независимых молекул, а все они связаны межмолекулярными силами в единое целое. В соответствии с этим, как было показано в предыдущем параграфе, в твердом теле распространяются упругие волны, длина которых варьируется от $2L$ до $2a$, а частоты — в соответствии с рис. 2.11. Число таких волн равно числу степеней свободы твердого тела, и в первом приближении (если пренебречь их ангармоничностью) эти колебания можно считать независимыми.

Если сохранить гипотезу Эйнштейна о квантованных значениях энергии осцилляторов, но в качестве независимых осцилляторов рассматривать эти волны, то энергию кристалла уже нельзя записать в виде

$$\mathcal{E} = \mathcal{E}_v 3N = \frac{3N h \nu}{e^{\frac{h\nu}{kT}} - 1}. \quad (2.160)$$

так как у различных волн частоты, а следовательно, и энергия фононов различны. В этом случае следует провести суммирование по различным частотам:

$$\mathcal{E} = \sum_{\nu_{\min}}^{\nu_{\max}} g(\nu_i) \frac{h\nu_i}{e^{\frac{h\nu_i}{kT}} - 1}. \quad (2.161)$$

где $g(\nu_i)$ — число колебаний с данной частотой ν_i .

Вычисление и экспериментальное нахождение этой суммы представляет для реальных кристаллов весьма большие трудности, и до настоящего времени оно проделано лишь для небольшого числа простейших структур.

Дебай же, для того чтобы решить задачу в общем виде, пошел на упрощение: заменил кристаллическую решетку непрерывной средой (континуумом) и пренебрег различием в частоте продольных и поперечных колебаний; при этом сумма (2.161) превратилась в интеграл:

$$\mathcal{E} = \int_{\nu_{\min}}^{\nu_{\max}} \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1} g(\nu) d\nu, \quad (2.162)$$

где $g(\nu) d\nu$ — число колебаний в интервале частот $d\nu$.

Для непрерывной среды можно показать, что $g(\nu) = Av^2 d\nu$, где A — некоторая постоянная*).

Дополнительные трудности возникают при вычислении верхнего предела интегрирования. Нижний предел в данном случае не играет роли; так как $g(\nu) \rightarrow 0$ при $\nu \rightarrow 0$, то $\nu_{\text{мин}}$ можно приравнять нулю. Сложнее обстоит дело с определением $\nu_{\text{макс}}$ и A . Эти трудности можно преодолеть, опираясь на следующие положения: а) при высоких температурах теплоемкость кристалла должна подчиняться закону Дюлонга и Пти: $C = 3R = 3Nk$, а следовательно, при $T \gg \theta$

$$\int_0^{\nu_{\text{макс}}} \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1} g(\nu) d\nu = 3NkT; \quad (2.163)$$

б) общее число степеней свободы грамм-атома должно быть $3N$:

$$\int_0^{\nu_{\text{макс}}} g(\nu) d\nu = 3N. \quad (2.164)$$

Для дебаевской модели так же, как это было сделано выше для модели Эйнштейна, можно ввести температуру Дебая, определив ее равенством $h\nu_{\text{макс}} = k\theta$ (т. е. как температуру, при которой будут возбуждаться самые высокие частоты) или $\nu_{\text{макс}} = k\theta/h$.

Интеграл (2.162) можно превратить в безразмерный, поделив и умножив его на θ^3 и вынеся все постоянные величины за знак интегрирования. После этого выражение (2.162) с учетом (2.163) и (2.164) принимает вид

$$\mathcal{E} = 3NkTD(t), \quad (2.165)$$

*) В одномерной цепочке каждой длине волны соответствуют три колебания с различным направлением поляризации; в трехмерном случае можно вычислить $g(\nu)d\nu$, рассмотрев число фазовых ячеек dN в слое $dp_\Phi = hdq$ фазового пространства для фононов:

$$dN = \frac{4\pi p_\Phi^2 dp_\Phi}{h^3} = 4\pi q^2 dq = \frac{4\pi \nu^2 d\nu}{\omega_\Phi^3}$$

(так как $q = \nu/\omega_\Phi$); таким образом, $g(\nu) = \frac{4\pi}{\omega_\Phi^3} \nu^2 d\nu$ для каждой ветви колебаний, и $A = 12\pi/\omega_\Phi^3$.

где $D(t)$ — так называемая функция Дебая.

$$D(t) = \frac{3}{t^3} \int_0^t \frac{x^3 dx}{e^x - 1} \quad \text{и} \quad t = \frac{\theta}{T}. \quad (2.166)$$

Выражение (2.166) для любой температуры может быть вычислено приближенно: для $t \ll 1$ (высокие температуры) $D(t) = 1$ и $\mathcal{E} = 3RT$; при $t \gg 1$ (низкие температуры) верхний предел в (2.166) можно заменить на ∞ , при этом:

$$\int_0^{\infty} \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15}, \quad (2.167)$$

$$\mathcal{E} = \frac{3\pi^4 NkT^4}{5\theta^3} \quad (2.168)$$

и

$$C = \frac{\partial \mathcal{E}}{\partial T} \approx \left(\frac{T}{\theta} \right)^3 \quad (2.169)$$

в согласии с экспериментальными данными. Результат, полученный Дебаем, имеет простую качественную интерпретацию. На рис. 1.12 представлены энергетический спектр кристалла по Эйнштейну (слева) и по Дебаю (справа). Как уже упоминалось, в модели Эйнштейна при $T < \theta$ вероятность возбуждения осциллятора убывает экспоненциально. Напротив, в модели Дебая, как бы ни была низка температура, всегда найдутся колебания, для которых $h\nu \ll kT$ и, следовательно, энергия которых $\bar{\epsilon} = kT$; число таких колебаний

$$N' \approx \int_0^{kT} 4\pi\nu^2 d\nu = \frac{4\pi}{3} (kT)^3, \quad (2.170)$$

полная энергия кристалла

$$\mathcal{E} \approx \bar{\epsilon} N' \sim T^4 \quad \text{и} \quad C \sim T^3. \quad (2.171)$$

Более детальная экспериментальная проверка теории Дебая показала, что для простейших кристаллов с одним атомом в элементарной ячейке она дает удовлетворительное согласие с опытом не только при $T \rightarrow 0$, но и в широком интервале температур, из чего можно сделать вывод, что модель Дебая удовлетворительно описывает акустические колебания.

Для кристаллов, содержащих несколько атомов в элементарной ячейке, теория Дебая дает значительно худшее совпадение с опытом. Это и понятно; как видно из рис. 2.12, частота оптических колебаний почти постоянна и, следовательно, их лучше описывать моделью Эйнштейна, чем Дебая. Следовательно, для сложных кристаллов энергию надо представить в виде суммы энергии нулевых колебаний, энергии трех ветвей акустических колебаний, рассчитанной как для непрерывного континуума по Дебаю, и энергии $3(m - 1)$ ветвей оптических колебаний (где m — число атомов в элементарной ячейке, рассчитанной по Эйнштейну):

$$\mathcal{E} = \mathcal{E}_0 + NkT \left[3D\left(\frac{\theta}{T}\right) + \sum_{i=1}^{i=3(m-1)} \frac{\frac{\theta_i}{T}}{e^{\frac{\theta_i}{T}} - 1} \right]. \quad (2.172)$$

Формула (2.172), в которой по существу имеется $3N$ произвольных постоянных (так как ни одна из входящих сюда температур Дебая точно не известна), хорошо согласуется с опытом.

Иногда выходят из положения другим путем — определяют энергию кристалла одной лишь формулой Дебая, но при этом температуру Дебая делают переменной величиной и подбирают θ при каждой температуре таким образом, чтобы теплоемкость, рассчитанная согласно (2.165), соответствовала опытным данным.

ОСНОВНЫЕ ПРЕДСТАВЛЕНИЯ ЭЛЕКТРОННОЙ ТЕОРИИ КРИСТАЛЛОВ

При изложении материала первой главы мы, стремясь к максимальной простоте, избегали применения аппарата квантовой механики, и в результате вынуждены были упустить целый ряд важных вопросов электронной теории, а другие описать чрезвычайно примитивно.

В настоящей главе мы попытаемся частично восполнить этот пробел.

3.1. АДИАБАТИЧЕСКОЕ ПРИБЛИЖЕНИЕ

Как уже было сказано выше, в квантовой механике решение задачи о стационарных состояниях и об энергетическом спектре любой системы частиц может быть принципиально получено путем решения уравнения Шредингера:

$$H\Psi = E\Psi, \quad (3.1)$$

где H — оператор Гамильтона и Ψ — волновая функция от координат всех частиц, вид которой, так же, как разрешенные значения энергии, определяется решением уравнения (3.1) с соответствующими дополнительными условиями (соответствие Ψ граничным условиям, симметрии задачи, однозначность и т. д.). Для кристалла Ψ является функцией от координат всех электронов r_i и атомных ядер R_α , $\Psi = \Psi(r_i, R_\alpha \dots)$, и оператор Гамильтона имеет вид

$$H = \sum_i \frac{1}{2m} p_i^2 + \sum \frac{1}{2M_\alpha} p_\alpha^2 + U(r_i, \dots, R_\alpha), \quad (3.2)$$

где p_i — оператор импульса i -го электрона ($p_i = \hbar/i \partial/\partial r_i$);
 p_α — аналогичный оператор импульса ядра α ;

U — потенциальная энергия взаимодействия всех электронов и ядер.

При подстановке (3.2) в (3.1) последнее примет вид

$$-\frac{\hbar^2}{2m} \sum_i \nabla_i^2 \Psi - \frac{\hbar^2}{2} \sum_\alpha \frac{1}{M_\alpha} \nabla_\alpha^2 \Psi + U(\mathbf{r}_i; \mathbf{R}_\alpha) \Psi = \mathcal{E} \Psi. \quad (3.3)$$

Решение этого уравнения дало бы совершенно исчерпывающие сведения о кристалле: о его структуре (т. е. о распределении плотности вероятности расположения ядер), о распределении электронных плотностей, об энергии системы в основном и возбужденных состояниях и т. д. К сожалению, эта задача безнадежно сложна и приближенное решение ее достигается лишь путем последовательных значительных упрощений.

Первое упрощение состоит в разделении частиц на легкие (валентные электроны) и тяжелые (атомные остатки)*).

Можно предполагать, что в состояниях наиболее вероятных, т. е. соответствующих минимуму свободной энергии, кинетические энергии и тех и других частиц одного и того же порядка, а следовательно, скорости электронов приблизительно на два порядка больше, чем скорости ядер. Если это так, то для любой, даже неравновесной, конфигурации ядер будет успевать устанавливаться соответствующее ей электронное равновесие, т. е. соответствующее данному расположению ядер распределение электронов по орбитам, которое будет медленно меняться вследствие движения ядер; иными словами, это значит, что в каждый данный момент можно рассматривать движение электронов в поле фиксированных ядер. Напротив, за время, пока заметно сместятся ядра, электроны успеют пробежать с соответствующей вероятностью все точки своей орбиты; поэтому движение ядер можно рассматривать не в поле мгновенной конфигурации электронов, а в поле, создаваемом средним пространственным распределением их зарядов. Такое представление исключает возможность обмена энергией между электронной и ядерной системами, поэтому это приближение называется адиабатическим.

*) «Атомные остатки», т. е. ядро вместе со всеми электронами, за исключением валентных, предполагается настолько жестко связанной системой, что она движется как единое целое; в дальнейшем атомный остаток будем сокращенно называть ядром.

При этих условиях движение электронов и ядер можно считать независимым — вероятность определенной конфигурации электронов и ядер равна произведению вероятностей, и волновую функцию системы можно при этом представить в виде произведения электронной и ядерной волновых функций:

$$\Psi(\mathbf{R}_\alpha, \mathbf{r}_i) = \xi(\mathbf{R}_\alpha) \Psi_n(\mathbf{r}_i). \quad (3.4)$$

Подставляя (3.4) в уравнение Шредингера и опуская члены, малые по сравнению с кинетической энергией электронов и слабо зависящие от их координат (см. [12]), можно приближенно разбить уравнение Шредингера (3.3) на два независимых, описывающих движения ядер и электронов; последнее из них будет иметь вид

$$-\frac{1}{2} \frac{\hbar^2}{m} \nabla_i^2 \Psi_n + \frac{1}{2} \sum_i \sum_j \frac{e^2}{r_{ij}} \Psi_n + U(\mathbf{r}_i; \mathbf{R}_\alpha) \Psi_n = \mathcal{E}_n(\mathbf{R}_\alpha) \Psi_n, \quad (3.5)$$

где координаты ядер \mathbf{R}_α уже не переменные, а параметры, от которых зависит потенциальная энергия U системы и собственные значения полной энергии $\mathcal{E}_n(\mathbf{R}_\alpha)$.

К сожалению, при упомянутой выше процедуре придется отбросить члены, среднее значение которых мало, но которые играют существенную роль при коррелированном движении электронов и ионов; между тем в ряде явлений такое движение имеет решающее значение.

Мы здесь лишь перечислим эти явления; подробнее см. [9, 12].

В ионных кристаллах электрон может поляризовать своим полем решетку и таким образом создать для себя локальное так называемое *поляронное* состояние, энергетически более выгодное, чем состояние в зоне проводимости. Если энергия взаимодействия электрона с вырытой им для себя потенциальной ямой будет больше, чем энергия теплового движения, то эта связь будет прочной и число электронов N_{Π} , находящихся в этих (энергетически более выгодных) поляронных состояниях, будет больше, чем в зоне проводимости N_{Θ} :

$$\frac{N_{\Theta}}{N_{\Pi}} \sim e^{-\frac{\mathcal{E}_{\Pi}}{kT}},$$

где \mathcal{E}_{Π} — энергия образования полярона.

В силу трансляционной симметрии решетки энергия \mathcal{E}_{Π} будет одинакова для любого эквивалентного положения полярона в решетке. Это значит, что все поляронные состояния могут объединиться в одно общее и что полярон (т. е. электрон вместе со своей ямой) может перемещаться по всему кристаллу.

Таким образом, образуется зона поляронных состояний, которая: 1) находится под дном зоны проводимости, 2) во много раз уже зоны проводимости и 3) отличается от зоны проводимости тем, что эффективная масса полярона, «везущего» за собой поляризационную «шубу», может быть во много раз больше, чем у зонного электрона.

Наконец, подвижность полярона во много раз меньше, чем у зонного электрона, и может иметь совершенно иную температурную зависимость.

С поляронными состояниями связан другой (т. е. отличный от зонного) механизм электропроводности, получивший название *перескокового*. Чтобы понять этот механизм, представим себе поляронный пакет, т. е. полярон, локализовавшийся вблизи какого-то узла кристаллической решетки. Такое состояние, не обладающее трансляционной симметрией, не будет устойчивым и поэтому для дальнейшей судьбы полярона есть две возможности:

1. Локализованный полярон в результате туннельного эффекта и связанного с ним расплывания волнового пакета «размажется» по всему кристаллу и в дальнейшем будет вести себя аналогично зонному электрону (но с много большей, чем у обычного электрона эффективной массой). Вероятность такого процесса ω_1 в первом приближении не зависит от температуры и характеризуется каким-то эффективным временем τ_1 .

2. В результате тепловых колебаний решетки вблизи одного из соседних ее узлов может образоваться потенциальная яма и полярон без затраты энергии перейдет в эту яму. Вероятность этого процесса будет также характеризоваться некоторым временем τ_2 . Однако так как вероятность образования такой «пустой» потенциальной ямы ω_2 будет экспоненциально расти с температурой

$$\omega_2 \sim e^{-\frac{\mathcal{E}'_{\Pi}}{kT}},$$

где \mathcal{E}'_{Π} — энергия, необходимая для образования этой (пустой) ямы, то τ_2 будет экспоненциально расти с температурой.

Конкуренцией этих двух процессов будет обусловлено совершенно различное поведение полярона. При достаточно низких температурах $\omega_1 \gg \omega_2$ и $\tau_1 \ll \tau_2$, поэтому случайно локализовавшийся полярон за время τ_1 опять превратится в зонный; перемещению зонного электрона мешают тепловые колебания решетки, поэтому подвижность полярона при этих условиях будет падать с ростом температуры.

Напротив, при достаточно высоких температурах ω_2 становится больше, чем ω_1 , и образование локализованных поляронов перестает быть случайностью; за счет роста интенсивности тепловых колебаний очень часто будет образовываться потенциальная яма и поляроны будут большую часть времени находиться в них в локализованном состоянии. Если $\omega_2 \gg \omega_1$, то полярон будет перемещаться (и при диффузии и под действием поля) в основном путем второго (названного перескоковым) механизма и подвижность будет экспоненциально расти с температурой:

$$u \sim e^{-\frac{\mathcal{E}'_{\Pi}}{kT}}. \quad (3.6)$$

Выражение (3.6) лишь по внешнему виду напоминает температурную зависимость подвижности ионов, и отсюда название — перескоковый механизм. В действительности же экспоненциальная зависимость подвижности в данном случае связана с вероятностью образования (по соседству) достаточно глубокой потенциальной ямы, и \mathcal{E}_Π — энергия, необходимая для образования этой ямы (а не высота потенциального горба).

В заключение заметим, что локализованные состояния полярона могут «застревать» вблизи какого-либо дефекта решетки; такой связанный полярон вполне аналогичен локализованному вблизи дефекта электрону.

Ясно, что приведенные выше рассуждения о том, что скорости движения электронов во много раз больше, чем скорости движения ионов, совершенно неприменимы к полярону, который движется как одно целое [9].

Корреляцией в движении электронов и ионов объясняются такие важные явления, как сверхпроводимость, *безызлучательные переходы* и др. Рассмотрение этих вопросов выходит, однако, за пределы настоящей книги.

Таким образом, адиабатическое приближение приводит к потере целого ряда важнейших физических явлений и тем не менее не устраняет основной трудности в решении задачи, остающейся по-прежнему задачей многих тел.

Следующее решающее упрощение и одновременно схематизация ведет к сведению задачи о движении многих электронов к одноэлектронному приближению.

3.2. ОДНОЭЛЕКТРОННОЕ ПРИБЛИЖЕНИЕ

Переход от многоэлектронной к одноэлектронной задаче достигается тем, что в уравнении (3.5) двойная сумма

$\sum_i \sum_j \frac{e^2}{r_{ij}}$ заменяется членом вида $\sum U'(\mathbf{r}_i)$, т. е. мгновенное

взаимодействие электрона со всеми остальными заменяется суммой потенциальных членов, каждый из которых описывает энергию одного из электронов в усредненном поле всех остальных. Если предположить, что заряд каждого электрона «размазан» по всему пространству с плотностью $e |\psi(\mathbf{r})|^2$, можно показать, что

$$U'(\mathbf{r}_i) \approx e^2 \sum_{k'}' \int \frac{|\Psi_{k'}(\mathbf{r}')|^2}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{v}', \quad (3.7)$$

где суммирование проводится по всем электронам, кроме данного (i -го).

Выражение (3.7) носит название потенциала Хартри; этот потенциал нужно добавить к потенциалу атомных остатков $U(\mathbf{r}_i, \mathbf{R}_\alpha)$ в (3.5). Теперь уже движение всех электронов независимо, и мы, следовательно, можем представить волновую функцию всех электронов в виде произведения одноэлектронных функций:

$$\Psi_n = \prod_i \psi_i(\mathbf{r}_i), \quad \text{иск. } \psi_i - \text{выр. в}$$

при этом уравнение (3.5) распадается на N независимых одинаковых уравнений для каждого из электронов:

$$-\frac{\hbar^2}{2m} \nabla^2 \psi_i + U(\mathbf{r}_i) \psi_i = \mathcal{E}_i \psi_i, \quad (3.8)$$

где $U = U(\mathbf{r}_i, \mathbf{R}_\alpha) + U'(\mathbf{r}_i)$ и \mathcal{E}_i — собственное значение энергии i -го электрона ($\mathcal{E}_n = \sum \mathcal{E}_i$).

Далее мы предположим, что ядра совершают малые колебания около равновесных положений, и заменим $U(\mathbf{r}_i, \mathbf{R}_\alpha)$ на $U_0(\mathbf{r}_i, \mathbf{R}_{\alpha 0})$; при этих условиях U будет попросту периодическим потенциалом решетки *), так как электронная плотность должна соответствовать симметрии распределения ядер.

Таким образом, мы свели задачу к движению электронов в периодическом поле $U(\mathbf{r})$ **). К сожалению, нахождение конкретного вида этого периодического потенциала представляет весьма большие трудности. Не говоря о суммировании поля всех атомных остатков в данной точке, остановимся коротко на вычислении $U'(\mathbf{r})$. В выражение (3.7) входит волновая функция $\psi(\mathbf{r}')$, которую мы можем узнать только в результате решения уравнения (3.8). Таким образом, задача должна решаться методом последовательных приближений: выбирается некоторый потенциал U' и находится соответствующая функция $\psi_{U'}$, которая подставляется в (3.7), и таким образом, находится лучшее выражение для потенциала, затем снова находится $\psi_{U'}$, и так до тех пор, пока решение не оказывается «самосогласованным», т. е. полученный потенциал не совпадает, по крайней мере приближенно, с заложенным в уравнение (3.8). Так как указанная процедура связана с чрезвычайно большими

*) В дальнейшем колебания ядер учитываются как малое возмущение, вызывающее рассеяние электронов.

**) В дальнейшем индекс i за ненадобностью будем опускать.

вычислительными трудностями, то зонная структура рассчитывалась описанным выше методом в очень небольшом числе случаев; обычно пользуются одним из четырех приближенных методов (см. ниже).

Приведенные выше рассуждения для обоснования выражения (3.7) самосогласованного потенциала не вполне строгие. В действительности электроны представляют собой не облака зарядов, размазанные по всему пространству, а являются точечными зарядами и их правильные волновые функции должны быть таковы, что вероятность найти два электрона близко друг от друга должна быть много меньше, чем в предположении о независимом их движении. Это вытекает из их кулоновского отталкивания и из принципа Паули или из антисимметричности полной волновой функции системы электронов [9]. Таким образом, если, вводя адиабатическое приближение, мы абстрагируемся от корреляции движения электронов и ядер, то вводя одноэлектронное приближение, мы отвлекаемся от корреляции в движении электронов и опять-таки от нашего внимания при этом ускользает целый ряд явлений, связанных с коррелированным движением электронов:

— дырочная проводимость, представляющая собой коллективное эстафетное движение электронов в почти заполненной зоне *),

— экситонные движения, которые можно рассматривать либо как эстафетное перемещение электронного возбуждения, либо как коллективное перемещение электрона и дырки.

Перейдем к конкретному рассмотрению отдельных методов анализа зонной структуры.

На рис. 3.1, а представлен примерно ход периодического потенциала вдоль одного из кристаллографических направлений ковалентного кристалла, на рис. 3.1, б дана аналогичная картина для кристалла с ионной связью. Как видно из этих рисунков, в обоих случаях ход потенциала слишком сложен для того, чтобы его можно было выразить простой аналитической функцией. Даже если предположить тем не менее, что потенциал точно известен (что в действительности никогда не имеет места) и что вид его можно выразить

*) Строгое обоснование введения дырок можно провести лишь при корректном решении многоэлектронной задачи. В одноэлектронном приближении дырки вводятся «на глаз» без серьезного обоснования.

простой аналитической функцией и подставить в уравнение Шредингера (3.8) для одноэлектронного приближения, то точное решение последнего оказывается все же слишком сложной задачей. Поэтому был разработан ряд приближенных методов решения. При этом оказалось, что некоторые основные закономерности, связанные с периодичностью потенциала, всегда сохраняются независимо от конкретного вида потенциала и метода приближенного решения задачи.

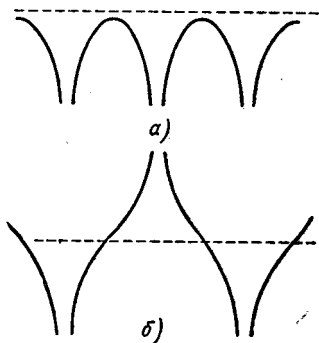


Рис. 3.1. Схема периодического потенциала:
a — в ковалентном кристалле, *b* — в ионном кристалле.

Волновая функция электрона в периодическом поле представляет собой плоскую волну с модулированной амплитудой, причем период модуляций пропорционален периоду решетки *). В зависимости от условий на границах кристалла это может быть бегущая плоская волна

$$\psi = u(\mathbf{r}) e^{2\pi i(\mathbf{k}\mathbf{r} - \nu t)} \quad (3.9)$$

или стоячая волна

$$\psi = u(\mathbf{r}) e^{2\pi i\nu t} \sin 2\pi\mathbf{k}\mathbf{r}. \quad (3.10)$$

Однако в обоих случаях $u(\mathbf{r})$ — периодическая функция с периодом решетки **)

$$u(\mathbf{r} + \mathbf{g}(\mathbf{n})) = u(\mathbf{r}),$$

где

$$\mathbf{g}(\mathbf{n}) = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \quad (3.11)$$

*) Эта теорема, впервые доказанная Блохом, носит его имя. Период модуляции амплитуды может быть равен периоду решетки или в целое число раз меньше него.

**) Или с периодом, меньшим периода решетки в целое число раз.

a_1, a_2, a_3 — базисные векторы решетки и n_1, n_2, n_3 — любые целые числа.

Энергетический спектр электрона всегда состоит из чередующихся полос разрешенных и запрещенных значений энергии; на верхней и нижней границе каждой полосы скорость электрона обращается в нуль, а в промежутке достигает максимального значения.

Эти две особенности волновых функций вытекают из периодичности потенциала (иными словами, трансляционной симметрии кристалла): в идентичных точках прямого и обратного пространства плотность электронного облака и зависимость энергии от волнового вектора должны быть одинаковы.

Но обычно кристаллическая решетка обладает, кроме трансляционной, еще целым рядом элементов симметрии группы вращений. Использование этого обстоятельства позволяет сделать ряд очень важных предсказаний и значительно уменьшить число необходимых вычислений. Поэтому во всех описанных ниже методах применяется теория групп для учета и использования симметрии кристалла.

3.3. ПРИБЛИЖЕНИЕ ПОЧТИ СВОБОДНЫХ ЭЛЕКТРОНОВ

В этом методе за исходные (т. е. нулевое приближение) берутся функции свободного электрона, а затем периодический потенциал учитывается как малое возмущение; поэтому мы вначале рассмотрим волновые функции и энергию свободных электронов.

Модель свободных электронов Зоммерфельда. Простейшее приближение состоит в пренебрежении в уравнении (3.5) всеми членами, за исключением кинетической энергии. В этом случае оно принимает вид

$$\sum_i \frac{\hbar^2}{2m} \nabla_i^2 \Psi(\mathbf{r}_1 \dots \mathbf{r}_n) = \mathcal{E} \Psi(\mathbf{r}_1 \dots \mathbf{r}_n) \quad (3.12)$$

и решение его имеет вид

$$\Psi = \prod_i \varphi_i(\mathbf{r}_i). \quad (3.13)$$

Здесь волновые функции отдельных электронов — плоские

ВОЛНЫ:

$$\varphi = V^{-1/2} e^{2\pi i k r}, \quad (3.14) *$$

где V — объем кристалла.

Энергия электрона

$$\varepsilon_i = \frac{p_i^2}{2m} = \frac{h^2 k_i^2}{2m} \quad (3.15)$$

и энергия всей системы

$$\mathcal{E} = \sum \varepsilon_i. \quad (3.16)$$

Однако волновая функция вида (3.13) не удовлетворяет принципу Паули (она не обращается в нуль), когда два электрона имеют одну и ту же координату. Поэтому для правильного описания системы надо составить линейную комбинацию из таких функций, которые бы удовлетворяли этому принципу и подчинялись статистике Ферми.

Нетрудно убедиться, что этим условиям удовлетворяет волновая функция, изображенная определителем

$$\Psi(1, 2, 3, \dots, N) = (N!)^{-1/2} \begin{vmatrix} \varphi_1(\mathbf{r}_1) & \varphi_1(\mathbf{r}_2) & \dots & \varphi_1(\mathbf{r}_N) \\ \varphi_2(\mathbf{r}_1) & \varphi_2(\mathbf{r}_2) & \dots & \varphi_2(\mathbf{r}_N) \\ \dots & \dots & \dots & \dots \\ \varphi_N(\mathbf{r}_1) & \varphi_N(\mathbf{r}_2) & \dots & \varphi_N(\mathbf{r}_N) \end{vmatrix} \quad (3.17)$$

Во многих случаях, однако, можно удовлетвориться мультипликативной функцией вида (3.13), помня при этом, что в ней может быть не более двух электронов с одинаковыми пространственными волновыми функциями, т. е. с одинаковыми волновыми векторами (но отличающимися направлением спина).

Модель свободных электронов была предложена Зоммерфельдом для описания свободных электронов в металлах и дала качественно удовлетворительные результаты для описания теплоемкости и явлений переноса (электропроводности, термо-э. д. с., гальвано- и термомагнитных явлений) в щелочных металлах. В ряде случаев эта модель

*) Множитель $V^{-1/2}$ берется из условия нормировки — вероятность найти электрон во всем объеме кристалла должна быть равна единице: $\int |\psi|^2 dV = 1$.

может дать удовлетворительную картину для описания свободных электронов в полуметаллах, если заменить в ней массу свободного электрона эффективной массой. Однако эта теория не может дать никаких сведений о зонной структуре твердого тела — для этого необходимо учесть периодический потенциал решетки.

Влияние периодического потенциала. Согласно сказанному выше периодический потенциал решетки может быть разложен в ряде Фурье:

$$U(\mathbf{r}) = \sum_{\mathbf{g}} U_{\mathbf{g}} e^{2\pi i \mathbf{g} \cdot \mathbf{r}}, \quad (3.18)$$

где \mathbf{g} — векторы обратной решетки, и коэффициенты $U_{\mathbf{g}}$ определяются по формуле обращения

$$U_{\mathbf{g}} = \frac{1}{V} \int U(\mathbf{r}) e^{2\pi i \mathbf{g} \cdot \mathbf{r}} dV. \quad (3.19)$$

Вычислим теперь матричные элементы $U(\mathbf{r})$ в системе невозмущенных функций, т. е. плоских волн вида (3.14):

$$U_{\mathbf{k}\mathbf{k}'} = \int \varphi_{\mathbf{k}}^*(\mathbf{r}) U(\mathbf{r}) \varphi_{\mathbf{k}'}(\mathbf{r}) dV. \quad (3.20)$$

Подставив в (3.20) разложение $U(\mathbf{r})$ (3.18), нетрудно убедиться, что матричные элементы отличны от нуля только в случае

$$\mathbf{k} - \mathbf{k}' = \mathbf{g} \quad (3.21)$$

и при этом $U_{\mathbf{k}\mathbf{k}'} = U_{\mathbf{g}}$.

Этих результатов нам достаточно, чтобы построить волновые функции в первом и энергию во втором приближении согласно (2.58) и (2.60):

$$\psi_{\mathbf{k}} = \varphi_{\mathbf{k}} + \sum_{\mathbf{g}} \frac{U_{\mathbf{g}} \varphi_{\mathbf{k}+\mathbf{g}}}{\hbar^2 [k^2 - (\mathbf{k} + \mathbf{g})^2] / 2m} \quad (3.22)$$

и

$$\mathcal{E}_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m} + U_0 + \sum_{\mathbf{g}} \frac{U_{\mathbf{g}}^2}{\hbar^2 [k^2 - (\mathbf{k} + \mathbf{g})^2] / 2m}; \quad (3.23)$$

член

$$U_0 = \int U(\mathbf{r}) |\psi_{\mathbf{k}}|^2 dV$$

есть средняя потенциальная энергия электрона в решетке.

Приближение свободных электронов по своему смыслу применимо тогда, когда периодический потенциал мал по сравнению с кинетической энергией электрона. Это значит, что в каждом члене суммы (3.23) числитель должен быть намного меньше знаменателя, т. е. амплитуда периодического потенциала V_g должна быть мала по сравнению со знаменателем — кинетической энергией электрона. Очевидно, что при этих условиях ψ_k будет мало отличаться от волновой функции свободного электрона, а \mathcal{E}_k — от энергии свободного электрона. Исключением будут те случаи, когда знаменатель в выражениях (3.22) и (3.23) обращается в нуль (или близок к нулю):

$$(\mathbf{k})^2 = (\mathbf{k}' + \mathbf{g})^2. \quad (3.24)$$

При этом мы имеем два вырожденных состояния с одинаковой энергией. В этом случае формулы (3.22) и (3.23) неприменимы и мы должны воспользоваться выражением (2.68). Тогда для определения поправки к энергии нулевого порядка мы должны сначала решить вековое уравнение

$$\begin{vmatrix} \frac{\hbar^2 \mathbf{k}^2}{2m} - \mathcal{E}, & U_g \\ U_g & \frac{\hbar^2 (\mathbf{k} + \mathbf{g})^2}{2m} - \mathcal{E} \end{vmatrix} = 0 \quad (3.25)$$

и только после этого учитывать возмущения более высоких порядков; из (3.25) получим

$$\mathcal{E} = \frac{1}{2} \left\{ \frac{\hbar^2 \mathbf{k}^2}{2m} + \frac{\hbar^2 (\mathbf{k} + \mathbf{g})^2}{2m} \pm \left[\left(\frac{\hbar^2 \mathbf{k}^2}{2m} - \frac{\hbar^2 (\mathbf{k} + \mathbf{g})^2}{2m} \right)^2 + 4 |U_g|^2 \right]^{1/2} \right\}. \quad (3.26)$$

Два корня уравнения (3.26) возникают из двух невозмущенных вырожденных уровней энергии; влияние возмущения в данном случае проявляется в снятии вырождения — расщеплении уровней. Величина расщепления, как видно из (3.26), составляет приблизительно $2U_g$. На рис. 3.2 для одномерного случая представлен энергетический спектр свободного электрона и электрона в твердом теле согласно (3.15). Разрывы в энергетическом спектре появляются всякий раз, когда $\mathbf{k} - \mathbf{k}' \equiv n/a$ и $|k| = |k'|$, т. е. при $k = n/2a$. Расщепления уровней, удовлетворяю-

щих условиях (3.24), на два и образование запрещенных зон качественно интерпретируется следующим образом. Суть дела состоит в том, что состояния A и A' на рис. 3.2 удовлетворяют условию Брегга (для линейного случая $\lambda = 2a$), поэтому они в результате отражения от решетки смешиваются, образуя две стоячие волны.

На рис. 3.3 представлено расположение этих волн и электронной плотности относительно узлов решетки; ясно, что

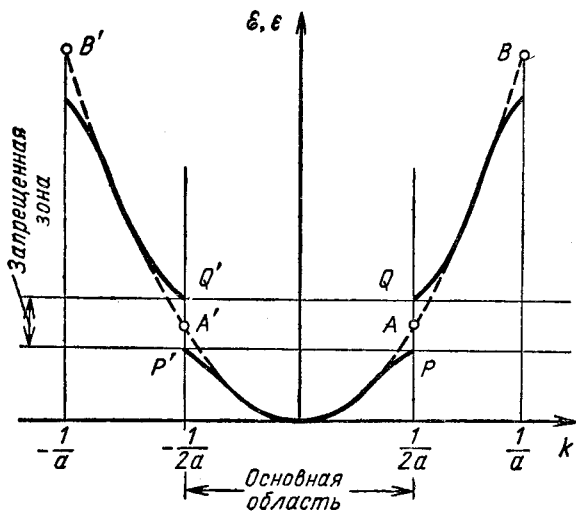


Рис. 3.2. Зависимость энергии от волнового вектора: — — — свободного электрона и — — — электрона в кристалле.

волна P дает энергетически более выгодное состояние, чем волна Q (так как в первом случае максимум электронной плотности соответствует минимуму электростатического потенциала). Таким образом, из двух вырожденных состояний A и A' (рис. 3.2) образуются два невырожденных P и Q . Точка P' (также на рис. 3.2) описывает то же состояние, что и точка P , так как в этой точке та же групповая скорость электрона $v = \partial v / \partial k = -1/h \partial \epsilon / \partial k = 0$ и она описывает такую же стоячую волну с той же энергией и тем же волновым вектором (в случае стоячей волны знак волнового вектора не играет роли). То же самое относится и к точкам Q и Q' .

Предложенное выше качественное объяснение относится и к плоской и объемной решетке с той лишь разницей, что в последнем случае векторы \mathbf{k} и \mathbf{k}' не обязательно совпадают по направлению. Однако и в этом случае разрыв в энергии

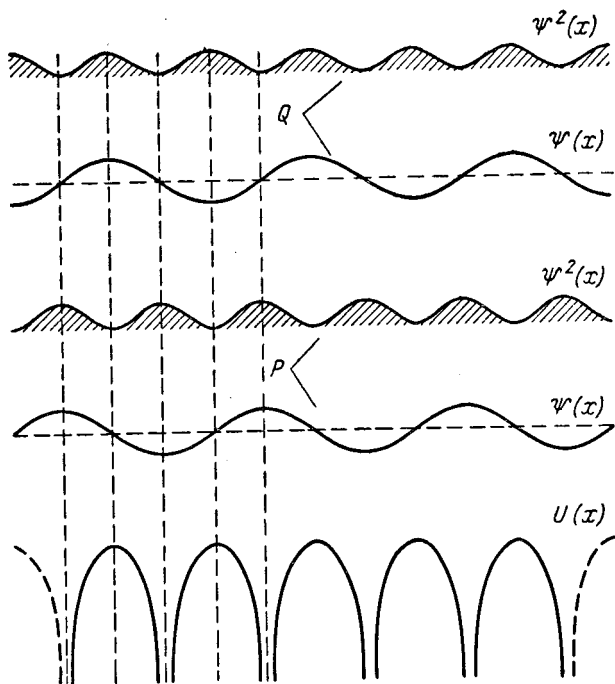


Рис. 3.3. Схема расположения волн электронной плотности, соответствующих верхнему краю валентной зоны (P) и нижнему краю свободной зоны (Q).

электрона наступает тогда, когда выполняются одновременно два условия:

$$\mathbf{k} - \mathbf{k}' = \mathbf{g} \quad \text{и} \quad |\mathbf{k}| = |\mathbf{k}'|,$$

где \mathbf{g} — любой вектор обратной решетки.

На рис. 3.4 представлена обратная решетка и показан вектор \mathbf{k} , удовлетворяющий условию (3.24). Из рисунка видно, что если за начало вектора \mathbf{k}' мы возьмем один из узлов обратной решетки O , то конец его должен лежать на прямой AB , а в пространстве — на плоскости, перпен-

дикулярной одному из векторов (OC) обратной решетки и делящей его в точке B пополам. Таким образом, все пространство обратной решетки делится этими прямыми (а в объеме — плоскостями) на области, внутри которых энергия меняется непрерывно, а на границах терпит разрыв. Эти области носят название зон Бриллюэна.

На рис. 3.5, *a* и *б* представлены первые зоны Бриллюэна для некоторых плоских и объемных решеток. В теории полупроводников часто пользуются схемой приведенных зон

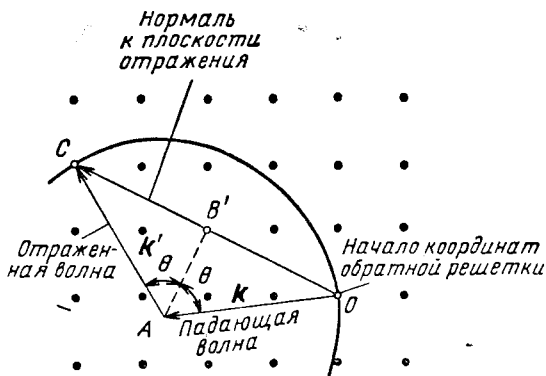


Рис. 3.4. Схема условий интерференции электронных волн.

Бриллюэна; на рис. 3.6 для линейной цепочки построена пунктиром вторая приведенная зона. Такое преобразование энергетического спектра электрона можно обосновать следующим образом: как мы уже упоминали, волновая функция электрона в периодическом поле представляет собой модулированную плоскую волну

$$\psi = u(\mathbf{r}) e^{2\pi i \mathbf{k} \cdot \mathbf{r}}, \quad (3.27) *$$

где $u(\mathbf{r})$ удовлетворяет единственному условию:

$$u(\mathbf{r} + \mathbf{g}) = u(\mathbf{r}).$$

Пусть волновая функция (3.27) соответствует состоянию d на рис. 3.6. Введем обозначение $\mathbf{k} = \mathbf{k}' + \frac{1}{a}$ (где a —

*) Временной множитель $e^{2\pi i \nu t}$ там, где он не нужен, будем опускать.

постоянная линейной решетки) и перепишем (3.27) в виде

$$\psi = u'(\mathbf{r}) e^{2\pi i \mathbf{k}' \cdot \mathbf{r}}, \quad (3.28)$$

где $u'(\mathbf{r}) = u(\mathbf{r}) e^{2\pi i \frac{\mathbf{r}}{a}}$ — также периодическая функция с периодом решетки, но $|\mathbf{k}'| < a/2$; таким образом, точка d' , описывающая электрон в состоянии (3.27), попала в первую зону Бриллюэна. Так как на $u(\mathbf{r})$ мы не накладывали ника-

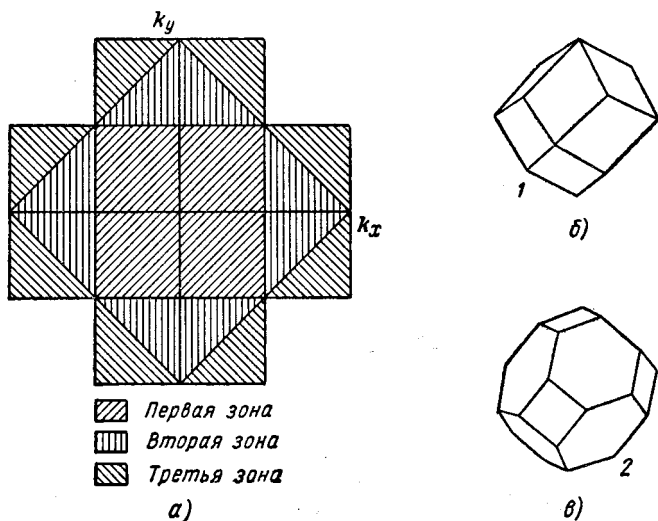


Рис. 3.5. Три первые зоны Бриллюэна для плоской квадратной решетки (а) и первые зоны Бриллюэна для кубической объемноцентрированной (б) и кубической гранецентрированной решеток (в).

ких ограничений, кроме сохраненной в $u'(\mathbf{r})$ периодичности, то оба описания электрона вполне адекватны. Очевидно, что аналогичную процедуру можно провести для любого состояния электрона и в плоском и в объемном случаях.

Как было видно из хода рассуждений, структура зон Бриллюэна ни в какой мере не связана с выбранным нами приближением, а целиком определяется геометрией обратной (а следовательно, и прямой) решетки; что касается ширины разрывов в энергетическом спектре и ширины и формы разрешенных полос, то они связаны с выбранным приближением, и поэтому применимость сделанных в этой области количественных выводов весьма ограничена; так

как в этом методе за нулевое приближение принимается волновая функция свободного электрона, а периодический потенциал учитывается затем как малое возмущение, то отсюда следует, что метод должен давать хорошие результаты в тех случаях, когда полная энергия электрона велика по сравнению с энергией связи с решеткой, т. е. он хорошо применим для электронов, энергия которых намного выше горбов U на рис. 3.1. Поэтому это приближение дает удовлетворительные результаты для возбужденных состояний

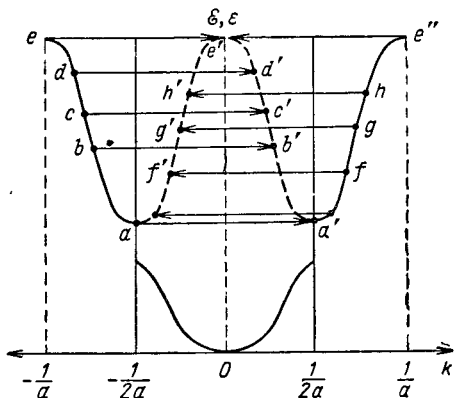


Рис. 3.6. Схема построения приведенных зон Бриллюэна.

(электронов в свободной и более высоких зонах) и в первую очередь для электронных пучков, падающих на кристалл. Для валентных электронов он может дать лишь качественно правильную картину. К этому надо добавить, что для вычисления волновых функций и энергетического спектра здесь необходимо знать Фурье-компоненты периодического потенциала. Практически это чрезвычайно сложная задача, так как вблизи ядер возмущающий потенциал очень велик и очень быстро меняется. Поэтому более удобным для вычислений является метод сильно связанных электронов, в котором за нулевое приближение принимается линейная комбинация волновых функций изолированных атомов (и тем самым автоматически учитывается вид потенциала вблизи ядра), а затем наличие других атомов учитывается как малое возмущение. К рассмотрению этого метода мы и перейдем в следующем параграфе.

3.4. ПРИБЛИЖЕНИЕ СИЛЬНО СВЯЗАННЫХ ЭЛЕКТРОНОВ

Как мы выше упоминали, в методе сильной связи за нулевое приближение берется линейная комбинация атомных функций. Для того чтобы она удовлетворяла теореме Блоха (т. е. ее амплитуда обладала периодичностью решетки), возьмем ее в виде

$$\psi(\mathbf{r}) = \sum_{\mathbf{l}} e^{2\pi i \mathbf{k} \cdot \mathbf{l}} \varphi_a(\mathbf{r} - \mathbf{l}), \quad (3.29)$$

где $\varphi_a(\mathbf{r} - \mathbf{l})$ — атомная волновая функция электрона для атома, находящегося в точке \mathbf{l} .

Энергия электрона в этом состоянии в первом приближении

$$\mathcal{E}_{\mathbf{k}} = \frac{\int \psi_{\mathbf{k}}^*(\mathbf{r}) H \psi_{\mathbf{k}}(\mathbf{r}) d\mathbf{v}}{\int \psi_{\mathbf{k}}^*(\mathbf{r}) \psi_{\mathbf{k}}(\mathbf{r}) d\mathbf{v}}, \quad (3.30)$$

где H — возмущенный гамильтониан (т. е. гамильтониан с учетом периодического потенциала):

$$H = \left[-\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) \right]. \quad (3.31)$$

Выражение (3.31) можно также переписать в виде

$$H = H_0 + [U_{\mathbf{k}}(\mathbf{r}) - U_a(\mathbf{r})] = H_0 + H_1, \quad (3.32)$$

где H_0 — гамильтониан изолированного атома и $H_1 = (U_{\mathbf{k}} - U_a)$ — возмущение, т. е. отличие потенциала в кристалле от атомного потенциала.

Если предположить, что в кристалле заметно перекрываются волновые функции электронов, принадлежащих только соседним атомам, то в (3.30) интегрирование даст члены, отличные от нуля, только для ближайших соседей; таким образом, мы получим

$$\mathcal{E}_{\mathbf{k}} = \mathcal{E}_a - \sum \mathcal{E}(\mathbf{l}) e^{2\pi i \mathbf{k} \cdot \mathbf{l}}, \quad (3.33)$$

где

$$\mathcal{E}(\mathbf{l}) = - \int \varphi_a^*(\mathbf{r} + \mathbf{l}) \{U_{\mathbf{k}}(\mathbf{r}) - U_a(\mathbf{r})\} \varphi_a(\mathbf{r}) d\mathbf{r} \quad (3.34)$$

и \mathcal{E}_a — энергетический уровень, соответствующий $\varphi_a(\mathbf{r})$ в изолированном атоме. Это выражение было впервые получено Блохом. Интеграл (3.34) носит название интеграла

перекрытия. Физически он характеризует вероятность перехода электрона от атома к его соседу. Величина k может меняться в пределах от 0 до $1/2 l$, что соответствует сдвигу по фазе между волновыми функциями соседних атомов от 0 до π .

Различным сдвигам по фазе соответствует различное значение ψ в промежутке между атомами, а следовательно, и различные значения интеграла перекрытия, т. е. раз-

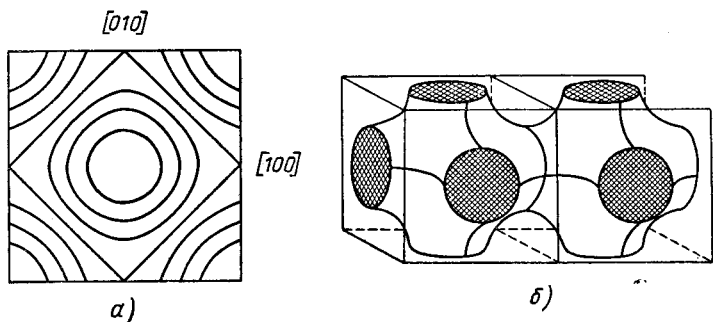


Рис. 3.7. Изоэнергетические поверхности для плоской квадратной (а) и объемной кубической (б) решетки.

личные уровни энергии. Таким образом атомные волновые функции расщепляются в зону *).

Рассмотрим в качестве примера простую кубическую решетку и предположим, что $\varphi_a(\mathbf{r})$ представляет собой s -состояние, т. е. сферически симметрична. Тогда, если ограничиться учетом в (3.30) только ближайших соседей, мы получим шесть одинаковых членов, т. е. \mathcal{E}_k будет иметь вид

$$\mathcal{E}_k = \mathcal{E}_a - \mathcal{E}(0) + 2\mathcal{E}(100) \{ \cos ak_1 + \cos ak_2 + \cos ak_3 \}. \quad (3.35)$$

На рис. 3.7, а и б представлены изоэнергетические поверхности, построенные согласно (3.35) для плоской квадратной и объемной кубической решеток.

В заключение упомянем два других приближенных метода нахождения волновых функций и энергетического спек-

*) Следует отметить, что из-за условия нормировки изменение значения ψ в промежутке между атомами влечет за собой соответствующее изменение ψ вблизи атома, что в свою очередь вызывает смещение атомных потенциалов \mathcal{E}_a .

тра электронов в кристалле в одноэлектронном приближении:

Метод ячеек. Кристалл разбивается на одинаковые многогранники плоскостями, делящими межатомные расстояния пополам; уравнение Шредингера решается для одного многогранника (ячейки), который заменяется сферой того же объема; полагая $U = -\frac{e^2}{r}$ и накладывая граничное условие, чтобы волновая функция и ее производная менялись непрерывно при переходе из одной ячейки в другую, решается уравнение Шредингера для этой ячейки.

Метод ПГВ — присоединенных плоских волн. Как видно из рис. 3.3, потенциал очень слабо меняется в промежутке между атомами, поэтому здесь можно достаточно точно описать волновую функцию плоской волной, а внутри сферы, окружающей ядро, — атомной волновой функцией и затем «сшить» на границе, т. е. наложить условие непрерывности на функцию и ее первую производную.

Существуют и другие приближенные методы, дающие, как и два предыдущих, достаточно точные результаты. При этом качественно результат для электронов в кристалле получается всегда один и тот же, энергетический спектр электрона состоит из чередующихся полос разрешенных и запрещенных значений энергии. На верхней и нижней границе каждой разрешенной полосы скорость электрона обращается в нуль, а в промежутке достигает максимума; более подробно поведение электрона в кристалле мы рассмотрим в следующем параграфе.

3.5. ОСНОВНЫЕ ОСОБЕННОСТИ СТРУКТУРЫ ЭНЕРГЕТИЧЕСКИХ ЗОН ПОЛУПРОВОДНИКОВ

ВИДЫ ЭНЕРГЕТИЧЕСКИХ ЗОН

Выше уже кратко упоминалось, что непосредственно из трансляционной симметрии кристалла вытекает ряд важных выводов относительно вида волновых функций и энергетического спектра электрона в кристалле. Остановимся на этом вопросе подробнее.

Энергетический спектр электрона в периодическом поле состоит из серии n квазинепрерывных полос (зон), разделенных запрещенными промежутками.

Волновые функции зонных электронов представляют собой модулированные плоские волны

$$\psi_n^k = u_n^k(\mathbf{r}) e^{2\pi i \mathbf{k} \mathbf{r}}, \quad (3.36)$$

где амплитуда модуляции зависит от вида периодического потенциала, номера зоны (n) и величины волнового вектора \mathbf{k} и является периодической функцией координат:

$$u_n^k(\mathbf{r} + \mathbf{g}_n) = u_n^k(\mathbf{r}); \quad (3.37)$$

\mathbf{g}_n — любой вектор решетки, т. е.

$$\mathbf{g}_n = \mathbf{a}_1 n_1 + \mathbf{a}_2 n_2 + \mathbf{a}_3 n_3, \quad (3.38)$$

где $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_3$ — базисные векторы и n_1, n_2, n_3 — любые целые числа.

Из (3.37) непосредственно следует, что амплитуда $u(\mathbf{r})$ для любого направления периодична с периодом решетки или с периодом, меньшим периода решетки в целое число раз.

Из периодичности амплитуды $u_n^k(\mathbf{r})$ как функции от \mathbf{r} в прямом пространстве непосредственно вытекает еще более важный вывод относительно периодичности энергии электрона для каждой (n -й) зоны в обратном пространстве, а именно: энергия должна быть периодической функцией от \mathbf{k} с периодом, равным периоду обратной решетки (или меньшим его в целое число раз).

Прямая решетка не только обладает трансляционной симметрией, но и принадлежит к какой-либо из точечных групп симметрии. Для того чтобы состояние электрона было устойчивым, волновая функция ψ_n^k должна быть симметричной или антисимметричной по отношению к определенным преобразованиям симметрии*).

Обратная решетка также обладает соответствующими элементами симметрии и из этого непосредственно следует, что энергия электрона должна быть симметричной относительно соответствующих преобразований (мы пока не будем рассматривать случай вырожденных зон).

Это значит, что энергия должна иметь одинаковое значение в эквивалентных точках обратной решетки, а на

*) Симметрия волновых функций может быть ниже, чем атома или кристалла. Подробнее см. [7].

середине отрезка, соединяющего эти точки, должна проходить через экстремум. Приведенные выше соображения позволяют сделать ряд выводов о возможных структурах зон любого кристалла.

Напомним, что каждый элемент симметрии (центр, ось, плоскость) размножает точки пространства, не принадлежащие этому элементу. Так, например, в элементарной ячейке плоской обратной квадратной решетки произвольная точка имеет восемь эквивалентных точек, точка на диагоналях — четыре и т. д. и лишь центр элементарной ячейки не имеет (в ее пределах) эквивалентных точек.

Отсюда вытекает ряд важных заключений.

1. В центре элементарной ячейки обратной решетки обязательно должен быть экстремум (максимум или минимум) энергии. Однако в полупроводнике зоны обычно либо почти пусты, либо почти заполнены, и нас поэтому интересуют главным образом абсолютные максимумы и минимумы. Если этот экстремум не является абсолютным, то мы можем «не заметить» его, т. е. он не будет отражаться на электрофизических свойствах полупроводника. Экстремум, находящийся в центре зоны Бриллюэна кубической решетки, должен обладать шаровой симметрией; в этом случае эффективная масса будет скалярной. Если экстремум лежит на одной из диагоналей, то таких экстремумов должно быть в плоской квадратной решетке 4, а в объемной кубической — 8. Если экстремум в объемной кубической ячейке лежит на направлении (110) , то таких экстремумов должно быть 6, и т. д.

2. Точечная подгруппа симметрии бесконечное число раз размножается трансляционной симметрией; этим обусловлена, как мы уже упоминали, периодичность волновой функции в обычном пространстве и энергии как функции от волнового вектора в обратном k -пространстве; именно поэтому мы обычно имеем возможность ограничиться рассмотрением первой зоны Бриллюэна.

При этом следует, однако, учитывать, что если экстремумы лежат на границах зоны Бриллюэна, то число их уменьшается вдвое, так как вторая эквивалентная половина экстремума попадает в следующую зону.

3. До сих пор мы рассматривали простые, невырожденные зоны; этот случай, однако, редко реализуется на практике. Очень часто в одной точке k пространства мы имеем несколько экстремумов, энергия которых в нулевом при-

ближении одинакова. Эти несколько экстремумов могут иметь различное происхождение:

— два экстремума могут появляться за счет подзон, соответствующих противоположным направлениям спина;

— если атомный уровень, из которого образовалась данная зона, не обладал шаровой симметрией (а шаровой симметрией обладают только s -состояния), то различной ориентации исходной атомной функции будут соответствовать различные зоны; таким образом образуется три p -зоны, пять d -зон и т. д.

Каждая из этих зон еще может расщепиться на две подзоны, соответствующие различным направлениям спина.

4. Энергия перечисленных выше подзон обычно бывает одинакова лишь в нулевом приближении. Если учесть возмущения (периодическое поле решетки и др.), то они обычно расщепляются — смещаются (по энергиям) друг относительно друга:

— зоны, соответствующие различным направлениям спина, смещаются относительно друг друга за счет спин орбитального взаимодействия, т. е. за счет того, что различным направлениям спина в магнитном поле орбиты соответствует различная энергия;

— зоны, соответствующие различной ориентации орбит, смещаются за счет различного взаимодействия с кристаллическим полем решетки и, в частности, за счет различного взаимодействия с ближайшими соседями. Поэтому зоны не только смещаются, но имеют различную ширину. Так, например, p -функции атомов углерода, направленные к соседним идентичным атомам углерода в органических молекулах, образуют прочные так называемые σ -связи, соответствующая зона лежит более низко и относительно широка, так как перекрытие волновых функций в этом случае велико, соответственно должна быть мала эффективная масса носителей.

Напротив, если p -функции расположены перпендикулярно линии, соединяющей атомы (π -связь), перекрытие волновых функций меньше, связь слабее, соответствующая зона должна быть уже и расположена выше. В различных кристаллах смещение и ширина зон, соответствующих различной ориентации орбит, может колебаться в весьма широких пределах.

5. В ряде случаев симметрия валентных орбит атомов, из которых построен кристалл, не соответствует симметрии

расположения ближайших соседей, и поэтому такие состояния не являются энергетически наиболее выгодными. В этом случае образуются гибридные состояния. Так, например, в алмазе и подобных решетках образуются *sp*-гибриды, волновые функции в этом случае обладают симметрией тетраэдра и образуют прочные ковалентные связи с соседними атомами.

ЭФФЕКТИВНАЯ МАССА

Как мы уже упоминали, экстремум, находящийся в центре зоны Бриллюэна кубического кристалла, должен обладать шаровой симметрией. Предположим, что это минимум, и разложим энергию вблизи него по степеням k :

$$\mathcal{E} = \mathcal{E}_0 + \frac{1}{2} \left(\frac{\partial^2 \mathcal{E}}{\partial k_x^2} k_x^2 + \frac{\partial^2 \mathcal{E}}{\partial k_y^2} k_y^2 + \frac{\partial^2 \mathcal{E}}{\partial k_z^2} k_z^2 \right). \quad (3.39)$$

В рассматриваемом нами случае кубического кристалла

$$\frac{\partial^2 \mathcal{E}}{\partial k_x^2} = \frac{\partial^2 \mathcal{E}}{\partial k_y^2} = \frac{\partial^2 \mathcal{E}}{\partial k_z^2} = \frac{\partial^2 \mathcal{E}}{\partial k^2}. \quad (3.40)$$

Если учесть это и, кроме этого, отсчитывать энергию от нижнего края зоны, то выражение (3.39) примет вид

$$\mathcal{E} = \frac{1}{2} \frac{\partial^2 \mathcal{E}}{\partial k^2} k^2. \quad (3.41)$$

Для свободного электрона связь энергии и импульса имеет вид

$$\mathcal{E} = \frac{1}{2} \frac{p^2}{m_0} \quad (3.42)$$

или, если учесть соотношение де Бройля ($p = \hbar k$), то (3.42) можно переписать в виде, аналогичном (3.41):

$$\mathcal{E} = \frac{1}{2} \frac{\hbar^2}{m_0} k^2. \quad (3.43)$$

Для свободного электрона в отсутствие внешних сил (электрического поля) импульс и энергия являются интегралами движения. Совсем по-иному обстоит дело для электрона, находящегося в периодическом поле кристалла. В этом случае и скорость, и импульс электрона меняются от точки к точке в весьма широких пределах, однако если учесть периодический характер потенциала, то непосредственно из закона сохранения энергии вытекает, что сред-

ние значения этих величин также сохраняют в отсутствие внешних полей постоянные значения.

Учитывая это, можно и для зонного электрона ввести по аналогии со свободным электроном понятие квазиимпульса, определив его соотношением

$$p = \hbar k. \quad (3.44)$$

Для того чтобы сделать выражение (3.41) подобным (3.42), можно также ввести понятие эффективной массы, определив ее соотношением

$$\frac{\hbar^2}{m} = \frac{\partial^2 \mathcal{E}}{\partial k^2} \quad \text{или} \quad m = \frac{\hbar^2}{\partial^2 \mathcal{E} / \partial k^2}. \quad (3.45)$$

При этом выражение (3.41) приобретает вид, совершенно аналогичный (3.42):

$$\mathcal{E} = \frac{1}{2} \frac{p^2}{m}. \quad (3.46)$$

Можно также показать, что если определить эффективную массу соотношением (3.45), то выражение для связи производной от средней скорости зонного электрона с полем приобретает вид, совершенно аналогичный второму закону Ньютона:

$$\frac{d\bar{v}}{dt} = \frac{1}{m} e\bar{E}, \quad (3.46a)$$

где \bar{E} — среднее (внешнее) поле.

Соотношения (3.44), (3.45) и (3.46) позволяют ввести понятие эффективной массы и широко использовать его в теории полупроводников.

Из приведенного выше определения (3.45) непосредственно вытекает, что эффективная масса совершенно не «обязана» быть равной массе свободного электрона.

Как мы уже упоминали в гл. 1, это неравенство является следствием того, что различным уровням в зоне соответствует не только различная кинетическая энергия, но и потенциальная, и под действием внешних сил меняются оба вида энергии.

Однако учет этого обстоятельства (т. е., иными словами, учет периодического потенциала) путем введения эффективной массы оказывается чрезвычайно плодотворным и чрезвычайно облегчает рассмотрение очень большого круга вопросов теории полупроводников.

Если в центре зоны расположен не минимум (как в рассмотренном выше случае), а максимум, то выражение для энергии примет вид, аналогичный (3.39):

$$\mathcal{E} = \mathcal{E}_m - \frac{1}{2} \left[\frac{\partial^2 \mathcal{E}}{\partial k_x^2} k_x^2 + \frac{\partial^2 \mathcal{E}}{\partial k_y^2} k_y^2 + \frac{\partial^2 \mathcal{E}}{\partial k_z^2} k_z^2 \right]. \quad (3.47)$$

Для зоны, обладающей шаровой симметрией, оно также приводится к виду

$$\mathcal{E} = \frac{p^2}{2m}, \quad (3.48)$$

если ввести обозначение

$$m = - \frac{\hbar^2}{\frac{\partial^2 \mathcal{E}}{\partial k^2}} \quad (3.49)$$

и энергию отсчитывать от верхнего края зоны; выражение для производной от средней скорости в этом случае имеет вид

$$\hbar^2 \frac{d\bar{v}}{dt} = - \frac{\partial^2 \mathcal{E}}{\partial k^2} e\bar{E}, \quad (3.50)$$

т. е., используя те же обозначения, получим выражение, аналогичное второму закону Ньютона:

$$\frac{d\bar{v}}{dt} = \frac{\bar{F}}{m}. \quad (3.51)$$

Таким образом, электрон вблизи верхнего края зоны ведет себя как частица с отрицательной массой, определяемой соотношением (3.49) (и, как всегда, с отрицательным зарядом).

Можно показать, что в случае почти заполненной зоны суммарный эстафетный механизм перемещения электронов, последовательно замещающих пустое место, эквивалентен перемещению одной частицы с положительным зарядом и положительной массой по абсолютной величине, определяемой соотношением (3.49).

Таким образом, вводится понятие «дырка», которая в этом случае ведет себя во всех отношениях так же, как электрон в почти пустой зоне.

Проведенные выше рассуждения можно обобщить на экстремумы, не обладающие шаровой симметрией. Если речь идет не о кубическом кристалле и мы имеем один минимум в центре зоны, то можно направить оси таким

образом (по главным осям эллипсоида энергии), чтобы выражение (3.39) по-прежнему было справедливым, но, разумеется, в этом случае равенство (3.40) уже не будет иметь места. Можно обобщить метод эффективной массы на этот случай, введя обозначения:

$$\frac{\partial^2 \mathcal{E}}{\partial k_x^2} = \frac{\hbar^2}{m_{xx}}, \quad \frac{\partial^2 \mathcal{E}}{\partial k_y^2} = \frac{\hbar^2}{m_{yy}}, \quad \frac{\partial^2 \mathcal{E}}{\partial k_z^2} = \frac{\hbar^2}{m_{zz}}. \quad (3.52)$$

Тогда выражение (3.39) принимает вид (полагая $\mathcal{E}_0 = 0$)

$$\mathcal{E} = \frac{1}{2} \left(\frac{p_x^2}{m_{xx}} + \frac{p_y^2}{m_{yy}} + \frac{p_z^2}{m_{zz}} \right). \quad (3.53)$$

Таким образом, эффективная масса становится тензором. В общем случае, если экстремум лежит не в центре зоны Бриллюэна, а в какой-либо другой точке, то таких экстремумов будет столько, сколько эквивалентных точек, и нет никаких оснований полагать, что они будут обладать шаровой симметрией. Мы можем, однако, придать выражению для энергии вид, аналогичный (3.39), расположив начало координат в точке экстремума (это нельзя, однако, сделать одновременно для всех экстремумов). Если он расположен на одной из осей симметрии более чем второго порядка, то он должен иметь вид эллипсоида вращения; число таких эллипсоидов будет равно числу эквивалентных осей. В этих случаях можно ввести продольную и поперечную эффективные массы, и если ось z , например, направить по главной оси одного из эллипсоидов, то выражение для энергии в нем примет вид

$$\mathcal{E} = \frac{1}{2} \left(\frac{p_x^2 + p_y^2}{m_{\perp}} + \frac{p_z^2}{m_{\parallel}} \right). \quad (3.54)$$

Следует, однако, заметить, что редко можно направить оси координат таким образом, чтобы они были главными осями для всех эллипсоидов, и очень часто приходится поэтому пользоваться выражением для энергии, не приведенным к главным осям:

$$\mathcal{E} = \mathcal{E}_0 + \frac{\hbar^2}{2} \sum_{xyz} \frac{k_x k_y}{m_{xy}} + \dots, \quad (3.55)$$

где тензор $1/m_{xy}$ носит название тензора эффективных масс, компоненты которого определяются соотношениями

$$\frac{1}{m_{xy}} = \hbar^2 \frac{\partial^2 \mathcal{E}}{\partial k_x \partial k_y} \text{ и т. д.}$$

СТАТИСТИКА ЭЛЕКТРОНОВ В ПОЛУПРОВОДНИКАХ

4.1. НЕКОТОРЫЕ ПОНЯТИЯ СТАТИСТИКИ И ТЕРМОДИНАМИКИ

Материал последующих глав не может быть понят без знания распределения электронов по энергии. Поэтому мы начинаем с рассмотрения некоторых положений термодинамики и физической статистики, частично знакомых читателю из общего курса физики.

Рассмотрим какую-либо физическую систему, например некоторое количество N молекул или ионов газа, заключенных в объем V .

Можно себе представить такое состояние системы, когда все молекулы или ионы будут находиться в небольшой части всего объема ΔV . Это, действительно, будет в самый первый момент, если мы впустим газ в объем через небольшое отверстие в одной из стенок. Но через некоторое время молекулы распределятся равномерно по всему объему. В принципе можно себе представить, что в дальнейшем они опять «случайно» соберутся в этом или каком-либо другом маленьком объеме ΔV . Однако при большом числе молекул (N) вероятность (ω) такого «случая» будет чрезвычайно мала:

$$\omega \sim \left(\frac{\Delta V}{V} \right)^N.$$

Например, приняв $N \approx 10^{22}$ (число молекул газа в объеме 1 л при нормальных условиях) и $\Delta V = 0,1V$, мы получим $\omega \sim 10^{-22}$, а вероятность равномерного распределения по всему объему будет практически равна единице. Таким

образом, мы можем утверждать, что через некоторое время после «впуска» газ распределится равномерно по всему объему и будет находиться в этом состоянии бесконечно долго *).

То же самое относится и к распределению молекул или ионов газа по скоростям. Если, например, газ состоит из одинаковых ионов и мы пропустим их через ускоряющее поле, то в первый момент все ионы могут иметь одинаковую по величине и направлению скорость. Но через некоторое время в результате столкновений со стенками и друг с другом ионы приобретут всевозможные по величине и направлению скорости. Мы можем опять-таки представить, что на какой-то момент «случайно» скорости ионов станут одинаковыми, но вероятность этого будет исчезающе мала.

После этих предварительных замечаний можно ввести некоторые определения статистики и термодинамики.

1. Логарифм вероятности данного состояния системы (т. е. данного распределения молекул газа по объему и по энергиям) $\ln \omega$, умноженный на постоянную Больцмана k , называется энтропией системы и обозначается буквой S ; $S = k \ln \omega$.

2. Состояние, в которое через некоторое время приходит и в котором остается как угодно долго система, находящаяся в постоянных внешних условиях, называется состоянием статистического (или термодинамического) равновесия. Время τ , в течение которого устанавливается это состояние, называется временем релаксации системы.

Данное выше определение времени релаксации весьма неточно. В действительности система возвращается в равновесное состояние постепенно; чем больше отступление от равновесия, тем скорее оно «рассасывается», но по мере возвращения к равновесию скорость приближения к нему все уменьшается, и потому к «идеальному» равновесию система придет лишь через бесконечно большой промежуток времени **).

Распределение частиц по объему (т. е. координатам), скоростям и энергиям можно охарактеризовать некоторой функцией, дающей вероятность нахождения частиц в данной точке фазового пространства (т. е. с данными координатами x, y, z и скоростями v_x, v_y, v_z); эта функция называется функцией распределения.

*) Мы здесь не говорим о небольших флуктуациях плотности и распределения по скоростям в небольших объемах, которые всегда будут иметь место.

**) При этом и в предельном «идеальном» равновесии все время будут возникать небольшие флуктуации — отступления от него.

В ряде случаев скорость изменения функции распределения f прямо пропорциональна отступлению ее от равновесного состояния f_0 :

$$\frac{df}{dt} = \frac{(f_0 - f)}{\tau}, \quad (4.1)$$

откуда

$$f = f_0 + (f - f_0) e^{-\frac{t}{\tau}}. \quad (4.2)$$

При этих условиях время релаксации τ можно определить более точно: это время, в течение которого отступление функции распределения от равновесного состояния убывает в e раз. В тех случаях, когда соотношения (4.1) и (4.2) не выполняются, введение понятия времени релаксации встречает большие затруднения; определив, например, его соотношением

$$\tau = \frac{df}{dt} / (f_0 - f),$$

мы получим для каждого момента времени другой результат.

Наконец, в ряде случаев время релаксации не только зависит от величины отступления от равновесия, но и от его конкретного вида — это еще более сложный случай для теоретического анализа. Таким образом, иногда время релаксации приходится вводить с большими оговорками. Тем не менее это понятие играет очень большую роль в теоретической физике и, в частности, в физике твердого тела.

Как мы уже упоминали, равновесное распределение по скоростям и энергиям устанавливается в результате столкновений. Поэтому время релаксации должно быть связано с вероятностью столкновений или числом их в единицу времени.

Действительно, теория показывает, что в простейших случаях время релаксации пропорционально, а в первом приближении просто равно среднему времени между столкновениями для данного сорта частиц. При этом следует, разумеется, учитывать столкновения рассматриваемых нами частиц (например, электронов) не только между собой, но и с другими частицами (атомами, ионами), квазичастицами (фотонами, фононами), дефектами и др. Более того, во многих случаях столкновения рассматриваемых частиц между собой играют менее важную роль, чем столкновения с другими частицами. Так, например, при наличии электрического поля столкновения электронов друг с другом не могут уменьшить их суммарный импульс, а могут только перераспределить его между собой; поэтому, если бы существовали только эти столкновения, электропроводность материала была бы бесконечна.

Точно так же нормальные столкновения фононов друг с другом не дают ограниченной теплопроводности, а приводят к перераспределению импульса между фононами и, таким образом, лишь косвенно влияют на теплопроводность.

В первом приближении можно для вычисления времени релаксации системы складывать все числа соударений между собой,

считая, что они дают одинаковый эффект:

$$\bar{\tau} = \frac{1}{\sum_i \nu_i} \approx \frac{1}{\sum_i \bar{\nu}_i} \quad (4.3)$$

или

$$\bar{\tau}^{-1} = \sum_i \bar{\nu}_i, \quad (4.4)$$

где $\sum_i \bar{\nu}_i$ — общее (среднее) число столкновений в единицу времени, а $\bar{\nu}_i$ — среднее число столкновений данного сорта (электронов с дефектами, электронов с фононами и т. д.). При наличии только одного вида столкновений

$$\bar{\tau}_i^{-1} = \bar{\nu}_i.$$

Выражение (4.4) можно также переписать в виде

$$\bar{\tau}^{-1} = \sum_i \bar{\tau}_i^{-1}. \quad (4.5)$$

Таким образом, соотношения (4.3) позволяют нам вычислить время релаксации для каждого конкретного случая, если известны $\bar{\nu}_i$. Однако вычисление ν_i встречает весьма большие трудности. В принципе в квантовой механике эти величины вычисляются через вероятности переходов частицы из одного стационарного состояния в другое под действием возмущения — взаимодействия с другой частицей.

Согласно сказанному выше, в состоянии равновесия энтропия изолированной системы достигает своего максимального значения.

Если будут изменяться внешние условия, в которых находится система (объем, давление и т. п.), то будет меняться и ее состояние: равновесное для одних внешних условий, оно не будет таким для других. Например, мы можем представить, что объем, занимаемый газом, резко увеличился и, таким образом, в первый момент часть нового объема оказалась пустой и лишь через некоторое время (время релаксации) газ приблизится к равномерному распределению по всему увеличенному объему.

Но можно себе представить и такой процесс, когда внешние параметры меняются достаточно медленно, чтобы система успевала приходить в равновесие во всех промежуточных состояниях между начальным и конечным; такой процесс называется квазистатическим.

В термодинамике также различают адиабатические, изотермические и политропические процессы. Адиабатическим мы называем такой процесс, во время которого система не

обменивается теплом с окружающей средой ($dQ = 0$). Этого можно достичь, например, теплоизоляцией или настолько быстрым изменением параметров, чтобы теплообмен, происходящий только через поверхность тела, практически не играл роли.

При изотермическом процессе температура системы поддерживается постоянной ($dT = 0$, $T = \text{const}$, $dQ \neq 0$), при политропическом меняется и теплосодержание и температура системы ($dQ \neq 0$, $dT \neq 0$). Разумеется, любой из этих процессов может быть и квазистатическим и неквазистатическим.

В состоянии равновесия внутренняя энергия системы \mathcal{E} определяется ее температурой и внешними условиями (параметрами): объемом, давлением, электрическим потенциалом, магнитным полем и т. д.

Если мы заставим систему совершать работу W (изменив, например, ее объем) и при этом оградим ее от окружающей среды теплонепроницаемыми стенками (адиабатические условия), то из закона сохранения энергии следует, что совершенная работа будет равна убыли внутренней энергии:

$$dW = -d\mathcal{E}.$$

Если же система обменивается теплом с окружающей средой, то она может совершать работу как за счет убыли внутренней энергии, так и за счет теплоты, полученной извне:

$$dW = -d\mathcal{E} + dQ \text{ или } d\mathcal{E} = dQ - dW. \quad (4.6)$$

В термодинамике доказывается, что теплота, получаемая системой при квазистатическом процессе, связана с возрастанием ее энтропии соотношением

$$dQ = T dS. \quad (4.7)$$

Следовательно, выражение (4.6) может быть переписано в виде

$$dW = -(d\mathcal{E} - T dS) = -d(\mathcal{E} - TS), \text{ при } T = \text{const}. \quad (4.8)$$

Таким образом, если работа теплоизолированной системы равна уменьшению ее внутренней энергии, то работа, совершаемая системой, обменивающейся теплом с окружающей средой (поддерживаемой благодаря этому

при постоянной температуре), равна убыли функции F :

$$F = \mathcal{E} - TS \quad (4.9)$$

(которая получила название свободной энергии).

Иными словами, если переписать (4.9) в виде

$$\mathcal{E} = F + TS, \quad (4.10)$$

то можно сказать, что энергия системы \mathcal{E} состоит из двух частей: свободной энергии («полезной») F , которую при изменении внешних условий (но при постоянной температуре) можно превратить в работу, и связанной («бесполезной») энергии TS , переходящей в тепло.

В этом смысле свободная энергия играет ту же роль для системы, находящейся в изотермических условиях, что полная — для адиабатических условий.

Согласно (4.6) внутренняя энергия системы \mathcal{E} может изменяться за счет совершенной работы или обмена теплом. Энергия системы может меняться также за счет изменения числа частиц (N) в ней (например, энергия полупроводника уменьшится при уходе электрона из него в металл или в вакуум). Следовательно, в более общем случае закон сохранения энергии (4.6) примет вид

$$d\mathcal{E} = dQ - dW + \mu dN, \quad (4.11)$$

где μ — некоторый коэффициент, равный изменению энергии системы на одну частоту*). В случае квазистатического процесса $dQ = T dS$ и

$$d\mathcal{E} = T dS - dW + \mu dN. \quad (4.12)$$

Величина μ в (4.12) называется химическим потенциалом (по аналогии с электростатическим потенциалом — изменение электростатической энергии равно произведению электростатического потенциала на изменение заряда). Вычитая из левой и правой частей (4.12) дифференциал связанной энергии

$$d(TS) = T dS + S dT,$$

получаем

$$dF = -S dT - dW + \mu dN. \quad (4.13)$$

*) Если при этом $dQ = 0$ (нет обмена теплом), и $dW = 0$ (не совершается работа).

Из (4.13) следует, что

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{V, T = \text{const}} \quad (4.14)$$

при $dT = 0$ и $dW = 0$, т. е. согласно (4.14) химический потенциал есть свободная энергия на одну частицу при определенной температуре (условие $T = \text{const}$) и определенных значениях внешних параметров (условие $dW = 0$ означает постоянство внешних параметров). Следовательно,

$$\mu = \frac{F}{N} = \bar{\mathcal{E}} - Ts, \quad (4.15)$$

где $\bar{\mathcal{E}}$ — энергия, приходящаяся на одну частицу (т. е. средняя энергия) и s — энтропия на одну частицу (удельная энтропия *).

В заключение этого раздела остановимся кратко на условиях термодинамического равновесия. Мы уже упоминали, что для адиабатически изолированной системы условием равновесия является максимум энтропии. Для системы, находящейся в термостате (в изотермических условиях), условием равновесия является минимальное значение свободной энергии. Последний и наиболее важный для нас случай (изотермические условия) позволяет сразу же вывести условия равновесия на контакте двух тел (полупроводника и металла или двух полупроводников и т. д.). Обозначим свободную энергию, число электронов и их химический потенциал для каждого из тел соответственно F_1, N_1, μ_1 и F_2, N_2, μ_2 . Свободная энергия всей системы

$$F = F_1 + F_2 = \mu_1 N_1 + \mu_2 N_2.$$

Но отсюда непосредственно следует, что, для того чтобы свободная энергия всей системы была минимальна, химические потенциалы электронов в подсистемах μ_1 и μ_2 должны быть одинаковы. Действительно, если, например, $\mu_1 > \mu_2$, то при переходе электронов из первой системы во вторую F будет уменьшаться и это состояние не будет равновесным.

*) Равенство (4.15), строго говоря, не является прямым следствием (4.14) и справедливо лишь приближенно для конденсированных систем — твердое тело и жидкость, объем которых почти не зависит от давления. Читатель, желающий познакомиться с этим вопросом более подробно, может обратиться к любому курсу физической статистики.

Выражение (4.13) легко обобщается на случай, когда рассматриваемые частицы имеют заряд e и находятся в точке с потенциалом φ . В этом случае

$$dF = -S dT - dW + (\mu - e\varphi) dN, \quad (4.13a)$$

а величина

$$\mu - e\varphi \quad (4.13b)$$

носит название электрохимического потенциала.

В случае заряженных частиц условием равновесия является постоянство электрохимического потенциала (или, иными словами, равенство его во всех подсистемах или частях системы, между которыми может происходить обмен частицами).

4.2. РАСПРЕДЕЛЕНИЕ ФЕРМИ

После предварительных замечаний, сделанных выше, мы можем приступить к отысканию распределения электронов по энергиям. Предположим, что электроны в нашей системе могут находиться на различных энергетических уровнях $(1, 2, \dots, i, \dots, m)$, причем на i -м уровне каждый электрон обладает энергией ε_i и всего может разместиться g_i электронов. т. е. этот уровень g -кратно вырожден.

Полная энергия системы будет

$$\mathcal{E} = \sum_{i=1}^m n_i \varepsilon_i, \quad (4.16)$$

где n_i — число электронов на i -м уровне.

Задача наша и заключается в том, чтобы найти все n_i . Условием равновесия для системы, находящейся в термостате, является минимум свободной энергии:

$$F = \mathcal{E} - TS. \quad (4.17)$$

Здесь

$$S = k \ln \omega = k \sum \ln \omega_i, \quad (4.18) *$$

где ω — вероятность данного распределения электрона по всем состояниям системы, а ω_i — вероятность того,

*) Равенство (4.18) следует из того, что вероятность одновременного осуществления независимых событий равна произведению их вероятностей: $\omega = \prod_i \omega_i$ и $\ln \omega = \sum \ln \omega_i$.

что n_i электронов находятся в i -м состоянии. Но w_i пропорционально числу способов, которым может быть выбрано n_i занятых мест на данном энергетическом уровне из общего числа мест g_i , т. е. числу сочетаний из g_i по n_i :

$$w_i = a_i \frac{g_i!}{n_i! (g_i - n_i)!} . \quad (4.19)$$

Здесь a_i — некоторая константа, не зависящая от n_i , следовательно, согласно (4.17), (4.18)

$$F = \sum_{i=1}^m \left[n_i \varepsilon_i - kT \ln \frac{g_i!}{n_i! (g_i - n_i)!} \right] + A, \quad (4.20)$$

где

$$A = -kT \sum_i \ln a_i.$$

Можно теперь найти распределение частиц по энергиям в состоянии равновесия, т. е. определить все n_i , отыскав минимум свободной энергии как функцию от n_i , при дополнительном условии, что общее число электронов в системе задано:

$$\sum_{i=1}^{i=m} n_i = N. \quad (4.21)$$

Для этого следует избавиться в (4.20) от факториалов с помощью формулы Стирлинга

$$\ln n! \approx n (\ln n - 1)$$

и воспользоваться методом неопределенных коэффициентов, т. е. продифференцировать по n выражения (4.20) и (4.21) и, умножив одно из них, например (4.21), на произвольный коэффициент μ^* , сложить друг с другом. После этого простые преобразования дают

$$\ln \frac{g_i - n_i}{n_i} - \frac{\varepsilon_i}{kT} + \mu^* = 0, \quad (4.22)$$

откуда

$$\frac{n_i}{g_i} = \frac{1}{e^{\frac{\varepsilon_i}{kT} - \mu^*} + 1} . \quad (4.23)$$

Отношение числа электронов n_i с энергией ϵ_i к общему числу состояний с данной энергией g_i называется функцией распределения Ферми и обозначается f_i . Иными словами, f_i — это вероятность того, что состояние с энергией ϵ_i занято электроном. Напротив, вероятность того, что данное состояние свободно, будет

$$f'_i = 1 - f_i.$$

В статистической физике доказывается, что введенная нами постоянная μ^* связана с химическим потенциалом μ соотношением

$$\mu = \mu^* kT \text{ или } \mu^* = \frac{\mu}{kT}. \quad (4.24)$$

В соответствии с этим величина μ^* называется приведенным химическим потенциалом (μ^* есть безразмерный химический потенциал, выраженный в единицах kT). Аналогично величина $x = \epsilon/kT$ называется приведенной энергией электрона. В дальнейшем мы будем индекс i опускать, так как f не зависит от введенной нами произвольной нумерации состояний, а целиком определяется энергией данного состояния ϵ .

Итак, мы доказали, что вероятность того, что состояние с энергией ϵ занято электроном, выражается формулой

$$f = \frac{1}{e^{\frac{\epsilon - \mu}{kT}} + 1}. \quad (4.25)^*$$

4.3. СТАТИСТИКА НЕВЫРОЖДЕННОГО ЭЛЕКТРОННОГО ГАЗА В ПОЛУПРОВОДНИКАХ

В первой главе уже упоминалось, что общее число состояний в свободной зоне твердого тела составляет примерно 10^{22} (на 1 см^3). Число же свободных электронов в полупроводниках колеблется обычно в пределах 10^{12} — 10^{18} см^{-3} . Это значит, что в отличие от металла доля занятых состояний в полупроводнике обычно ничтожно мала или, иными словами, обычно $f \ll 1$ для всех электронных состоя-

*) В тех случаях, когда необходимо подчеркнуть, что f — это равновесная функция распределения, мы будем отмечать ее индексом 0.

ний, включая самые нижние, примыкающие ко дну зоны * (мы в дальнейшем для удобства примем нуль отсчета энергии совпадающим с дном свободной зоны, т. е. для этих электронов будем считать $\varepsilon = 0$). Но если $f \ll 1$, то это значит, что знаменатель (4.25) $1 + e^{\frac{\varepsilon - \mu}{kT}} \gg 1$ или $\frac{\varepsilon - \mu}{kT} \gg 0$, т. е. уровень химического потенциала проходит значительно ниже дна зоны. В этом случае для всех состояний в зоне проводимости мы можем пренебречь единицей в знаменателе выражения (4.25) и, следовательно,

$$f \approx e^{\frac{\mu - \varepsilon}{kT}}. \quad (4.26)$$

Приближенное выражение (4.26) называется распределением Максвелла — Больцмана. Такое состояние электронного газа, в котором $f \ll 1$ и которое описывается приближенной формулой (4.26), называется невырожденным. Мы увидим несколько позже, что в этом случае в отличие от вырожденного состояния средняя энергия электронов растет пропорционально температуре (и это является одним из основных признаков отсутствия вырождения).

Замена точного выражения для функции распределения (4.25) приближением (4.26) значительно упрощает все расчеты, целый ряд вычислений в этом случае можно в общем виде довести до конца.

Подсчитаем сначала число электронов в зоне проводимости полупроводника, энергии которых находятся в интервале $\varepsilon, \varepsilon + d\varepsilon$. Для этого мы должны умножить число состояний в этом интервале $g(\varepsilon) d\varepsilon$ на вероятность их заполнения f :

$$dn = f(\varepsilon) g(\varepsilon) d\varepsilon. \quad (4.27)$$

Мы знаем, что общее число состояний в любой зоне кристалла (объемом 1 см^3) равно числу атомов N_0 , умноженному на кратность g_a атомного уровня, образовавшего данную зону; следовательно,

$$\int_{\varepsilon_1}^{\varepsilon_2} g(\varepsilon) d\varepsilon = g_a N_0, \quad (4.28)$$

*) Бывают, впрочем, случаи, когда условие $f \ll 1$ для самых нижних уровней зоны и в полупроводнике не выполняется, такие полупроводники называются вырожденными (см. ниже).

где \mathcal{E}_1 и \mathcal{E}_2 — энергия нижнего и верхнего края зоны, но конкретный вид самой функции $g(\epsilon)$ нам неизвестен.

Нахождение точного вида этой функции является очень сложной задачей. Однако если учесть, что функция распределения f очень круто спадает по мере удаления (вверх) от дна зоны проводимости, то мы можем удовлетвориться знанием величины $g(\epsilon)$ вблизи дна зоны. Для таких значений энергии интересующая нас функция относительно легко вычисляется, как мы это уже делали для фононов:

$$dN = g(\epsilon) d\epsilon = \frac{8\pi p^2 dp}{h^3}. \quad (4.29)$$

Выражая в (4.29) импульс через энергию $p = \sqrt{2m_n\epsilon}$ и $p dp = m_n d\epsilon$, получаем

$$dN = \frac{4\pi}{h^3} (2m_n)^{3/2} \epsilon^{1/2} d\epsilon.$$

Таким образом, в случае одной простой сферической зоны получаем

$$g(\epsilon) = \frac{4\pi}{h^3} (2m_n)^{3/2} \epsilon^{1/2}. \quad (4.30)$$

Точно так же можно показать, что в случае анизотропной эффективной массы

$$g(\epsilon) = \frac{4\pi}{h^3} 2^{3/2} \sqrt{m_1 m_2 m_3} \epsilon^{1/2}, \quad (4.31)$$

где m_1, m_2, m_3 — главные значения тензора эффективной массы.

Наконец, в случае нескольких (N) эллипсоидов

$$g(\epsilon) = \frac{4\pi}{h^3} 2^{3/2} N \sqrt{m_1 m_2 m_3} \epsilon^{1/2} \quad (4.32)$$

можно все эти случаи объединить в один, если ввести понятие эффективной массы плотности состояний ($m_{\text{пс}}$) и определить эту массу соотношением

$$m_{\text{пс}} = N^{2/3} \sqrt[3]{m_1 m_2 m_3}. \quad (4.33) *$$

*) Все приведенные выше формулы для плотности состояний верны лишь для параболических зон, т. е. таких, в которых энергию можно представить в виде $p^2/2m$, где эффективная масса $m = \text{const}$ не зависит от энергии. За последнее время исследован целый ряд полупроводников (с узкой запрещенной зоной), в которых это условие не выполняется.

Тогда для всех случаев, рассмотренных выше,

$$g(\varepsilon) = \frac{4\pi}{h^3} (2m_{nc})^{3/2} \varepsilon^{1/2}. \quad (4.34)$$

Зная обе функции, $g(\varepsilon)$ и $f(\varepsilon)$, мы можем теперь получить общее число электронов в зоне проводимости

$$n = \int_{\mathcal{E}_1}^{\mathcal{E}_2} f(\varepsilon) g(\varepsilon) d\varepsilon \approx \int_0^{\infty} \frac{4\pi}{h^3} (2m_n)^{3/2} \varepsilon^{1/2} e^{\frac{\mu - \varepsilon}{kT}} d\varepsilon$$

(верхний предел интегрирования мы вправе продлить до бесконечности, так как $f(\varepsilon)$ с ростом ε очень быстро убывает).

Для того чтобы вычислить n , вынесем из-под интеграла все величины, не зависящие от энергии, и превратим подынтегральное выражение в безразмерное, для чего поделим и умножим его на $(kT)^{3/2}$. Тогда выражение для n примет вид

$$n = \frac{4\pi}{h^3} (2m_n kT)^{3/2} e^{\frac{\mu}{kT}} \int_0^{\infty} e^{-x} x^{1/2} dx = \frac{2 (2\pi m_n kT)^{3/2} e^{\frac{\mu}{kT}}}{h^3}, \quad (4.35)^*$$

откуда

$$\mu^* = \frac{\mu}{kT} = \ln \frac{h^3 n}{2 (2\pi m_n kT)^{3/2}}. \quad (4.35a)$$

Аналогичные вычисления дают число дырок p в заполненной зоне

$$p = \frac{2 (2\pi m_p kT)^{3/2} e^{\frac{\mu'}{kT}}}{h^3}, \quad (4.36)$$

где m_p — эффективная масса дырки и μ' — расстояние от уровня химического потенциала до верхнего края заполненной зоны:

$$\mu + \mu' = -\Delta\mathcal{E}_0 \quad \text{или} \quad \mu' = -\Delta\mathcal{E}_0 - \mu. \quad (4.37)$$

*) Так как $\int_0^{\infty} e^{-x} x^{1/2} dx = \frac{1}{2} \sqrt{\pi}$.

Следовательно, мы можем также записать (4.36) в виде

$$p = \frac{2(2\pi m_p kT)^{3/2}}{h^3} e^{-\frac{\Delta \mathcal{E}_0}{kT}} e^{-\frac{\mu}{kT}}. \quad (4.36a)$$

Значительно проще вычисляется число электронов n_d на донорных уровнях; согласно сказанному выше оно равно числу этих уровней, умноженному на функцию распределения:

$$n_d = \frac{N_d}{1 + e^{\frac{-\Delta \mathcal{E}_1 - \mu}{kT}}}, \quad (4.38)$$

где $\Delta \mathcal{E}_1$ — энергия активации для отрыва электрона с этого уровня.

Выведенные выше выражения показывают, что если мы будем знать μ , то сможем определить число электронов в зоне проводимости на локальных уровнях и число дырок в заполненной зоне. Но и сама величина μ вычисляется из того условия, что общее число электронов в системе нам известно. Действительно, электроны в свободной зоне появляются по двум причинам: за счет перебросов с донорных уровней и за счет перебросов из заполненной зоны. Следовательно, число свободных электронов должно равняться сумме числа дырок в заполненной зоне и числа свободных мест на донорных уровнях (числа «связанных дырок») ($N_d - n_d$):

$$n = p + (N_d - n_d) \text{ или } n - p + n_d = N_d. \quad (4.39)$$

Подставляя в (4.39) выражения для n и p из (4.35), (4.37) и (4.38), получаем уравнение, из которого может быть определено μ :

$$\begin{aligned} \frac{2(2\pi m_n kT)^{3/2}}{h^3} e^{\frac{\mu}{kT}} - \frac{2(2\pi m_p kT)^{3/2}}{h^3} e^{-\frac{\Delta \mathcal{E}_0}{kT}} e^{-\frac{\mu}{kT}} + \\ + \frac{N_d}{e^{\frac{-\Delta \mathcal{E}_1 - \mu}{kT}} + 1} = N_d. \end{aligned} \quad (4.40)$$

Однако легко убедиться, что (4.40) представляет собой уравнение третьей степени относительно $e^{-\mu/kT}$. Поэтому обычно его не решают в общем виде, а ограничиваются рассмотрением частных случаев *).

ОБЛАСТЬ ПРИМЕСНОЙ ПРОВОДИМОСТИ

Предположим, что температура недостаточно высока для того, чтобы в заметном количестве наблюдались перебросы электронов из заполненной зоны, т. е. что электроны в зоне проводимости появляются в основном за счет переходов с примесных уровней, а число дырок в заполненной зоне мы можем пренебречь. Тогда уравнение (4.39) заменится более простым:

$$n + n_d = N_d \quad (4.41)$$

или согласно (4.35) и (4.38)

$$\frac{2(2\pi m_n kT)^{3/2}}{h^3} e^{\frac{\mu}{kT}} + \frac{N_d}{e^{-\frac{\Delta \mathcal{E}_1 - \mu}{kT}} + 1} = N_d. \quad (4.42)$$

Уравнение (4.42) — уже квадратное уравнение относительно $e^{\frac{\mu}{kT}}$. Решая его, получаем

$$\mu = -\Delta \mathcal{E}_1 + kT \ln \frac{1}{2} \left[\sqrt{1 + \frac{2e^{\frac{\Delta \mathcal{E}_1}{kT}} N_d h^3}{(2\pi m_n kT)^{3/2}}} - 1 \right]. \quad (4.43)$$

Выражение (4.43) все еще слишком сложно для его непосредственной интерпретации, поэтому рассмотрим его предельные значения в различных интервалах температур.

Область низких температур:

$$e^{\frac{\Delta \mathcal{E}_1}{kT}} \gg \frac{2(2\pi m_n kT)^{3/2}}{N_d h^3}. \quad (4.44)$$

*) Еще более сложный вид принимает выражение для определения химического потенциала при наличии нескольких сортов донорных уровней (с различной энергией активации) или доноров и акцепторов.

Пренебрегая единицами под знаком логарифма, получаем

$$\mu \approx -\frac{\Delta\mathcal{E}_1}{2} + kT \ln \left[\frac{N_d h^3}{2 (2\pi m_n kT)^{3/2}} \right]^{1/2}. \quad (4.45)$$

Выражение (4.45) показывает, что при абсолютном нуле температуры уровень химического потенциала проходит строго посередине между дном запрещенной зоны и донорными уровнями. Второй член в (4.45) при низких температу-

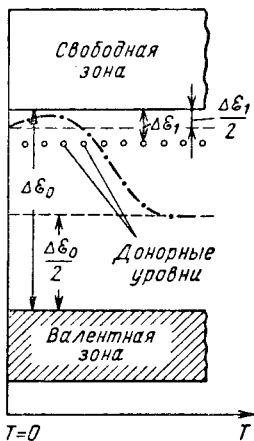


Рис. 4.1. Температурная зависимость уровня химического потенциала в электронном полупроводнике.

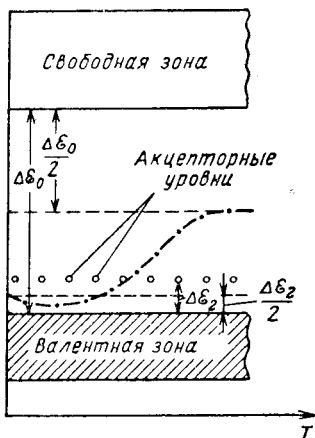


Рис. 4.2. Температурная зависимость уровня химического потенциала в дырочном полупроводнике.

рах [т. е. пока $(2\pi m_n kT)^{3/2} < N_d h^3$] положительный, а при более высоких становится отрицательным. Поэтому уровень химического потенциала сначала поднимается, а затем начинает опускаться (рис. 4.1). Подставляя (4.45) в (4.35), получаем концентрацию свободных электронов:

$$n = (2N_d)^{1/2} \frac{(2\pi m_n kT)^{3/4}}{h^{3/2}} e^{-\frac{\Delta\mathcal{E}_1}{2kT}} = A e^{-\frac{\mathcal{E}_1}{2kT}}, \quad (4.46)$$

где $A = (2N_d)^{1/2} \frac{(2\pi m_n kT)^{3/4}}{h^{3/2}}$ — величина, слабо зависящая от температуры (по сравнению со вторым экспоненциаль-

ным множителем). Выражение (4.46) показывает, что (в этом интервале температур) концентрация электронов в зоне проводимости растет экспоненциально с ростом температуры. Прологарифмировав (4.46), получим

$$\ln n = \ln A - \frac{\Delta \mathcal{E}_1}{2k} \cdot \frac{1}{T}. \quad (4.47)$$

Следовательно, если мы будем на графике откладывать по оси ординат $\ln n$, а по оси абсцисс — обратную температуру ($1/T$), то мы получим прямую с наклоном $\Delta \mathcal{E}_1/2k$.

Область «высоких» температур:

$$e^{\frac{\Delta \mathcal{E}_1}{kT}} \ll \frac{2(2\pi m_n kT)^{3/2}}{h^3 N_d}. \quad (4.48)$$

Термин «высокие температуры» здесь следует принимать лишь в смысле выполнения неравенства (4.48), но мы по-прежнему будем предполагать, что собственная проводимость еще не играет существенной роли. Учитывая (4.48) и разлагая радикал в выражении (4.43) в ряд, получаем

$$\mu \approx kT \ln \frac{N_d h^3}{2(2\pi m_n kT)^{3/2}}. \quad (4.49)$$

Согласно (4.48) и (4.49) в этом интервале температур уровень химического потенциала уже спустился ниже донорных уровней, и, следовательно, почти все они должны быть пусты.

Действительно, подставляя (4.49) в (4.35), получаем

$$n = N_d,$$

т. е. все (или точнее почти все, так как мы все время пользуемся приближенными формулами) электроны перешли в зону проводимости. Концентрация электронов проводимости в этом интервале температур будет оставаться постоянной, что соответствует горизонтальному участку кривой рис. 1.7.

Заметим в заключение, что совершенно аналогичные выводы можно получить для примесного дырочного полупроводника. В этом случае в области низких температур

$$\mu' \approx \frac{\Delta \mathcal{E}_2}{2} - kT \ln \frac{N_d h^3}{2(2\pi m_p kT)^{3/2}} \quad (4.50)$$

где $\Delta\mathcal{E}_2$ — энергия, необходимая для заброса электрона из заполненной зоны на акцепторный уровень, т. е. уровень химического потенциала проходит примерно посередине между верхним краем валентной зоны и акцепторным уровнем (рис. 4.2). При этом концентрация дырок меняется экспоненциально:

$$p = (2N_a)^{1/2} \frac{2 (2\pi m_p kT)^{3/4}}{h^{3/2}} e^{-\frac{\Delta\mathcal{E}_2}{kT}}. \quad (4.51)$$

В области более высоких температур уровень химического потенциала поднимается выше акцепторных уровней, и все они в силу этого оказываются забитыми электронами, а число дырок в заполненной зоне становится постоянным:

$$p = N_a.$$

При дальнейшем подъеме температур и в случае электронного и в случае дырочного полупроводника основную роль начинает играть собственная проводимость.

ОБЛАСТЬ СОБСТВЕННОЙ ПРОВОДИМОСТИ

Рассмотрим теперь второй предельный случай, т. е. предположим, что число перебросов электронов из заполненной зоны в свободную очень велико и по сравнению с этим примесные электроны играют второстепенную роль. Тогда мы можем принять

$$n = p. \quad (4.52)$$

Подставляя (4.35) и (4.36а) в (4.52), получаем

$$\mu = -\frac{\Delta\mathcal{E}_0}{2} + \frac{3}{2} kT \ln \frac{m_p}{m_n}, \quad (4.53)$$

т. е. если $m_n \approx m_p$ уровень химического потенциала проходит приблизительно посередине запрещенной зоны. Подставляя (4.53) в (4.35) и (4.36а), найдем концентрацию электронов и дырок:

$$n = p = \frac{2 (2\pi \sqrt{m_n m_p} kT)^{3/2}}{h^3} e^{-\frac{\Delta\mathcal{E}_0}{2kT}}. \quad (4.54)$$

В полулогарифмическом масштабе

$$\ln n = f\left(\frac{1}{T}\right), \quad (4.55)$$

$$\ln n = \ln A - \frac{\Delta \mathcal{E}_0}{2kT}, \quad (4.56)$$

где A приблизительно постоянно, таким образом, (4.56) изобразится прямой с наклоном $\Delta \mathcal{E}_0/2k$.

4.4. ЭНЕРГИЯ ЭЛЕКТРОНОВ В ЗОНЕ ПРОВОДИМОСТИ, ВЫРОЖДЕНИЕ

Остановимся коротко на термине «вырождение», который в теоретической физике имеет свой особый смысл.

Вырожденными в квантовой механике называются системы, энергия которых не зависит от какого-либо параметра. Если при каких-либо изменениях в условиях существования системы ее энергия начинает зависеть от этого параметра, то мы говорим, что «вырождение снимается».

ВЫРОЖДЕНИЕ В АТОМЕ

В водородном атоме, т. е. при кулоновской зависимости потенциальной энергии от расстояния до ядра ($U = -e^2/r$), энергия электрона не зависит от эксцентриситета орбиты (орбитального квантового числа l), а зависит только от размера ее большой полуоси — главного квантового числа n . Все подоболочки: p , d , f , s , внутри одной оболочке вырождены, т. е. имеют одну и ту же энергию. В более сложных атомах поле других электронов экранирует поле ядра; поэтому по мере удаления от ядра потенциальная энергия убывает быстрее, чем в атоме водорода. В непосредственной близости от ядра $U = -ze^2/r$ постепенно убывает до $U' = -e^2/r$ на больших расстояниях. Поэтому энергия валентных электронов тем меньше (или энергия связи тем больше), чем ближе данная орбита подходит к ядру. Таким образом, в результате взаимодействия электронов друг с другом вырождение снимается, энергия подоболочек в порядке возрастания располагается в следующей последовательности: s , p , d , f .

Второй тип вырождения, имеющегося в любом атоме, так называемое ориентационное. Благодаря шаровой симметрии атомного потенциала энергия электронов на всех

орбитах одной подоболочки одинакова независимо от их ориентации. Этот тип вырождения снимается при наличии магнитного поля (эффект Зеемана) или электрического поля (эффект Штарка) и, в частности, периодического поля в кристалле.

ВЫРОЖДЕНИЕ В КРИСТАЛЛЕ

В системе из большого числа N одинаковых невзаимодействующих (т. е. достаточно удаленных друг от друга) атомов для электронов существует так называемое трансляционное вырождение — в силу трансляционной симметрии энергия электрона не зависит от того, на каком из атомов он (электрон) находится. Таким образом, каждый атомный уровень становится N -кратно вырожденным.

Энергия системы не зависит также от того, какой атом возбужден, ионизован и т. д.; таким образом, все возбужденные состояния кристалла также вырождены. Если расстояние между атомами уменьшается и из изолированных атомов образуется конденсированная система — твердое тело (или жидкость), то в результате взаимодействия атомов трансляционное вырождение снимается. Из N независимых атомов образуется одна гигантская молекула, кристалл, и N -кратно вырожденный уровень объединяется в зону из N уровней, единых для всего кристалла и расположенных очень близко один над другим.

Мы уже упоминали, что нижнему и верхнему уровням в зоне соответствуют состояния электрона, в которых скорость его поступательного движения по кристаллу (или групповая скорость волны $v = dv/dk = h^{-1} \cdot \partial \epsilon / \partial k$) равна нулю, в промежутке скорость проходит через максимальное значение. На корпускулярном языке различные уровни энергии в зоне обусловлены различными скоростями поступательного движения электронов — кинетической энергией, а также их потенциальной энергией. Это видно из того, что каждому значению скорости соответствуют два состояния с различной потенциальной энергией.

На волновом языке (если исходить из приближения сильно связанных электронов, т. е. за исходные состояния брать атомные функции) различные состояния в зоне отличаются различным сдвигом по фазе между волновыми состояниями соседних атомов: для самого нижнего и самого верхнего этот сдвиг по фазе равен 0 и π соответственно, а для остальных пробегает все промежуточные значения.

Состояние электронного газа называется вырожденным, если его энергия не зависит от температуры или, точнее, слабо зависит от температуры — слабее, чем это следует из законов классической физики *).

Рассмотрим вначале качественно температурную зависимость энергии электронного газа в полупроводнике и металле, а затем введем необходимые математические соотношения.

При своем движении в зоне проводимости невырожденного полупроводника, рассмотренного выше, электроны испытывают соударения с атомами, совершающими тепловые колебания, и обмениваются с ними кинетической энергией; поэтому кинетическая энергия этих электронов в среднем, так же как кинетическая энергия атомов, равна $\frac{3}{2} kT$ (при комнатной температуре 0,04 эв).

В силу этого энергия теплового движения электронов, так же как энергия частиц обычного газа, растет пропорционально температуре, скорость — пропорционально квадратному корню из температуры.

Совершенно по-иному ведут себя валентные электроны в металле. Вернемся к примеру щелочных металлов. Как уже упоминалось, в этом случае валентная зона кристалла оказывается наполовину заполненной. Ширина валентной зоны составляет несколько электрон-вольт, и, следовательно, все энергетические уровни в ней до уровня с энергией по крайней мере в один электрон-вольт заполнены уже при абсолютном нуле. Поэтому тепловые колебания атомов не могут оказать значительного влияния на распределение электронов по энергиям.

Действительно, электроны находящиеся не на самых верхних уровнях, т. е. в нижней половине зоны, не могут изменить своей энергии под действием столкновения с атомами, так как при этом они должны перейти на другой уровень, а все соседние уровни заняты. Электроны же, находящиеся на верхних уровнях, имеют энергию не менее 1 эв, в то время как средняя энергия тепловых колебаний атомов составляет при комнатной температуре всего лишь 0,04 эв,

*) Состояние, в котором энергия совершенно не зависит от температуры, называется состоянием полного вырождения.

при температуре 1000°K (700°C)— $0,1 \text{ \AA}$, т. е. по крайней мере в 10 раз меньше. Поэтому при столкновении с атомом электрон не может увеличить свою энергию, с другой стороны, он не может также и передать свою энергию атому, так как при этом он должен был бы перейти в энергетически более низкое состояние, а все нижние энергетические уровни заняты.

Все сказанное выше относится, впрочем, лишь к столкновениям электронов с атомами со средней энергией. Наряду с этим всегда при температуре, отличной от абсолютного нуля, имеется некоторое число атомов, энергия которых значительно больше средней; под действием столкновений с такими атомами электроны могут переходить в верхнюю половину зоны. Так как число таких атомов очень мало, вероятность этих переходов тоже очень мала, но все же благодаря им при температуре, отличной от абсолютного нуля, всегда имеется (и растет с температурой) некоторое число электронов в верхней половине валентной зоны и соответствующее число пустых мест в нижней.

Итак, в силу того, что в щелочных металлах заполнена электронами половина валентной зоны, средняя энергия электронов уже при абсолютном нуле значительно выше, чем энергия тепловых колебаний атомов при любых температурах, вплоть до точки плавления металла. Поэтому энергия электронов не может значительно изменяться под действием соударений с атомами. Все сказанное выше относится и к щелочно-земельным металлам.

Выше мы уже упоминали, что такое состояние свободных электронов, когда их энергия почти не зависит от температуры, называется вырожденным. В полупроводниках энергия электронов растет пропорционально температуре, такое состояние электронного газа называется невырожденным.

Из сказанного выше ясно, что степень вырождения зависит от числа электронов в зоне проводимости. В полупроводниках это число обычно очень мало, и все электроны могли бы разместиться на самых низких уровнях в виде тонкого и плотного слоя «электронной жидкости». Но под действием теплового движения электроны «испаряются», т. е. забрасываются на более высокие уровни, расположенные в энергетическом интервале порядка kT , и так как число уровней в этом интервале во много раз больше, чем число электронов, то «электронная жидкость» превращается в разрежен-

ный газ *). В металлах же зона проводимости заполнена до половины плотной «электронной жидкостью», над которой при температуре, отличной от нуля, имеется тонкий слой «электронного пара». Для того чтобы вся эта «жидкость» превратилась в газ, нужна очень высокая температура (в зависимости от ширины зоны от 5000 до 10 000°).

В 1 см³ металла содержится приблизительно 10²² свободных электронов, в полупроводниках концентрация свободных электронов колеблется обычно в зависимости от чистоты, ширины запрещенной зоны и температуры в пределах от 10¹² до 10¹⁸ см⁻³. Наконец, в полуметаллах и в очень «грязных» полупроводниках мы имеем дело с промежуточным случаем: в них число свободных электронов меняется в зависимости от содержания примесей от 10¹⁸ до 10²¹ см⁻³. Предположим для определенности, что оно равно 10¹⁹ см⁻³. Такое число электронов при плотной упаковке заняло бы все уровни в зоне приблизительно до энергии 0,03 эв. Как мы уже упоминали, при комнатной температуре энергия тепловых колебаний атомов составляет примерно столько же. Поэтому в данном случае при температурах значительно ниже комнатной электронный газ будет находиться в вырожденном состоянии; при температурах значительно выше комнатной — в невырожденном. В интервале температур приблизительно от 100 до 400° К электроны будут находиться в состоянии частичного вырождения, т. е. постепенно переходить из вырожденного состояния в невырожденное. В соответствии с этим энергии электронов при низких температурах будут очень слабо зависеть от температуры, в промежуточной области эта зависимость постепенно усилится и, наконец, при дальнейшем нагревании средняя энергия электронов начнет расти пропорционально температуре.

Все сказанное выше о полуметаллах с концентрацией электронов 10¹⁹ см⁻³ будет относиться ко всем полуметаллам (в том числе и дырочным) с той лишь разницей, что в зависимости от концентрации электронов или дырок переходная область будет сдвигаться соответственно в сторону либо более высоких, либо более низких температур.

Более строгий анализ показывает, что и в полупроводниках иногда наблюдается состояние частичного вырожде-

*) Или, как ее называют, «фермиевская жидкость», так как она подчиняется статистике Ферми — превращается в «максвелловский газ».

ния. Действительно, мы установили в предыдущем параграфе, что при абсолютном нуле температуры уровень химического потенциала проходит посередине между дном зоны проводимости и примесными уровнями, а затем начинает подниматься. При этом если энергия активации примесных уровней мала, а число их велико, то уровень химического потенциала может приблизиться к дну зоны и даже пересечь его так, что неравенство $f \ll 1$ нарушится и все выводы предыдущего параграфа, основанные на нем, будут по крайней мере неточны.

Качественно причина этого заключается в следующем. В примесном полупроводнике с ростом температуры одновременно происходят два процесса: 1) заброс электронов с примесных уровней в зону проводимости, что увеличивает плотность электронного газа (в энергетическом пространстве), т. е. степень вырождения, и 2) расселение под действием все усиливающегося теплового движения этих электронов в более широком интервале энергий, которое приводит к обратному результату.

При низких температурах первый процесс преобладает — плотность электронного газа растет и уровень Ферми поднимается; при высоких температурах примесные уровни начинают истощаться, а скорость расселения электронов в зоне роста, плотность электронов в энергетическом пространстве падает, степень вырождения уменьшается и уровень Ферми начинает понижаться.

Отметим в заключение, что, как следует из самого определения функции распределения

$$f(\epsilon) = \frac{n(\epsilon)}{g(\epsilon)}, \quad (4.57)$$

степень вырождения зависит не только от числа электронов с данной энергией, но и от плотности состояний [см. (4.30) и (4.33)].

В свою очередь $g(\epsilon)$ определяется эффективной массой. Поэтому чем меньше эффективная масса, тем при меньшей концентрации электронов наступает вырождение.

В полупроводниках с резко отличающейся массой электронов и дырок (например, в InSb) вырождение может иметь место даже в области собственной проводимости: если например, $m_n \ll m_p$, то даже небольшое число электронов, перейдя из валентной зоны в свободную, могут заполнить

последнюю настолько, что электроны в ней будут в вырожденном состоянии.

Перейдем теперь к количественному рассмотрению вопроса.

*Концентрация носителей тока **

В общем случае согласно (4.25) и (4.33)

$$n = \int_0^{\infty} f(\varepsilon) dG(\varepsilon) \approx \frac{4\pi}{h^3} (2m_{\text{пс}})^{3/2} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\frac{\varepsilon-\mu}{kT}} + 1}. \quad (4.58)$$

Поделив и умножив правую часть на $(kT)^{3/2}$, можно привести интеграл к безразмерному виду

$$n = \frac{4\pi}{h^3} (2m_{\text{пс}}kT)^{3/2} F_{1/2}(\mu^*), \quad (4.59)$$

где $F_{1/2}(\mu^*)$ — так называемый интеграл Ферми:

$$F_{1/2}(\mu^*) = \int_0^{\infty} \frac{x^{1/2} dx}{e^{x-\mu^*} + 1}, \quad (4.60)$$

x — приведенная (безразмерная) энергия $x = \varepsilon/kT$ и μ^* — приведенный химический потенциал $\mu^* = \mu/kT$. Интеграл Ферми $F_i(\mu^*)$ для целочисленных и полуцелочисленных значений i протабулированы (см. приложения).

В случае, если электронный газ не вырожден, $e^{x-\mu} \gg 1$, и единицей в (4.60) можно пренебречь, выражение (4.60) легко интегрируется (см. выше).

В случае полного вырождения ($T = 0$) $f = 1$ при $x < \mu^*$ и $f = 0$ при $x > \mu^*$. В соответствии с этим

$$n \approx 2 \frac{4\pi}{3} p_{\text{макс}}^3 / h^3 = \frac{8\pi}{3h^3} (2m_{\text{пс}}\mu_0)^{3/2}, \quad (4.61)$$

где μ_0 — значение уровня химического потенциала при абсолютном нуле; $p_{\text{макс}} = \sqrt{2m\mu_0}$ — граничное значение квазиимпульса.

Физический смысл (4.61) очевиден — все электроны при абсолютном нуле размещаются внутри сферы

*) Все рассуждения в этом параграфе мы проводим на примере электронов, однако все здесь сказанное в равной мере относится и к дыркам.

радиуса $\rho_{\text{макс}}$; согласно (4.61)

$$\mu_0 = \frac{\hbar^2}{2m_{\text{пс}}} \left(\frac{3n}{8\pi} \right)^{2/3}. \quad (4.62)$$

В случае сильного (но не полного) вырождения (металл при температуре, отличной от абсолютного нуля) выражение (4.60) может быть приближенно вычислено интегрированием по частям [11]:

$$n \approx \frac{8\pi}{3\hbar^3} (2m_{\text{пс}})^{3/2} \left[\mu^{3/2} + \frac{\pi^2}{8} \frac{(kT)^2}{\mu^{1/2}} \right] \quad (4.63)$$

и

$$\mu = \mu_0 \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{\mu_0} \right)^2 \right]. \quad (4.64)$$

Формула (4.35) дает удовлетворительную точность ($\approx 5\%$) при $\mu^* < -2$, формулы (4.63) и (4.64) — при $\mu^* > 10$; при промежуточных значениях $-2 < \mu^* < 16$ следует пользоваться (4.59) и протабулированными значениями $F_{1/2}(\mu)$.

Согласно (4.35) при отсутствии вырождения

$$\mu^* \approx \ln \frac{\hbar^3 n}{2 (2\pi m k T)^{3/2}}. \quad (4.65)$$

Иногда (главным образом в старых книгах) величину $\eta = \hbar^3 n / 2 (2\pi m k T)^{3/2}$ называют критерием вырождения, имея в виду при этом, что при $n < 1$ и $\mu^* < 0$ электронный газ не вырожден, а при $\eta > 1$ и $\mu^* > 0$ — частично вырожден.

Этим критерием можно пользоваться лишь для весьма грубых оценок; в частности, следует заметить, что сама формула (4.65) при $\mu^* = 0$ дает ошибку в определении $\mu^* \approx 30\%$.

Полная энергия электронов в зоне проводимости

$$\mathcal{E} = \int \epsilon f(\epsilon) g(\epsilon) d\epsilon = \frac{4\pi}{\hbar^3} (2m_{\text{пс}})^{3/2} kT F_{3/2}(\mu^*); \quad (4.66)$$

средняя

$$\bar{\epsilon} = \frac{\mathcal{E}}{n} = kT \frac{F_{3/2}(\mu^*)}{F_{1/2}(\mu^*)}. \quad (4.67)$$

В случае невырожденного электронного газа

$$\bar{\varepsilon} = kT \frac{\int_0^{\infty} x^{3/2} e^{-x} dx}{\int_0^{\infty} x^{1/2} e^{-x} dx} = \frac{3}{2} kT. \quad (4.68)$$

Таким образом, средняя кинетическая энергия электронов в невырожденном полупроводнике выражается точно такой же формулой, как кинетическая энергия одноатомных молекул газа.

В случае полного вырождения согласно (4.67) $\bar{\varepsilon} = \frac{3}{5} \mu_0$ и в случае сильного, но не полного вырождения ($\mu^* > 10$)

$$\bar{\varepsilon} = \frac{3}{5} \mu_0 \left[1 + \frac{5}{12} \pi^2 \left(\frac{kT}{\mu_0} \right)^2 \right]. \quad (4.69)$$

В промежуточных случаях $-2 < \mu^* < 10$ следует пользоваться формулой (4.67) и таблицами интегралов Ферми (см. приложение).

НЕКОТОРЫЕ ВОПРОСЫ ТЕОРИИ ЯВЛЕНИЙ ПЕРЕНОСА

До сих пор мы изучали поведение полупроводников в условиях равновесия: в третьей главе был рассмотрен энергетический спектр электронов, в четвертой главе — их статистика, т. е. расселение по энергетическим уровням, во второй — соответствующие вопросы по отношению к фотонам.

Все последующие главы связаны с поведением электронов в неравновесных условиях. Важную часть этого круга вопросов составляют так называемые явления переноса, т. е. ряд явлений, обусловленных электрическими и тепловыми потоками, возникающими в среде при наличии электрических, тепловых и магнитных полей. Сюда относятся электро- и теплопроводность, термоэлектрические, гальвано- и термомагнитные явления. При этом предполагается, что концентрация носителей остается равновесной или почти равновесной *) (хотя она и может меняться от точки к точке при наличии градиента температуры) таким образом, что ее изменения не откладывают заметный отпечаток на картину явлений. Поэтому, рассматривая явления переноса, мы исключаем контактные явления, в которых нарушение концентрационного равновесия играет существенную роль. Обычно явления переноса (или по крайней мере их теория) изучаются в стационарных условиях. Полученные результаты потом без труда обобщаются на случай переменных (но не слишком быстро меняющихся) полей.

*) Исключением являются явления в сильных электрических полях (см. ниже).

Изучение явлений переноса (или, как их также называют, кинетических эффектов) мы начнем с самого упрощенного, качественного подхода, а затем введем некоторые уточнения.

5.1. ЭЛЕМЕНТАРНЫЙ РАСЧЕТ ЭЛЕКТРОПРОВОДНОСТИ И ПОДВИЖНОСТИ

Как уже упоминалось в предыдущей главе, в состоянии равновесия все свободные электроны участвуют в беспорядочном тепловом движении. При этом, так как функция распределения симметрична, то количество электронов, движущихся в противоположных направлениях, всегда одинаково, поэтому макроскопические электрические токи отсутствуют. Более того, это равновесие детально; это значит, что не только общее число электронов, движущихся «туда и обратно», одинаково, но это равенство соблюдается для любой данной энергии и скорости, поэтому отсутствуют не только электрические, но и тепловые потоки.

Таким образом, в условиях равновесия $\bar{v}_x = \bar{v}_y = \bar{v}_z = 0$, $j_x = j_y = j_z = 0$, $q_x = q_y = q_z = 0$, где j — плотность электрического тока и q — плотность потока энергии.

Представим себе теперь, что к рассматриваемому нами полупроводниковому образцу приложено электрическое поле E_x . Тогда каждый свободный электрон под влиянием силы $F_x = eE_x$, действующей на него, получает ускорение $a_x = eE_x/m$ и приращение скорости

$$\Delta v_x = \frac{eE_x}{m} t. \quad (5.1)$$

Если бы при этом никакой процесс не препятствовал увеличению скорости электрона, то она возрастала бы беспредельно. Но в действительности электроны время от времени испытывают столкновения (друг с другом, с тепловыми колебаниями решетки и дефектами), при которых теряют приобретенную в поле энергию и направленную скорость. Поэтому мы предположим, как это уже делали раньше, что можно ввести некоторое среднее время релаксации $\bar{\tau}$ с учетом всех процессов столкновений и считать, что в течение этого времени электрон беспрепятственно ускоряется, а в момент столкновения полностью теряет

направленную составляющую скорости, после чего весь процесс повторяется сначала.

Если число таких столкновений (или вероятность столкновения) за одну секунду ω , то $\bar{\tau} = 1/\omega$.

Предположим также, что направленная добавка к скорости $\bar{\Delta v} = \bar{\tau} a_x$, приобретаемая за время пробега электрона *), много меньше v_{0x} . Средняя скорость движения электрона в направлении поля будет равна

$$\bar{v}_x = \frac{\bar{v}_{0x} + \bar{v}_{\tau x}}{2} = \frac{\bar{\Delta v}}{2} = \frac{e\bar{\tau}E_x}{2m} \quad (5.2)$$

(так как $\bar{v}_{0x} = 0$).

При более строгом расчете (с учетом распределения электронов по временам свободного пробега)

$$\bar{v}_x = \frac{e\bar{\tau}}{m} E_x = u E_x, \quad (5.3)$$

где

$$u = \frac{e}{m} \bar{\tau} = \frac{e}{m} \bar{l} \quad (5.3a)$$

— подвижность, т. е. дрейфовая скорость электронов, приобретаемая ими в поле, равном единице.

Если число свободных электронов в одном кубическом сантиметре n , все они движутся со скоростью \bar{v}_x и каждый переносит заряд e , то плотность электрического тока

$$j_x = en\bar{v}_x = \frac{e^2 n \bar{\tau}}{m} E_x. \quad (5.3b)$$

Таким образом, электропроводность

$$\sigma = \frac{e^2 n \bar{\tau}}{m} = enu. \quad (5.4)$$

В случае двух знаков носителей

$$\sigma = e(nu_n + pu_p). \quad (5.5)$$

В случае нескольких (k) сортов носителей (например, тяжелых и легких дырок и электронов)

$$\sigma = e^2 \sum_k \frac{n_k \bar{\tau}_k}{m_k}, \quad (5.6)$$

где суммирование проводится по всем k зонам.

*) При этом ограничении мы можем считать, что $\bar{\tau}$ не зависит от поля E_x .

Мы уже упоминали в гл. 3, что если электроны, или, в более общем виде, носители k -го сорта испытывают столкновения различного рода, например столкновения с тепловыми колебаниями, с ионами примеси и нейтральными дефектами, то для каждого типа дефектов можно ввести свое время свободного пробега — соответственно τ_T , $\tau_{и}$, $\tau_{н}$ и т. д.; обозначим в общем виде время свободного пробега носителей k -го сорта, соответствующее определенному (i -му) сорту соударений, через τ_{ik} ; результирующее время свободного пробега τ_k будет, разумеется, меньше любого τ_i . Чтобы получить τ_k , мы должны просуммировать все столкновения за 1 сек:

$$\bar{\omega}_k = \overline{\sum_i \omega_{ik}} = \sum_i \frac{1}{\tau_{ik}},$$

и, следовательно,

$$\frac{1}{\tau_k} = \sum_i \frac{1}{\tau_{ik}}. \quad (5.7)$$

Соотношение (5.7) также не является строгим и точным; в дальнейшем мы должны будем учесть, что не каждое столкновение приводит к полной потере направленной составляющей скорости и что время свободного пробега электрона для каждого типа столкновений по-своему зависит от его энергии, и внести в формулу (5.7) соответствующие коррективы.

В приведенных выше формулах время релаксации τ может быть выражено через среднюю длину свободного пробега и среднюю тепловую скорость:

$$\tau = \left(\frac{l}{v} \right) \approx \left(\frac{\bar{l}}{\bar{v}} \right),$$

соответственно подвижность

$$u \approx \frac{e}{m} \frac{\bar{l}}{\bar{v}} \quad (5.8)$$

и электропроводность

$$\sigma \approx \frac{e^2}{m} \frac{\bar{l}}{\bar{v}} n. \quad (5.9) *$$

*) Знак «приблизленно» в последних двух формулах поставлен, так как $(\bar{l}/\bar{v}) \approx \bar{l}/\bar{v}$.

Таким образом, на первый взгляд может показаться, что проблема решена: получены формулы для подвижности и электропроводности — остается только пользоваться этими формулами. Однако, не преуменьшая значения формул (5.3), (5.4) и (5.9), связывающих электропроводность и подвижность, а также выражающих макроскопические параметры вещества через микроскопические параметры носителей, следует подчеркнуть, что целый ряд вопросов остался нерешенным.

Некоторые из этих вопросов попытаемся решить или, по крайней мере, отметить путь решения, по поводу других укажем лишь соответствующую литературу.

УТОЧНЕНИЕ ПОНЯТИЯ О СТОЛКНОВЕНИИ

Во-первых, мы предположили, что движение электрона можно рассматривать в виде чередующихся периодов сво-

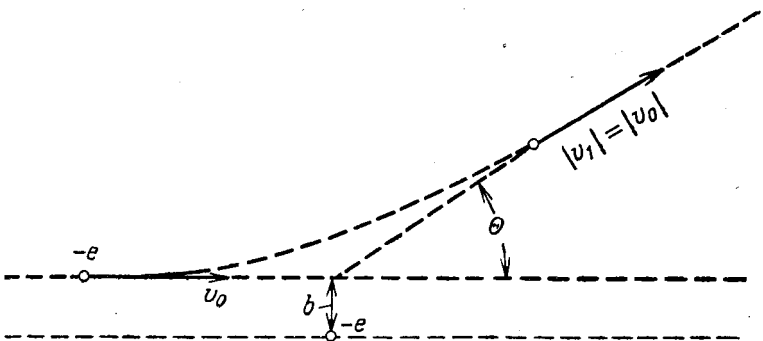


Рис. 5.1. Траектория движения электрона вблизи примесного иона.

бодного полета и мгновенных соударений, что далеко не всегда можно сделать. Так, даже в таком простейшем случае, как рассеяние на ионах примеси, взаимодействие никак нельзя считать мгновенным. В этом случае электрон постепенно отклоняется от направления своего первоначального движения все время, пока находится в поле иона; при этом, если прицельное расстояние b (рис. 5.1) будет сравнимо со средним расстоянием между ионами (а это как раз наиболее типичный случай), то отклонения

на любых участках траектории будут сравнимы, и понятия столкновения и времени свободного пробега должны быть существенно пересмотрены.

Существует ряд случаев, когда теория требует более радикального пересмотра. Наибольшие трудности в этом отношении представляют вещества с малой подвижностью; в частности, случай, когда длина свободного пробега электрона меньше его длины волны,

$$l < \lambda, \quad (5.10)$$

значительно более сложен.

Уже само соотношение (5.10) показывает, что ситуация требует тщательного анализа — трудно представить волну, которая рассеивается (или затухает), не совершив ни одного колебания; но неравенству (5.10) можно придать еще более наглядный вид, воспользовавшись соотношением де Бройля $\lambda = \frac{h}{mv}$ и заменив длину свободного пробега через время $l = v\tau$:

$$v\tau < \frac{h}{mv} \quad \text{или} \quad mv^2 < \frac{h}{\tau}. \quad (5.11)$$

Однако согласно принципу неопределенности h/τ есть неопределенность энергии электрона. Таким образом, согласно (5.11) неопределенность энергии электрона становится больше его кинетической энергии. Это значит, что взаимодействие электрона с дефектами или колебаниями решетки (фононами) так велико, что электрон уже нельзя рассматривать как самостоятельную свободную частицу (даже с поправкой на эффективную массу), а само взаимодействие — как последовательность независимых столкновений (переходов). Если причиной неравенств (5.10) и (5.11) является взаимодействие электрона с примесными ионами, то мы должны учесть их влияние на энергетический спектр электрона *).

*) Мы уже упоминали, что за последнее время рядом авторов разработана теория ионных полупроводников с малой подвижностью, в которой учитывается сильное взаимодействие электронов с фононами, приводящее к образованию так называемых поляронных состояний с характерной для этого случая экспоненциальной зависимостью подвижности от температуры (в определенном интервале температур) [9].

Выше, при вычислении средней скорости электрона, мы сразу же заменили время свободного пробега его средним значением, никак это не обосновав. В действительности же время свободного пробега может зависеть от целого ряда факторов (эффективной массы и энергии носителей, температуры, характера и числа рассеивающих центров), и выявление этих зависимостей — одна из главных задач теории явлений переноса.

Мы вернемся к этому в следующем параграфе, а теперь начнем с учета статического разброса времени свободного пробега. При этом мы сделаем простейшее и наиболее естественное допущение: предположим, что вероятность того, что электрон столкнется в течение промежутка времени dt , пропорциональна dt/τ , где $1/\tau$ — пока неизвестный параметр.

Обозначим вероятность того, что электрон летел в течение времени t без столкновения $p(t)$. Тогда вероятность того, что электрон пролетел время t без столкновений, а затем испытал столкновение за время dt , будет равна произведению вероятностей этих двух событий:

$$dw(t) = p(t) \frac{dt}{\tau} .$$

Но это $dw(t)$ и есть уменьшение $p(t)$ за время dt : $dw = -dp$; следовательно,

$$dp = -p \frac{dt}{\tau} \text{ и } p = e^{-\frac{t}{\tau}} \quad (5.12)$$

(если начать отсчет с момента после столкновения, когда $p = 1$).

Теперь мы подготовлены к тому, чтобы вычислить все величины:

$$\bar{t} = \int_0^{\infty} t dw(t) = \tau, \quad (5.13)$$

$$\begin{aligned} \bar{v}_x &= \int_0^{\infty} v_{xt} dw(t) = \int \left[\bar{v}_{0x} + \frac{eEt}{m} \right] e^{-\frac{t}{\tau}} \frac{dt}{\tau} = \\ &= \bar{v}_{0x} + \frac{eE\tau}{m} = \frac{eE\tau}{m} . \end{aligned} \quad (5.14)$$

Таким образом, действительно, двойка в знаменателе выражения для подвижности (5.2) была лишней.

Мы видим из (5.13) и (5.14), что введенный нами параметр τ приобретает двойкий смысл: это и среднее время между столкновениями, и время релаксации системы; мы к этому вопросу вернемся еще раз позднее, а теперь введем еще некоторые уточнения в понятие соударения.

УПРУГИЕ И НЕУПРУГИЕ СТОЛКНОВЕНИЯ

При своем движении в направлении электрического поля электрон приобретает не только дополнительную скорость

$$d\mathbf{v} = \frac{e\mathbf{E}}{m} dt,$$

но и энергию

$$d\varepsilon = (\mathbf{F} d\mathbf{S});$$

поэтому для того чтобы поддерживалось стационарное состояние, необходимо существование таких соударений, при которых электрон мог бы терять не только направленный импульс, но и дополнительную энергию. Столкновения первого типа, т. е. без значительной потери энергии, называются упругими, второго — неупругими.

Покажем, что столкновения электронов с примесными атомами (без возбуждения последних) и с фононами являются почти упругими.

В тепловом равновесии средние кинетические энергии электронов и атомов должны быть одинаковы и, следовательно, средние скорости атомов \mathcal{V} в $\sqrt{m/M}$ раз меньше (M — масса атома), чем скорость электронов v_0 *), поэтому пренебрежем кинетической энергией атома до соударения (т. е. положим $\mathcal{V}_0 \approx 0$). Тогда законы сохранения энергии и импульса примут вид

$$m|\mathbf{v}_0 - \mathbf{v}_1| = M\mathcal{V}_1, \quad (5.15)$$

$$\frac{mv_0^2}{2} - \frac{mv_1^2}{2} = \frac{M\mathcal{V}_1^2}{2}. \quad (5.16)$$

*) Для вырожденного электронного газа это отношение будет еще больше.

Из (5.16) находим

$$\Delta |v| \approx \sqrt{\Delta v^2} = \sqrt{\frac{M}{m}} \mathcal{V}_1 \quad (5.17)$$

и, подставляя \mathcal{V}_1 из (5.17) в (5.15), получаем

$$\frac{|\Delta v|}{\Delta v} \approx \sqrt{\frac{M}{m}} \gg 1. \quad (5.18)$$

Соотношение (5.18) показывает, что векторное изменение скорости электрона во много раз больше, чем такое ее изменение по абсолютной величине — это и значит, что столкновение почти упругое.

Этот же результат можно получить в явном виде, получив v_1 из (5.15), подставив его в (5.16) и обозначив угол между v_0 и \mathcal{V}_1 через θ :

$$2v_0 \cos \theta = -\mathcal{V}_1 \left(\frac{M}{m} \pm 1 \right). \quad (5.19)$$

Очевидно, что в (5.19) можно пренебречь единицей по сравнению с M/m ; но если читатель возьмет на себя труд проделать промежуточные выкладки, то он убедится, что эта единица «происходит» от кинетической энергии атома в (5.16), иначе говоря, мы тем самым предполагаем соударение упругим. Полагая в (5.19) $\cos \theta = 1$, получаем

$$\mathcal{V}_{\text{макс}} = \frac{2vm}{M} \quad (5.20)$$

или, в общем случае, так как $\cos \theta \leq 1$, то

$$M\mathcal{V}_1 \leq 2mv_0. \quad (5.21)$$

Неравенство (5.21) имеет очень простой физический смысл. Из закона сохранения энергии следует, что $v_1 < v_0$ *), следовательно, максимальное изменение импульса электрона при упругом лобовом ударе $\Delta p \leq (mv_0) - (mv_1) = 2mv_0$. Это и есть максимальный импульс, который может получать атом ($M\mathcal{V}_1$). Из (5.21) находим максимальную энергию отдачи атома:

$$T_{\text{макс}} = \frac{1}{2} M (\mathcal{V}_{\text{макс}})^2 = 4 \frac{m}{M} \left(\frac{mv_0^2}{2} \right). \quad (5.22)$$

*) Мы предположили, что скорость атома до столкновения $v_0 = 0$, следовательно, электрон передает часть энергии атому, и его скорость после столкновения $v_1 < v_0$.

Выражение (5.22) еще раз показывает, что атом, отобрав у электрона весь импульс, получает лишь ничтожную часть его энергии.

Полученные соотношения характерны для столкновений электрона (или любой другой частицы) с более тяжелой частицей («стенкой») — он охотно отдает импульс, но не может отдать ей (без возбуждения внутренних степеней свободы тяжелой частицы) своей кинетической энергии.

Обратные соотношения имеют место при столкновениях зонных электронов с фотонами видимой и инфракрасной частей спектра. Импульс этих квазичастиц $p_{\phi} = h\nu/c$ очень мал [так как скорость света много больше (примерно в 10^8 раз) скорости звука, а энергия сравнима с энергией электронов], поэтому электрон при излучении или поглощении такого фотона может легко изменить свою энергию, но всегда появляются трудности со «сбытом» импульса. Отсюда возникают правила отбора или необходимость в участии третьей частицы — примесного атома, фонона или др.

Рентгеновские и еще более жесткие фотоны являются уже более «тяжелыми» частицами, так как их энергия и соответственно импульс во много раз больше, поэтому они могут взять на себя и импульс электрона.

К вопросу о взаимодействии электрона с излучением мы вернемся позднее (см. гл. 9), а теперь рассмотрим столкновение электронов с акустическими фононами. Для этого опять напишем законы сохранения энергии и импульса электрона, но на этот раз выразим энергию электрона через его квазиимпульс:

$$\frac{h^2 k_1^2}{2m} = \frac{h^2 k_0^2}{2m} \pm h\nu, \quad (5.23)$$

$$h\mathbf{k}_1 = h\mathbf{k}_0 \pm h\mathbf{q}. \quad (5.24)$$

Здесь ν и $h\mathbf{q}$ — соответственно частота и квазиимпульс фонона, связанные для акустической ветви:

$$\nu = \omega_{\phi} q, \quad (5.25)$$

где ω_{ϕ} — фазовая скорость фонона (скорость звука).

Определив \mathbf{k}_1 из (5.24), подставив его в (5.23), после простых преобразований получим соотношение, совершенно

аналогичное (5.19):

$$q = \mp 2k_0 \cos \Theta \pm \frac{2m\omega_{\Phi}}{h}, \quad (5.26)$$

где Θ — угол между k_0 и q .

Второй член в правой части (5.26) происходит от энергии фонона ($h\nu$) в (5.23); покажем, что при обычных условиях он мал и что, следовательно, столкновения электронов с фононами также являются упругими. Для этого оценим отношение этого члена к q :

$$\frac{m\omega_{\Phi}}{hq} = \frac{\omega_{\Phi}}{v_0} \frac{mv_0}{hq}. \quad (5.27)$$

(Правая часть выражения (5.27) отличается от левой тем, что мы умножили и разделили ее на скорость электрона и таким образом получили отношение импульсов.) Для того чтобы рассеяние было эффективным, необходимо, чтобы импульс фонона был одного порядка с импульсом электрона *). Следовательно,

$$\frac{mv_0}{hq} \approx 1 \text{ и } \frac{mv_0}{hq} \approx \frac{\omega_0}{v_0} \ll 1.$$

Так как фазовая скорость упругих волн в обычных полупроводниковых кристаллах $\omega_0 \approx 10^5$ см/сек, а средняя скорость теплового движения электронов при комнатной температуре $v_0 \approx 10^7$ см/сек, следовательно, при обычных условиях второй член в (5.26) очень мал по сравнению с q и им можно пренебречь, т. е. столкновение с фононом также является почти упругим и согласно (5.26)

$$q \approx \mp 2k_0 \cos \Theta. \quad (5.28)$$

С другой стороны, из (5.26) следует, что при достаточно низких температурах (т. е. таких, при которых $v_0 \approx \omega_0$) столкновения электронов с фононами перестают быть упругими — фонон перестает быть «тяжелой частицей» по сравнению с электроном **).

*) Из (5.24) следует, что если $q \ll k_0$, то $k_1 = k_0$, т. е. электрон просто не «почувствует» столкновения с фононом.

**) Последний результат качественно можно было предсказать с самого начала. Действительно, как мы уже упоминали, для того чтобы соударение электрона с фононом было эффективным, их импульсы должны быть одного порядка; при этих условиях для того чтобы их энергии были сравнимы, необходимо (и достаточно) равенство их скоростей. Оценка показывает, что это имеет место при температурах порядка 1° К.

Из соотношения (5.28) следует очень важный вывод: так как $\cos \theta \leq 1$, то

$$q_{\text{макс}} \leq 2k. \quad (5.29) *$$

Это значит, что электроны могут взаимодействовать только со сравнительно длинноволновыми фононами. Это иллюстрируется рис. 5.2, на котором приведены кривые дисперсии электронов и фононов и показано, с какими акустическими и оптическими фононами могут взаимодействовать тепловые электроны.

Напротив, оптические фононы, удовлетворяющие этому неравенству (см. тот же рисунок), обладают сравнительно большой энергией, порядка $k\theta_0$ (θ_0 — температура Дебая для оптических колебаний); поэтому столкновения с оптическими фононами можно считать приблизительно упругими лишь при таких температурах, при которых энергия электронов $\epsilon \gg k\theta_0$ (или при наличии сильного вырождения). При низких температурах столкновения с оптическими фононами также не могут играть существенной роли, так как число их $N_0 e^{-\theta/T}$ очень мало. Поэтому существует промежуточная область температур, в которой неупругие столкновения с оптическими фононами могут играть существенную роль и вызвать ряд аномалий:

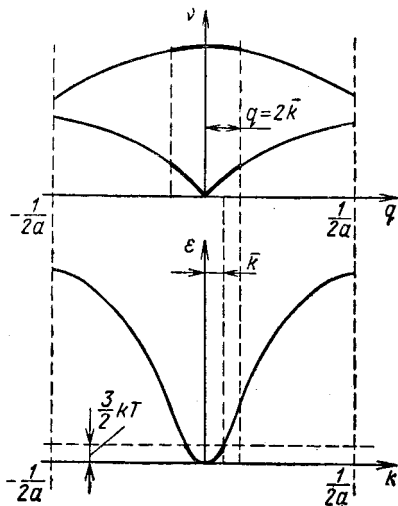


Рис. 5.2. Кривые дисперсии электронов $\epsilon(k)$ и фононов $\nu(q)$.

Жирными линиями обведены части спектра, соответствующие тепловым электронам ($\epsilon \approx \frac{3}{2} kT$) и фононам, которые удовлетворяют неравенству $q < 2k$ (т. е. могут взаимодействовать с этими электронами).

*) Соотношение (5.29) также легко наглядно интерпретировать: если столкновение упругое, то $|k_1| = |k_0|$ и при «лобовом» столкновении с фононом $k_1 = -k_0$ и $|q| = |2k_0|$. Во всех остальных случаях из условий и построения треугольника импульсов следует, что $q \leq 2k$.

а) резкое падение подвижности с ростом температуры;
б) отступления от закона Видемана и Франца (это обусловлено тем, что при наличии неупругих столкновений нельзя ввести одно и то же время релаксации для явлений в электрических и тепловых полях).

Подвижность электронов ограничивают в основном столкновения, лишь связанные с рассеянием на большие углы (приблизительно 180°). Для теплопроводности (и для всех процессов, связанных с переносом тепла) существенны также так называемые «вертикальные переходы», т. е. рассеяние с небольшой потерей импульса, но с потерей энергии *); поэтому время релаксации для теплопроводности становится меньше, чем для подвижности и соответственно уменьшается отношение этих величин ($\kappa_{эл}/\sigma$).

Вторым видом неупругих столкновений являются межэлектронные. Эти столкновения также совершенно по-разному влияют на электрические и тепловые явления.

При наличии только электрического и магнитного полей межэлектронные столкновения приводят лишь к перераспределению импульса между электронами, но так как общий импульс при этом сохраняется, то сами по себе они не могут ограничить подвижность (и электропроводность). Если же имеется дополнительно другой механизм рассеяния носителей с большой дисперсией (т. е. зависимостью длины свободного пробега от энергии), то эти столкновения могут косвенно повлиять на подвижность и уменьшить ее, перераспределив энергию между электронами. Поэтому роль межэлектронных столкновений проявляется при «промежуточных» концентрациях электронов приблизительно 10^{19} и средних температурах, т. е. когда концентрация электронов достаточно велика, чтобы вероятность межэлектронных столкновений была значительна в то время, когда нет вырождения (при наличии вырождения столкновения уже не могут изменить распределения по энергиям).

Значительно сильнее межэлектронные столкновения влияют на тепловые процессы и термомагнитные явления. Здесь мы имеем дело со встречными потоками электронов (от горячего конца образца к холодному и обратно),

*) Формально это проявляется в том, что в данном случае в интегралах переноса появляется множитель $\varepsilon - \mu$, который меняет знак при пересечении поверхности Ферми.

и поэтому эти столкновения приводят непосредственно к потере импульса, ограничивают теплопроводность и существенно изменяют термомагнитные коэффициенты.

Межэлектронные столкновения, так же, как и столкновения электронов с оптическими фононами, приводят к уменьшению числа Лоренца *).

Таким образом, мы рассмотрели два типа почти упругих столкновений — с дефектами и акустическими фононами, и два неупругих — с оптическими фононами и межэлектронные, и убедились, что неупругие столкновения играют существенную роль лишь в ограниченном интервале температур и концентраций носителей. Поэтому (а также потому, что рассмотрение неупругих процессов весьма сложно) мы их в дальнейшем учитывать не будем и будем считать, таким образом, что наиболее важные для рассеяния электронов столкновения (с тепловыми колебаниями и примесными атомами) являются почти упругими. Из проведенных выше выкладок следует, что при расчете времени, длины свободного пробега и подвижности можно пренебречь изменением абсолютной величины скорости и квазиимпульса электрона. Но из этого отнюдь не следует, что такое пренебрежение правомерно при анализе всех процессов. Как уже упоминалось выше, при движении в электрическом поле электроны непрерывно увеличивают свою кинетическую энергию, и для поддержания теплового равновесия необходимо, чтобы существовал механизм «сброса» этой избыточной энергии.

Анализ, проведенный выше, показывает, что обмен энергией между электронами, с одной стороны, и акустическими тепловыми колебаниями, с другой, весьма затруднен, но все же он происходит и именно он обеспечивает сохранение теплового равновесия между электронным газом и кристаллом в целом при отсутствии более эффективных (неупругих) столкновений.

Впрочем, при достаточно сильных полях электрон настолько быстро набирает энергию, что не успевает отдавать ее решетке, и с этим связан целый ряд весьма интересных явлений (зависимость подвижности и концентрации носителей от поля и др.), которые мы рассмотрим в конце этой

*) Различное влияние разного вида соударений на различные процессы приводит к тому, что показатель степени ν в зависимости длины свободного пробега электрона от его энергии может быть неодинаковым в различных кинетических коэффициентах.

главы. Пока же отметим, что в обычных условиях электрон при столкновениях с акустическими фонами приблизительно в сто раз медленнее теряет свою избыточную энергию, чем свой направленный импульс. При этом не каждое столкновение приводит сразу же к потере направленного импульса; анализу этого важного уже сейчас для нас вопроса посвящен следующий раздел.

БОЛЕЕ ТОЧНЫЙ ПОДСЧЕТ ПОТЕРИ НАПРАВЛЕННОЙ СОСТАВЛЯЮЩЕЙ СКОРОСТИ

В начале этой главы мы предположили, что каждое столкновение приводит к полной потере направленной (вдоль поля) составляющей скорости, т. е. что направление скорости электрона после столкновения совершенно не зависит от направления скорости до столкновения. В действительности это не всегда так. Например, при

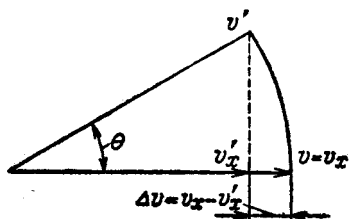


Рис. 5.3. Изменение направленной составляющей скорости электрона при упругом рассеянии.

рассеянии электрона на примесном ионе (см. рис. 5.1) угол отклонения электрона θ зависит от прицельного расстояния b и от скорости электрона (чем медленней движется электрон, тем больше времени он находится вблизи иона и тем сильнее отклоняется), от массы электрона и диэлектрической постоянной кристалла.

Из проведенного выше рассмотрения взаимодействия электронов с фонами следует, что угол отклонения электрона также зависит от импульса (волнового вектора) фона. Внесем соответствующее обобщение в понятие времени релаксации.

Предположим, что электрон до столкновения имел направленную вдоль поля скорость v и при столкновении отклонился на угол θ (рис. 5.3) (будем считать, что абсолютное значение скорости при этом не изменилось — столкновение упругое). Так как при этом скорость электрона по абсолютной величине не изменится (столкновение упругое), то мы найдем время релаксации лишь для установления равновесия электронов в пределах определенного шарового слоя в пространстве импульсов (т. е. для электронов с заданной по абсолютной величине скоростью). В дальнейшем нам еще предстоит провести усреднение по энергиям.

В результате столкновения направленная составляющая скорости v_x уменьшится на величину

$$\Delta v = v [1 - \cos(\theta)]. \quad (5.30)$$

Если вероятность такого столкновения (или, точнее, число таких столкновений в 1 сек) $w(v, v')$, то соответствующее уменьшение скорости за 1 сек будет

$$\Delta v w(v, v') = v [1 - \cos(v, v')] w(v, v'). \quad (5.31)$$

Для того чтобы получить полное уменьшение скорости за 1 сек, мы должны просуммировать (или проинтегрировать) выражение (5.31) по всевозможным столкновениям, соответствующим отклонениям на разные углы:

$$\Delta \bar{v} = \sum_{v'} v [1 - \cos(v, v')] w(v, v') \quad (5.32)$$

или

$$\Delta \bar{v} = \sum_{\theta} v (1 - \cos \theta) w(\theta), \quad (5.33)$$

где

$$\theta = \angle(v, v').$$

Если за одну секунду направленная составляющая скорости v уменьшается на величину $\Delta \bar{v}$, то время, необходимое для ее полного исчезновения, будет $v/\Delta \bar{v}$; это время мы будем в дальнейшем считать эффективным временем свободного пробега или временем релаксации τ . Таким образом,

$$\tau = \frac{v}{\Delta \bar{v}} = \frac{1}{\sum_{\theta} (1 - \cos \theta) w(\theta)} \quad (5.34)$$

или

$$\frac{1}{\tau} = \sum_{\theta} (1 - \cos \theta) w(\theta). \quad (5.35)$$

Выражение (5.35) можно вывести точно так же и в волновой форме, т. е. получить время релаксации не как функцию скорости, а как функцию волнового числа. Действительно, предположим, что вероятность перехода электрона из состояния с импульсом k в состояние с импульсом k' будет $w(k, k')$; уменьшение направленного импульса при этом

$$\Delta k = k [1 - \cos(k, k')].$$

Не повторяя выкладок и рассуждений, приведенных для корпускулярного представления, получим

$$\frac{1}{\tau(k)} = \sum_{k'} [1 - \cos(k, k')] w(k, k') \quad (5.36)$$

или

$$\frac{1}{\tau(k)} = \sum_{k'} \frac{\Delta k}{k} w(k, k'). \quad (5.37)$$

Если $\omega(\Theta)$ — число столкновений, вызывающих отклонение электрона на угол Θ , то время между такими столкновениями $\tau(\Theta) = \frac{1}{\omega(\Theta)}$; таким образом, формулу (5.37) можно записать и в следующем виде:

$$\frac{1}{\tau(k)} = \sum \overline{(1 - \cos \Theta)} \frac{1}{\tau(\Theta)}. \quad (5.38)$$

Выражения (5.36) и (5.37) могут быть написаны и в интегральной форме, например:

$$\frac{1}{\tau(k)} = \int_{k'} \omega(k, k') [1 - \cos(k, k')] dk'. \quad (5.39)$$

Формула (5.38) является обобщением (5.7) с учетом того, что столкновения определенного сорта не обязательно приводят к полной потере скорости *).

Если ввести для каждого вида столкновений эффективное время свободного пробега и определить его следующим образом:

$$\tau_{i \text{ эф}} = \frac{\tau_i(\Theta)}{(1 - \cos \Theta)}, \quad (5.40)$$

то (5.38) примет вид, совершенно аналогичный (5.7):

$$\frac{\bar{1}}{\tau_{\text{эф}}} = \sum_i \frac{\bar{1}}{\tau_{\text{эф} i}}. \quad (5.40a)$$

Соотношения (5.38) и (5.39) все еще не являются строгими и точными — в них не учтено то обстоятельство, что для различных типов столкновений время свободного пробега электрона по-разному зависит от его энергии. Учет энергетической зависимости $\tau(\epsilon)$ и вычисление средних значений с учетом этой зависимости и является задачей следующих параграфов.

5.2. КИНЕТИЧЕСКОЕ УРАВНЕНИЕ (УЧЕТ ЭНЕРГЕТИЧЕСКОЙ ЗАВИСИМОСТИ ВРЕМЕНИ РЕЛАКСАЦИИ)

В самом начале этой главы мы убедились в том, что неправильный учет статистического разброса времени релаксации сразу же приводит к ошибке в абсолютном значении подвижности в 2 раза.

Еще к более серьезным последствиям может привести неучет (или неправильный учет) энергетической зависимости

*) Если мы положим средний угол отклонения $\Theta = 90^\circ$ (что и означает полную потерю направленной составляющей скорости), то (5.38) сразу же переходит в (5.7).

времени релаксации, так как этой зависимостью определяется температурный ход подвижности, зависимость подвижности от степени вырождения (концентрации носителей) и от характера и числа рассеивающих центров.

Выше мы уже упоминали, что при рассеянии на тепловых колебаниях и на ионах примеси зависимость времени релаксации от энергии электронов имеет совершенно различный характер. В первом случае, чем больше энергия (и волновой вектор) электрона, тем с большим числом фононов он может взаимодействовать и, следовательно, тем больше число соударений и меньше длина и время свободного пробега, во втором — чем больше энергия и скорость, тем меньше электрон отклоняется в поле иона, тем меньше $\cos \Theta$ в формуле (5.38) и тем больше, следовательно, эффективное время релаксации.

В дальнейшем нам предстоит рассмотреть конкретно время релаксации для каждого механизма рассеяния отдельно, сейчас же мы предположим, что вид зависимости $\tau(\epsilon)$ уже известен, и посмотрим, как вычислить $\bar{\tau}(\epsilon)$, т. е. произвести усреднение по всем энергиям. Казалось бы, что вопрос решается просто:

$$\bar{\tau} = \frac{\int \tau(\epsilon) dn(\epsilon)}{\int dn(\epsilon)}, \quad (5.41)$$

где $dn(\epsilon)$ — число электронов с данной энергией, определяется плотностью состояний $dG(\epsilon)$ и функцией распределения Ферми f_0

$$dn(\epsilon) = f_0(\epsilon) dG(\epsilon), \quad (5.42)$$

но в действительности формулы (5.41) и (5.42) в данном случае неприменимы. В самом деле, ведь нас интересует среднее значение времени релаксации для электронов, участвующих в электрическом токе (j_x^*), а следовательно, и усреднение надо проводить соответствующим образом:

$$\tau = \frac{\int \tau(\epsilon) dj_x(\epsilon)}{\int dj_x(\epsilon)} = \frac{\int \tau(\epsilon) v_x dn(\epsilon)}{\int v_x dn(\epsilon)}, \quad (5.43)$$

так как $dj_x = ev_x dn$ и $j_x = e \int v_x dn$, и при этом инте-

*) Поскольку везде будем рассматривать ток, проходящий через 1 см^2 поперечного сечения, то мы не будем делать различия между плотностью тока j и током I .

гирование необходимо вести не по энергиям, а по скоростям.

Нетрудно убедиться, что если мы воспользуемся для dn выражением (5.42), то получим в числителе и знаменателе (5.43) нули. Это и неудивительно: ведь функция Ферми именно и характеризует равновесное распределение электронов, в котором все макроскопические токи и потоки отсутствуют.

Равновесная функция распределения электронов по скоростям для двух случаев: вырожденного электронного газа $\mu > 0$ и невырожденного $\mu < -2$ представлена на рис. 5.4. В обоих случаях она симметрична относительно оси ординат: $dn(v_x) = dn(-v_x)$, и никакой ток в этих условиях не может существовать. На этом же рисунке показано изменение функции распределения под действием электрического поля.

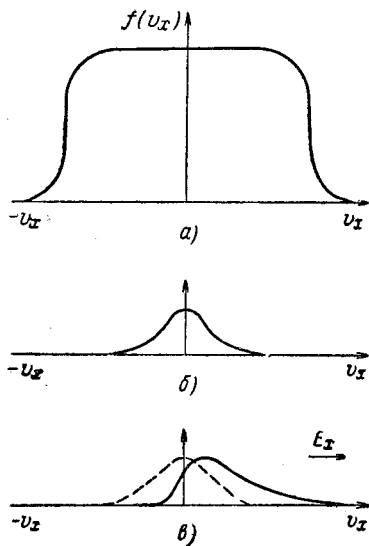


Рис. 5.4. Функция распределения электронов по скоростям для вырожденного электронного газа (а), для невырожденного электронного газа (б) и изменение функции распределения под действием электрического поля (в).

Электрический ток при наличии разности потенциалов возникает в результате того, что функция распределения в этом случае «деформируется», это проиллюстрировано рис. 5.4, в, где для определенности рассмотрен случай невырожденного электронного газа.

Рассмотрим состояние электронного газа при наличии разности потенциалов более подробно. Под действием электрического поля (рис. 5.4, в) все электроны получают дополнительную скорость в направлении v_x — вся функция распределения смещается в этом направлении. Если бы не было столкновений, то этот процесс ничем бы не компенсировался, а вся кривая $f(v)$ сместилась бы в направлении $+\infty$ по оси v_x . Но в действительности при каждом столкновении (или в результате нескольких столкновений)

любой электрон теряет свою направленную составляющую скорости, возвращаясь на свое «равновесное место», и весь процесс повторяется сначала. Если бы все электроны испытывали столкновения одновременно, то функция распределения все время пульсировала бы между двумя крайними состояниями — равновесным f_0 и смещенным f_τ — на отрезок $v_x = (eE/m) \tau$ из состояния равновесия. В действительности же устанавливается динамическое равновесие между действием столкновений и поля, т. е. какое-то новое промежуточное стационарное распределение f , которое можно найти из того условия, что действие соударений и поля на него должны компенсировать друг друга:

$$\frac{df}{dt} = \left(\frac{\partial f}{\partial t} \right)_{\text{поля}} + \left(\frac{\partial f}{\partial t} \right)_{\text{столкн}} = 0. * \quad (5.44)$$

Время релаксации для электронов с различной энергией различно, поэтому под действием поля функция распределения не только смещается, но и деформируется; так, например, при рассеянии на ионах больше сместится часть функции, соответствующая электронам с большими скоростями (так как $\tau(v)$ для них в этом случае больше), при рассеянии на акустических колебаниях, наоборот, больше сместятся электроны с малыми скоростями.

Подставляя в (5.43) $dn(\epsilon)$, выраженное через неравновесную функцию f ,

$$dn(\epsilon) = f dG(\epsilon),$$

получаем

$$\tau = \frac{\int \tau v_x f dG(\epsilon)}{\int v_x f dG(\epsilon)} = \frac{j_1(f)}{j_2(f)}. \quad (5.43a)$$

В дальнейшем будем предполагать, что направленная добавка к скорости электрона, получаемая в поле, мала по сравнению с тепловой скоростью и, следовательно, функция f должна несильно отличаться от равновесной функции распределения f_0 , поэтому мы можем представить ее в виде

$$f = f_0 + f_1, \quad (5.45)$$

*) Условие (5.44) для определения функции распределения в неравновесных условиях f было впервые сформулировано Больцманом и называется кинетическим уравнением Больцмана.

где f_1 — небольшая добавка. Но если мы подставим (5.45) в (5.43а), то интегралы в числителе и знаменателе (5.43а) $j_1(f)$ и $j_2(f)$ разобьются на два:

$$j_1(f) = j_1(f_0) + j_1(f_1) \quad (5.46)$$

и

$$j_2(f) = j_2(f_0) + j_2(f_1);$$

при этом, так как f_0 характеризует равновесное состояние,

$$j_1(f_0) = j_2(f_0) = 0.$$

Следовательно,

$$\bar{\tau} = \frac{\int \tau v_x f_1 dg(\epsilon)}{\int v_x f_1 dg(\epsilon)}. \quad (5.47)$$

Можно точно так же показать, что все кинетические коэффициенты (электро- и теплопроводность, коэффициенты в термоэлектрических, гальвано- и термомагнитных явлениях) будут выражаться через f_1 ; поэтому нахождение этой небольшой неравновесной добавки к равновесной функции распределения является основной задачей теории явлений переноса*).

Для того чтобы найти f_1 , напишем уравнение Больцмана более подробно, считая при этом, что в полупроводнике существует не только электрическое поле, но и градиент температуры. Это обобщение нам понадобится для исследования электронной теплопроводности и термоэлектрических явлений.

Рассмотрим вначале «полевую» производную от функции распределения, входящую в (5.44), т. е. изменение функции распределения под действием поля.

Будем рассматривать стационарное состояние, следовательно, время явно не может входить в выражение для f , а будет влиять на нее через скорости электронов и их координаты, которые уже явно зависят от времени $dg/dt = \Rightarrow v$ и $dv/dt = w = F/m$, где F — полная сила, действующая на электрон, т. е. сила Лоренца:

$$F = e \left\{ E + \frac{1}{c} [vH] \right\}. \quad (5.48)$$

*). Второй, не менее важной задачей является нахождение конкретного выражения τ — времени релаксации, как функции от всех микро- и макроскопических параметров: энергии, эффективной массы электрона, температуры, характера и числа дефектов и т. д.

Электроны перемещаются в пространстве координат за время dt на отрезок $d\mathbf{r} = \mathbf{v} dt$ и в пространстве скоростей на отрезок $d\mathbf{v} = (\mathbf{F}/m) dt$.

Следовательно, распределение, которое в начальный момент t было в точке 1 ($\mathbf{v} - \mathbf{w} dt$, $\mathbf{r} - \mathbf{v} dt$), через промежуток времени dt попадает в точку 2 (\mathbf{v} , \mathbf{r}); это значит, что

$$f_{t+dt}(\mathbf{v}, \mathbf{r}) = f_t(\mathbf{v} - \mathbf{w} dt, \mathbf{r} - \mathbf{v} dt). \quad (5.48a)$$

Согласно (5.48a) изменение (точнее, производная) f по времени за счет ее перемещения под действием поля (скоростей и ускорений) из точки 1 в точку 2 будет иметь вид

$$\left(\frac{\partial f}{\partial t}\right)_{\text{поля}} = \frac{f(t+dt) - f(t)}{dt} = -\mathbf{v}\nabla_{\mathbf{r}}f - \mathbf{w}\nabla_{\mathbf{v}}f, \quad (5.49)^*$$

или, подставляя вместо ускорения \mathbf{w} его выражение, получим окончательно

$$\left(\frac{\partial f}{\partial t}\right)_{\text{поля}} = \mathbf{v}\nabla_{\mathbf{r}}f - \frac{e}{m} \left\{ \mathbf{E} + \frac{1}{c} [\mathbf{v}\mathbf{H}] \right\} \nabla_{\mathbf{v}}f. \quad (5.49a)$$

Перейдем теперь ко второму члену в (5.44) — изменению функции распределения под действием столкновений. Полное изменение функции распределения в данной точке (точнее, в единице объема, так как f — это плотность) фазового пространства за счет столкновений состоит из двух частей:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{столкн}} = b - a, \quad (5.50)$$

где b — число частиц, пришедших в результате столкновений из всех точек фазового пространства в данную точку, и a — число частиц, ушедших под действием столкновений из данной точки во все точки фазового пространства. Нетрудно в общем виде написать число и тех и других переходов (но неизмеримо трудней вычислить). Число переходов из точки \mathbf{v}' в точку \mathbf{v} будет пропорционально вероятности такого перехода $\omega(\mathbf{v}', \mathbf{v})$ и числу частиц в точке \mathbf{v}' : $dn = f(\mathbf{v}') dg(\mathbf{v}')$, а также числу пустых мест в точке \mathbf{v} , т. е. $[1 - f(\mathbf{v})]$.

Таким образом,

$$db = f(\mathbf{v}') \omega(\mathbf{v}', \mathbf{v}) dg(\mathbf{v}') [1 - f(\mathbf{v})]. \quad (5.51)$$

*) Здесь и в дальнейшем выражения $\nabla_{\mathbf{r}}f$ и $\nabla_{\mathbf{v}}f$ означают градиенты функции f в пространстве координат и импульсов.

Для того чтобы получить b , мы должны проинтегрировать это выражение по всему фазовому пространству:

$$b = \int_{v'} \omega(\mathbf{v}', \mathbf{v}) f(\mathbf{v}') [1 - f(\mathbf{v})] dg(\mathbf{v}'). \quad (5.52)$$

Число обратных переходов da пропорционально их вероятности $\omega(v, v')$ плотности частиц в точке v , т. е. $f(v)$, и числу пустых мест в точке v' , т. е. $[1 - f(v')] dg$:

$$da = \omega(\mathbf{v}, \mathbf{v}') f(\mathbf{v}) [1 - f(\mathbf{v}')] dg(\mathbf{v}'). \quad (5.53)$$

Суммируя переходы из рассматриваемой точки v во все точки фазового пространства v' , получаем

$$a = \int_{v'} \omega(\mathbf{v}, \mathbf{v}') f(\mathbf{v}) [1 - f(\mathbf{v}')] dg(\mathbf{v}'). \quad (5.54)$$

Можно показать из очень общих соображений так называемого принципа микроскопической обратимости, что вероятности прямых и обратных переходов должны быть одинаковы: $\omega(v', v) = \omega(v, v')$; согласно (5.52) и (5.54) полное изменение функции распределения в результате столкновений

$$\left(\frac{\partial f}{\partial t} \right)_{\text{столкн}} = \int \{ f(\mathbf{v}') [1 - f(\mathbf{v})] - f(\mathbf{v}) [1 - f(\mathbf{v}')] \} \times \\ \times \omega(\mathbf{v}, \mathbf{v}') dg(\mathbf{v}') = \int [f(\mathbf{v}') - f(\mathbf{v})] \omega(\mathbf{v}, \mathbf{v}') dg(\mathbf{v}'). \quad (5.55)$$

В ряде случаев удобнее рассматривать функцию распределения f как функцию не от скоростей $(\mathbf{v}, \mathbf{v}')$, а от волновых векторов $(\mathbf{k}, \mathbf{k}')$; проделав совершенно аналогичные выкладки, получим (здесь для сокращения записи волновые векторы \mathbf{k} и \mathbf{k}' вынесены в индексы)

$$\left(\frac{\partial f}{\partial t} \right)_{\text{столкн}} = \int [f_{\mathbf{k}'} (1 - f_{\mathbf{k}}) - f_{\mathbf{k}} (1 - f_{\mathbf{k}'})] \omega_{\mathbf{k}\mathbf{k}'} dg_{\mathbf{k}'} = \\ = \int (f_{\mathbf{k}'} - f_{\mathbf{k}}) \omega_{\mathbf{k}\mathbf{k}'} dg_{\mathbf{k}'}. \quad (5.55a)$$

Подставляя (5.49a) и (5.55a) в (5.44), получаем интегральное уравнение для определения неравновесной функции распределения $f_{\mathbf{k}}$:

$$\mathbf{v} (\nabla_{\mathbf{r}} f_{\mathbf{k}}) + \frac{e}{m} \left[\mathbf{E} + \frac{1}{c} (\mathbf{v}_{\mathbf{k}} \mathbf{H}) \right] (\nabla_{\mathbf{v}} f_{\mathbf{k}}) = \\ = \int (f_{\mathbf{k}'} - f_{\mathbf{k}}) \omega_{\mathbf{k}'\mathbf{k}} dG_{\mathbf{k}'}. \quad (5.56)$$

Решение уравнения (5.56) в таком общем виде представляет весьма большие трудности, однако при некоторых упрощающих предположениях оно допускает точное аналитическое решение. Полагая $f_k = f_{0k} + f_{1k}$ и учитывая, что для равновесного состояния $(\partial f / \partial t)_{\text{столкн}} = 0$, можно переписать (5.55) в виде

$$\left(\frac{\partial f_k}{\partial t} \right)_{\text{столкн}} = \int (f_{1, k'} - f_{1, k}) \omega_{kk'} dG_{k'}, \quad (5.57)$$

исключив, таким образом, из правой части (5.56) f_0 . Напротив, из левой части (5.56) при известных условиях можно исключить f_1 . Действительно, все члены в $(\partial f / \partial t)_{\text{поля}}$, за исключением члена, содержащего магнитное поле (см. ниже), не обращаются в нуль при замене $f_k \rightarrow f_{0k}$, поэтому мы можем (учитывая, что $f_{0k} \gg f_{1k}$) переписать (5.56) в виде

$$\bar{v} \left(\nabla_r f_{0k} + \frac{eE}{m} (\nabla_v f_{0k}) + \right. \\ \left. + \frac{1}{c} [\mathbf{v}_k \mathbf{H}] (\nabla_v f_{1k}) \right) = \int (f_{1k'} - f_{1k}) \omega_{kk'} dg_{k'}, \quad (5.58)$$

и в случае отсутствия магнитных полей можно переписать кинетическое уравнение в виде

$$v (\nabla_r f_{0k}) + \frac{eE}{m} (\nabla_v f_{0k}) = \int [f_{1k'} - f_{1k}] \omega_{kk'} dg_{k'}. \quad (5.59)$$

Даже в таком упрощенном виде решение кинетического уравнения представляет большие трудности.

Действительно, в правую часть (5.59) входит интеграл по всему фазовому пространству, это значит, что изменение функции распределения под влиянием столкновений в точке k зависит от ее вида не только в данной точке, но и во всех точках фазового пространства; этим самым предположение о существовании времени релаксации τ , сделанное нами в гл. 4 и в начале настоящей главы, становится весьма сомнительным. Действительно, согласно (4.1) изменение функции f под действием столкновений запишется так:

$$\left(\frac{\partial f}{\partial t} \right)_{\text{столкн}} = \frac{f_0 - f}{\tau}. \quad (5.60)$$

Но в свете сказанного выше определенное таким образом τ должно зависеть от значений f во всем фазовом простран-

стве, т. е. должно быть функцией от времени и от мгновенного распределения электронов по скоростям; если это действительно так, то введение времени релаксации как некоторой характеристики системы теряет смысл. Однако в ряде случаев есть достаточно серьезные основания для такого упрощения задачи [т. е. введения понятия времени релаксации и выражения $(\partial f / \partial t)_{\text{столкн}}$ в виде (5.60)].

Во-первых, следует учесть, что, как это было показано выше, в ряде интересующих нас случаев столкновения являются почти упругими, при этом: а) мы должны проводить интегрирование не по всему фазовому пространству, а лишь по сфере постоянной энергии (точнее, по шаровому слою толщиной $\Delta \mathcal{E}$), характеризующему «неупругость» соударений); б) так как столкновения происходят только внутри этого слоя, то $(\partial f / \partial t)_{\text{столкн}}$ зависит только от распределения электронов в нем. При этих условиях может существовать время релаксации для данного (и любого другого) слоя, являющееся функцией только энергии.

При этом остается открытым вопрос: как же устанавливается равновесное распределение электронов по энергиям, если электроны в разных слоях не обмениваются энергией друг с другом? Дело в том, что соударения являются все же почти упругими, и поэтому обмен энергией все же происходит, но значительно медленней (т. е. со значительно большим временем релаксации), чем обмен импульсами. К этому надо добавить, что межэлектронные столкновения, когда они происходят в достаточном количестве, значительно ускоряют установление равновесного распределения по энергиям.

Мы покажем, что для упругих соударений время релаксации вводится строго; в ряде других случаев этот метод решения кинетического уравнения не может быть строго обоснован, но тем не менее он дает хорошие результаты.

Если учесть, что $f_1 = f - f_0$ есть малое изменение функции распределения, то мы можем разложить его в ряд по v и ограничиться первыми исчезающими членами. Так как поле направлено по оси x и функция f смещена и деформирована по оси v_x , то очевидно, что в первом члене должно быть v_x в первой степени, а v_y и v_z — не ниже второй, так как функция распределения должна сохранять симметрию по этим осям; иными словами, f_1 может быть только функцией v_x и энергии. Таким образом,

в этом приближении

$$f_1(v_x) = v_x \chi(\varepsilon), \quad (5.61)$$

где $\chi(\varepsilon)$ — пока неизвестная функция энергии.

Далее, учитывая (5.61) и то, что соударения упругие (рис. 5.3), получим

$$f(v') = v'_x \chi(\varepsilon) = v_x \cos \Theta \chi(\varepsilon). \quad (5.62)$$

Предположим также, что вероятность рассеяния является только функцией от угла:

$$\omega(v, v') = \omega(\Theta).$$

Подставляя $f(v)$ и $f(v')$ в интеграл столкновений (5.57), получаем

$$\left(\frac{\partial f}{\partial t}\right)_{\text{столкн}} = \int \omega(v, v') \chi(\varepsilon) v_x (1 - \cos \theta) dG(v'). \quad (5.62a)$$

Так как интегрирование происходит по v' и при $\varepsilon = \text{const}$, то можно $\chi(\varepsilon) v_x = f_1$ вынести за знак интеграла; таким образом, получаем

$$\left(\frac{\partial f}{\partial t}\right)_{\text{столкн}} = f_1 \int \omega(v, v') (1 - \cos \theta) dG(v'). \quad (5.63)$$

Интеграл в (5.63) в точности соответствует выражению для обратного времени релаксации (5.39):

$$\tau^{-1} = \int \omega(v, v') [1 - \cos(v, v')] dG(v'), \quad (5.64)$$

следовательно,

$$\left(\frac{\partial f}{\partial t}\right)_{\text{столкн}} = \frac{f_1}{\tau} = \frac{f - f_0}{\tau}. \quad (5.65)$$

Таким образом при сделанных нами предположениях можно не только ввести время релаксации, но и получить для него явное выражение (5.64).

Подставляя (5.65) в (5.59) (и опуская для простоты индексы k), получаем явное выражение для f_1 :

$$f_1 = \left(\frac{\partial f}{\partial t}\right)_{\text{поля}} \tau = \tau \left[\mathbf{v} \nabla_r f_0 + \frac{e}{m} (\mathbf{E} \nabla_v f_0) \right]. \quad (5.66)$$

Учитывая, что

$$\nabla_r f_0 = \frac{\partial f_0}{\partial T} \nabla T + \frac{\partial f_0}{\partial \mu} \nabla \mu = \frac{\partial f_0}{\partial \varepsilon} \left[\frac{\varepsilon - \mu}{T} \nabla T - \nabla \mu \right]$$

и

$$\nabla_{\mathbf{v}} f_0 = \frac{\partial f_0}{\partial \varepsilon} m \mathbf{v} \text{ и } \mathbf{E} = -\nabla \Phi,$$

после простых преобразований получаем

$$f_1 = \tau \left[\frac{\mu - \varepsilon}{T} \nabla T - \nabla (\mu - e\Phi) \right] \frac{\partial f_0}{\partial \varepsilon} v_x. \quad (5.67)$$

5.3. ФЕНОМЕНОЛОГИЧЕСКИЙ АНАЛИЗ ЯВЛЕНИЙ ПЕРЕНОСА

Знание неравновесной функции распределения позволяет вычислить все кинетические коэффициенты: электро- и теплопроводность, термоэлектрические коэффициенты (Зеебека, Пельтье и Томсона), коэффициенты гальвано- и термомагнитных явлений. Оставив изучение двух последних классов явлений на будущее (что мы вынуждены сделать, так как исключили из кинетического уравнения члены, содержащие магнитное поле), напомним выражение для плотностей электрического тока и потока энергии при отсутствии магнитного поля:

$$\mathbf{j} = \overline{e n \mathbf{v}} = e \int \mathbf{v} f_1 dG(\varepsilon) \quad (5.68)$$

и

$$\mathbf{q} = \overline{\varepsilon n \mathbf{v}} = \int \varepsilon \mathbf{v} f_1 dG(\varepsilon). \quad (5.69)$$

Если считать, что время релаксации существует, то можно сразу же написать феноменологические выражения (5.67) для f_1 и (5.66) и (5.63) для плотностей тока и потока энергии. Таким образом, весь вопрос сводится к нахождению выражения для времени релаксации. Так как f_1 согласно (5.67) является линейной функцией градиентов потенциала и температуры, то электрический ток и поток энергии должны быть линейными функциями от этих параметров. Поэтому в самом общем виде можно написать

$$\mathbf{j} = L_{EE} \mathbf{E} + L_{ET} \nabla T \quad (5.70)$$

и

$$\mathbf{q} = L_{TE} \mathbf{E} + L_{TT} \nabla T, \quad (5.71)$$

где все L — кинетические коэффициенты, подлежащие определению.

По определению, коэффициент пропорциональности между напряжением и током при $\nabla T = 0$ называется электропроводностью σ , следовательно,

$$\sigma = L_{EE}. \quad (5.72)$$

Коэффициент термо-э. д. с. α измеряется при наличии градиента температуры и отсутствии тока (по существу, это является определением термоэлектродвижущей силы). Таким образом, приравнявая в (5.70) \mathbf{j} нулю, находим

$$\mathbf{E} = -\frac{L_{ET}}{L_{EE}} \nabla T, \quad (5.73)$$

следовательно,

$$\alpha = \frac{L_{ET}}{L_{EE}}. \quad (5.74)$$

Коэффициент теплопроводности измеряется также при отсутствии тока в цепи. Подставив \mathbf{E} , соответствующее $\mathbf{j} = 0$, из (5.73) в (5.71), находим

$$\mathbf{q} = \left(-\frac{L_{TE}L_{ET}}{L_{EE}} + L_{TT} \right) \nabla T. \quad (5.75)$$

Следовательно, удельная теплопроводность

$$\kappa = L_{TT} - \frac{L_{TE}L_{ET}}{L_{EE}}. \quad (5.76)$$

Аналогично можно получить феноменологические выражения для коэффициентов Пельтье и Томсона. Для того чтобы получить конкретные выражения для всех перечисленных выше кинетических коэффициентов, необходимо подставить в (5.68) и (5.69) выражения для f_1 и провести интегрирование по энергиям. Но для проведения этой программы мы должны получить зависимость времени релаксации τ от энергии в явном виде — следующий параграф и посвящен рассмотрению этого вопроса.

5.4. ВЫЧИСЛЕНИЕ ВРЕМЕНИ РЕЛАКСАЦИИ

РАССЕЯНИЕ НА ИОНАХ ПРИМЕСИ

Вычисление времени релаксации является одной из наиболее сложных задач теории явлений переноса, поэтому мы здесь рассмотрим лишь основные вехи расчета для наи-

более простого случая — рассеяния электронов на ионах примеси.

Анализ рассеяния электронов на ионах примеси при известных допущениях можно проводить по классической схеме, т. е. рассматривая электрон как точечный заряд e с массой m , движущейся в поле другого неподвижного заряда — иона, масса которого много больше массы первого (см. рис. 5.1). При этом для вычисления времени релаксации мы воспользуемся формулой (5.34):

$$\frac{1}{\tau} = \sum_{\Theta} (1 - \cos \Theta) \omega(\Theta), \quad (5.77)$$

где заменим

$$1 - \cos \Theta = 2 \sin^2 \frac{\Theta}{2} = \frac{2}{1 + \operatorname{ctg}^2 \frac{\Theta}{2}}.$$

Зависимость угла отклонения Θ от прицельного расстояния b была выведена впервые Резерфордом (для анализа рассеяния α -частиц):

$$\operatorname{tg} \frac{\Theta}{2} = \frac{e^2}{\kappa m v^2 b}, \quad (5.78)$$

где κ — диэлектрическая постоянная среды; следовательно, согласно (5.77)

$$1 - \cos \Theta = \left(-\frac{\Delta v_x}{v_x} \right) = \frac{2}{1 + \left(\frac{\kappa m v^2 b}{e^2} \right)^2}.$$

Таким образом, вероятность того, что электрон отклонится на угол Θ , свелась к определению вероятности того, что он пройдет на расстоянии b от иона. Последнюю же можно приближенно оценить как отношение площади кольца $dS = 2\pi b db$ к площади окружности, очерченной радиусом, равным среднему расстоянию между ионами ($S = \pi a^2/4$), умноженной на число таких «столкновений» в единицу времени, т. е. v/a .

Таким образом,

$$\omega = \frac{v}{a} \frac{2\pi b db}{\pi a^2/4} = \frac{4v}{a^3} 2b db. \quad (5.79)$$

Подставив (5.79) в (5.77) и заменив сумму интегралом, мы можем теперь получить явное выражение для τ^{-1} :

$$\tau^{-1} = \frac{4v}{a^3} \int_0^{\frac{a}{2}} \frac{2b \, db}{1 + \left(\frac{\kappa m v^2}{e^2} b \right)^2} \quad (5.80)$$

или после интегрирования

$$\begin{aligned} \tau^{-1} &= \frac{4v}{a^3} \frac{e^4}{\kappa^2 m^2 v^4} \ln \left[1 + \frac{m v^2 / e^2}{\kappa a / 2} \right] = \\ &= \frac{2v}{a} \left(\frac{\varepsilon_i}{2\varepsilon} \right)^2 \ln \left[1 + \left(\frac{2\varepsilon}{\varepsilon_i} \right) \right], \end{aligned} \quad (5.81)$$

где $\varepsilon = m v^2 / 2$ — кинетическая энергия электрона и $\varepsilon_i = e^2 / (\kappa a / 2)$ — потенциальная энергия электрона в поле иона при максимальном прицельном расстоянии.

Таким образом,

$$\tau = \frac{a}{2v} \left(\frac{2\varepsilon}{\varepsilon_i} \right)^2 \left[\frac{1}{\ln \left(1 + \frac{2\varepsilon}{\varepsilon_i} \right)} \right] \quad (5.82)$$

и длина свободного пробега

$$l = \tau v \approx A \varepsilon^2, \quad (5.83)$$

где A — величина, которую в первом приближении можно считать постоянной (выражение, стоящее под знаком логарифма, можно в первом приближении считать постоянным, так как логарифм — слабо меняющаяся функция).

При выводе формул (5.82) и (5.83) мы сделали ряд упрощающих предположений:

- пренебрегли всеми квантовыми эффектами (это возможно, когда $\varepsilon \gg \varepsilon_i$);
- не учли экранировку поля ионов свободными электронами, что допустимо при малой концентрации последних;
- считали, что рассеяние около каждого иона происходит независимо, в действительности же, если построить из ионов правильную решетку, то они вообще перестанут быть рассеивающими центрами (а лишь изменят энергетический спектр электрона);
- заменили кристалл непрерывной средой с диэлектрической постоянной κ , что справедливо при малой концентрации ионов и не слишком большой диэлектрической постоянной.

Поэтому полученный нами результат имеет ограниченную область применений; тем не менее он очень важен, так как в ряде случаев дает картину явлений, правильную не только качественно, но и количественно.

РАССЕЯНИЕ ЭЛЕКТРОНОВ НА ТЕПЛОВЫХ КОЛЕБАНИЯХ РЕШЕТКИ

Этот случай значительно более сложен и может быть рассмотрен последовательно только квантовомеханическим путем. Схема расчета качественно сводится к следующему. Мы уже упоминали, что в идеальном кристалле (при строго периодическом поле) зонные электроны описываются незатухающими модулированными плоскими волнами. Это значит, что длина свободного пробега электрона при этих условиях бесконечна, модулированные волны

$$\psi_1 = u_1(\mathbf{r}) e^{2\pi i(\mathbf{k}_1 \mathbf{r} - \nu_1 t)}, \quad \psi_2 = u_2(\mathbf{r}) e^{2\pi i(\mathbf{k}_2 \mathbf{r} - \nu_2 t)} \quad (5.84)$$

распространяются беспрепятственно по всему кристаллу, вероятность рассеяния, т. е. перехода одной волны в другую, равна нулю. Причиной рассеяния могут явиться любые нарушения периодического потенциала (заряженные и нейтральные примеси, точечные, линейные, плоские и объемные дефекты и, наконец, тепловые колебания), так как если мы их включим в уравнение Шредингера, то решения типа (5.84) перестанут быть стационарными. В данном случае нас интересует рассеяние на тепловых колебаниях, т. е. мы должны вычислить вероятность перехода из одного электронного состояния в другое под действием нарушения периодического потенциала $\Delta V = V_{\text{д}} - V_{\text{и}}$ (индексами «д» и «и» обозначены потенциал деформированной и идеальной решеток), вызванного смещением атома из положения равновесия на величину u . Таким образом, в нашем случае возмущающий потенциал

$$\Delta V \approx V_{\text{д}}(r) - V_{\text{и}}(r). \quad (5.85)$$

Вероятность такого перехода пропорциональна квадрату матричного элемента возмущения:

$$\omega \sim [(\Delta V)_{1,2}]^2 = \left| \int \Phi_1 \Delta V \Phi_2 dv \right|^2, \quad (5.86)$$

где Φ_1 и Φ_2 — волновые функции системы до и после перехода; так как в процессе рассеяния меняется не только

электронное, но и фоновое состояние, то в Φ должны быть включены волновые функции электронов и фононов:

$$\Phi_{1,2} = [u_k(r) e^{2\pi i k r} \prod_q \psi_{qN}]_{1,2}, \quad (5.87)$$

где ψ_{qN} — фоновые волновые функции, описывающие нормальные колебания решетки; смещения атомов и соответственно ΔV должны быть также разложены по нормальным колебаниям решетки.

Определение матричных элементов и вероятностей переходов согласно (5.85)—(5.87) приводит к весьма громоздким вычислениям и проводится по-разному для рассеяния на акустических и оптических колебаниях и в атомных и ионных кристаллах, поэтому мы их здесь опустим и приведем лишь основные результаты расчета.

Время релаксации при рассеянии акустической ветви тепловых колебаний

$$\tau = \frac{9\pi}{4\sqrt{2}} \frac{M\omega^2\hbar^4}{\Omega_0 C^2 m^{3/2} kT} \frac{1}{\sqrt{\epsilon}}, \quad (5.88)$$

$$l = v\tau = \frac{9\pi}{4} \frac{M\omega^2\hbar^4}{\Omega_0 C^2 m kT}, \quad (5.89)$$

где Ω_0 — объем элементарной ячейки; ω — скорость продольной звуковой волны; M — масса атома и C — константа взаимодействия электронов с фононами, характеризующая вероятность перехода ($\omega_{1,2}$) состояния ψ_{k_1} в состояние ψ_{k_2} под действием тепловых колебаний. Нахождение C в результате интегрирования выражения (5.86) и является центральной и наиболее сложной задачей при изучении рассеяния носителей на тепловых колебаниях решетки.

Бардиным и Шокли был развит другой метод вычисления времени релаксации для атомной решетки, так называемый метод деформационного потенциала, сущность которого заключается в следующем.

Как известно (см. гл. 1), положение и ширина зон в кристалле определяются расстоянием между соседними атомами (или постоянной решетки). При продольных колебаниях это расстояние меняется, что, в свою очередь, вызывает колебания дна зоны проводимости, которые равносильны возникновению потенциальных скачков на пути электрона, а следовательно, должны вызывать его рассеяние. При этом подходе вероятность рассеяния выражается

через так называемую константу деформационного потенциала E , характеризующую сдвиг дна зоны проводимости (для электронов) или сдвиг верхнего края заполненной зоны (для дырок) при однородной деформации кристалла; константа E с точностью до коэффициента равна константе взаимодействия C , характеризующей взаимодействие электронов с фононами.

Время релаксации при рассеянии на оптических колебаниях ионных кристаллов:

— в области низких температур ($T \ll \theta$)

$$\tau = \frac{3\sqrt{2}}{4\pi} \frac{Ma^3 (\hbar\omega_0)^{\frac{3}{2}}}{z^2 e^4 m^{\frac{1}{2}}} e^{\frac{\hbar\omega_0}{kT}}, \quad (5.90)$$

$$l = \tau v = \frac{3a}{2\pi} \left(\frac{M}{m}\right) \left(\frac{\hbar\omega_0}{ze^2}\right)^2 e^{\frac{\hbar\omega_0}{kT}} \sqrt{\frac{e}{kT}}, \quad (5.90a)$$

где z — заряд ионов; ω_0 — частота оптических колебаний и M — средняя масса ионов; a — постоянная решетки;

— в области высоких температур ($T \gg \theta$)

$$\tau = \frac{\sqrt{2}}{4\pi} \frac{Ma^3 (\hbar\omega_0)^2 e^{\frac{1}{2}}}{z^2 e^4 m^{\frac{1}{2}} kT}, \quad (5.91)$$

$$l = \tau v = \frac{a}{2\pi} \frac{M}{m} \left(\frac{\hbar\omega_0}{ze^2}\right)^2 \frac{e}{kT}. \quad (5.92)$$

Почти во всех явлениях переноса весьма серьезную роль играет зависимость длины свободного пробега электрона от энергии (так как она определяет среднюю энергию электронов в потоке). Как мы видели, во всех рассмотренных выше случаях эта зависимость носит степенной характер, таким образом, мы можем написать

$$l \sim \varepsilon^r. \quad (5.93)$$

При рассеянии на ионах примеси $r = 2$, при рассеянии на акустических колебаниях $r = 0$, при рассеянии на оптических колебаниях ионных кристаллов $r = 1/2$ при $T < \theta$ и $r = 1$ при $T > \theta$.

Получив явные выражения для τ (ϵ), мы можем их подставить в (5.67), а затем f_1 — в (5.68) и таким образом вычислить плотность тока, электропроводность и подвижность.

При этом для невырожденного электронного газа с изотропной эффективной массой получают следующие выражения для подвижности.

Подвижность при рассеянии на акустической ветви тепловых колебаний

$$u = \frac{4}{3} \frac{e}{\sqrt{\pi}} \frac{e}{m} \frac{\tau_{0\text{ак}}}{(kT)^{\frac{3}{2}}}, \quad (5.94)$$

где $\tau_{0\text{ак}}$ — коэффициент в выражении для времени релаксации, не зависящий от температуры и энергии,

$$\tau_{0\text{ак}} = \frac{9\pi}{4} \frac{M\omega^2 \hbar^4}{\sqrt{2} \Omega_0 C^2 m^{\frac{3}{2}}}. \quad (5.95)$$

Подвижность при рассеянии на ионах примеси

$$u = \frac{8}{\sqrt{\pi}} \frac{e}{m} \tau_{0\text{и}} (kT)^{\frac{3}{2}}, \quad (5.96)$$

где [см. (5.82) и (5.83)]

$$\tau_{0\text{и}} = \frac{\kappa^2 m^2 A}{2\pi N_{\text{и}} e^4} \quad (5.97)$$

($N_{\text{и}}$ — число ионов в единице объема $N_{\text{и}} = a^{-3}$).

Подвижность при рассеянии на оптических колебаниях ионных кристаллов:

— в области низких температур ($T < \theta$)

$$u = \frac{e}{m} \tau_{0\text{опт}} e^{\frac{\hbar\omega_0}{kT}}, \quad (5.96a)$$

где

$$\tau_{0\text{опт}} = \frac{3\sqrt{2}}{4\pi} \frac{M a^3 (\hbar\omega_0)^{3/2}}{z^2 e^4 m^{1/2}}; \quad (5.97a)$$

— в области высоких температур ($T \gg \theta$)

$$u = \frac{8}{3} \frac{e}{\sqrt{\pi}} \frac{e}{m} \frac{\tau_{0\text{опт}}}{(kT)^{1/2}}, \quad (5.96b)$$

где

$$\tau_{0\text{опт}} = \frac{\sqrt{2}}{4\pi} \frac{Ma^3 (\hbar\omega_0)^2}{z^2 e^4 m^{1/2}}. \quad (5.976)$$

При произвольной степени вырождения и простой зонной структуре из (5.68) в общем случае получаем следующие выражения для электропроводности

$$\sigma = \frac{j}{E} = \frac{16\pi m e^2}{3\hbar^3} l_0(T) (kT)^{r+1} F_r(\mu^*) \quad (5.68a)$$

и подвижности

$$u = \frac{\sigma}{en}, \quad (5.68б)$$

где

$$n = 4\pi \left(\frac{2mkT}{\hbar^2} \right)^{3/2} F_{1/2}(\mu^*).$$

5.5. ЯВЛЕНИЯ В СИЛЬНЫХ ЭЛЕКТРИЧЕСКИХ ПОЛЯХ

Все выводы относительно подвижности и электропроводности, сделанные выше, основывались на двух предположениях: 1) направленная добавка к скорости Δv_x [см. формулу (5.2)] мала по сравнению с тепловой скоростью v_0 и 2) концентрация носителей электричества не зависит от поля и остается равновесной, определяемой статистикой Максвелла — Больцмана или Ферми. В действительности оба эти условия соблюдаются в полупроводниках до определенных для каждого материала и температуры критических полей; новые эффекты, которые происходят при нарушении хотя бы одного из этих условий, называются явлениями в сильных полях. На них мы и остановимся коротко в этом параграфе.

В первую очередь — это отступления от закона Ома, т. е. нарушение линейной зависимости тока от напряжения. Согласно формулам

$$j = \sigma E \text{ и } \sigma = en\mu \quad (5.98)$$

это будет иметь место, когда либо подвижность, либо концентрация носителей начнут зависеть от поля.

Рассмотрим каждое из этих явлений в отдельности.

Согласно выражению

$$u = \frac{e}{m} \frac{l}{v} \quad (5.99)$$

подвижность начинает зависеть от поля с того момента, как скорость v в (5.3а) перестает быть постоянной, т. е. когда добавкой Δv к скорости за счет поля нельзя пренебречь по сравнению с тепловой скоростью v_0 , иными словами, критерием слабого поля является условие

$$\Delta v \ll v_0. \quad (5.100)$$

Добавку к тепловой скорости Δv можно оценить, исходя из закона сохранения энергии (в дальнейшем мы увидим, что эта оценка не точна):

$$\frac{m}{2} [(\overline{v_0 + \Delta v})^2 - \overline{v_0^2}] = eE\bar{l} \quad (5.101)$$

или, так как $\overline{\Delta v \cdot v_0} = 0$ (поскольку $\overline{v_0}$ направлено беспорядочно, а Δv — по полю), то

$$\frac{m}{2} \Delta \overline{v^2} = eE\bar{l}. \quad (5.102)$$

Таким образом, сформулированный выше критерий слабых полей будет иметь вид

$$eE\bar{l} \ll \frac{3}{2} kT. \quad (5.103)$$

При комнатной температуре $\frac{3}{2} kT \approx 0,04$ эв, полагая тепловую скорость электрона $\overline{v_0} \approx 10^7$ см/сек и время между соударениями $\bar{\tau} \approx 10^{-13}$ сек, получаем

$$l \approx \tau v \approx 10^{-6} \text{ см.} \quad (5.104)$$

В этом случае критическое поле приблизительно равно 10^4 в/см.

При понижении температуры средняя кинетическая энергия электронов падает пропорционально температуре, длина свободного пробега, если исключить рассеяние на ионах примеси (на котором мы остановимся отдельно), либо остается постоянной, либо растет с понижением температуры; поэтому согласно (5.103) критическое поле, как правило, довольно резко падает с понижением температу-

ры. Так, например, в достаточно чистом германии подвижность при гелиевых температурах начинает зависеть от поля при полях меньше 10 в/см .

Здесь следует подчеркнуть, что критические поля в неоднородных полупроводниках могут проявляться при очень малых разностях потенциала. Так, например, в кристалле, в котором имеется p - n переход, на него, как правило, падает все напряжение (так как сопротивление p - n перехода много больше сопротивления толщи полупроводника). Толщина p - n перехода d меняется в различных приборах от 10^{-3} до 10^{-7} см . При $d = 10^{-7} \text{ см}$ и разности потенциалов 1 в поле достигнет колоссальных размеров:

$$E = \frac{V}{l} = 10^7 \frac{\text{в}}{\text{см}}. \quad (5.105)$$

Следовательно, в тех случаях, когда эффекты, связанные с сильными полями, нежелательны, p - n переход следует делать достаточно толстым. Это, в частности, относится к силовым выпрямителям; одним из условий, необходимых для того, чтобы они выдерживали большое обратное напряжение (порядка 10^3 в/см), является достаточная толщина p - n перехода ($d \sim 10^{-4} \text{ см}$). При толстом p - n переходе прямая ветвь вольтамперной характеристики хуже, чем при тонком, однако на это приходится идти, чтобы обеспечить высокое обратное напряжение. Напротив, в приборах, основанных на явлениях в сильных полях (туннельных диодах и приборах, основанных на лавинном эффекте), p - n переход должен быть тонким *).

В поликристаллических образцах также очень часто сопротивление прослоек между зернами бывает во много раз больше сопротивлений самих зерен. Размеры зерен обычно колеблются в пределах от 10^{-2} до 1 мм , а толщина прослоек — в пределах 10^{-3} — 10^{-6} см . Следовательно, и в этом случае поле в прослойках может быть в 10 — 10^6 раз больше, чем поле, приложенное к образцу. На этом «усилении поля» основан целый ряд полупроводниковых приборов: действие нелинейных сопротивлений из карбида

*) Толщина p - n перехода также играет весьма существенную роль в его частотных характеристиках. Чем меньше она, тем больше емкость p - n перехода, что ухудшает его частотные характеристики; с другой стороны, при уменьшении толщины перехода уменьшается время пролета носителей, что весьма важно для работы на высоких частотах.

кремния и других материалов, фоточувствительных пленок из сернистого свинца и т. д.

После этого краткого отступления вернемся к влиянию поля на подвижность — проанализируем качественно, как это влияние зависит от механизма рассеяния носителей.

Рассеяние на тепловых колебаниях акустической ветви атомной решетки. Согласно формуле (5.53)

$$l \sim \varepsilon^r \quad (5.106)$$

и, следовательно,

$$u \sim \tau = \frac{l}{v} \sim \varepsilon^{r-\frac{1}{2}}. \quad (5.107)$$

В рассматриваемом нами случае $r=0$, следовательно,

$$u \sim \varepsilon^{-\frac{1}{2}} \sim v^{-1}. \quad (5.108)$$

Согласно (5.102) в сильном поле $v = \sqrt{E}$, следовательно,

$$u \sim E^{-\frac{1}{2}}; \quad (5.109)$$

к этому же выводу пришел Л. Д. Ландау, исходя из строгого анализа.

Рассеяние на ионах примеси ($r=2$)

$$l \sim \varepsilon^2 \sim v^4 \text{ и } u \sim v^3. \quad (5.110)$$

Следовательно, в этом случае подвижность должна расти с ростом поля, что действительно наблюдалось в германии при температурах порядка 20°K и поле, приблизительно равном 100 в/см . Однако при дальнейшем росте поля подвижность начинала падать. Это странное на первый взгляд явление имеет простое качественное объяснение.

Действительно, при температурах, отличных от абсолютного нуля, наряду с примесным рассеянием имеет место рассеяние на тепловых колебаниях решетки, однако с понижением температуры интенсивность тепловых колебаний падает, одновременно с этим падает и средняя тепловая скорость электронов, а следовательно, соответственно растет эффективность рассеяния на ионах примеси, и при низких температурах это рассеяние начинает играть преобладающую роль. Поэтому при наличии ионов примеси темпе-

температурная зависимость подвижности в широком интервале температур имеет вид, схематически представленный на рис. 5.5. В области рассеяния на тепловых колебаниях решетки подвижность падает с ростом температуры по закону $\mu \sim T^{-\frac{3}{2}}$, при низких температурах преобладает рассеяние на ионах примеси и подвижность растет пропорционально $\mu \sim T^{\frac{3}{2}}$. При этом излом в зависимости подвижности от температуры проявляется при тем более высокой темпе-

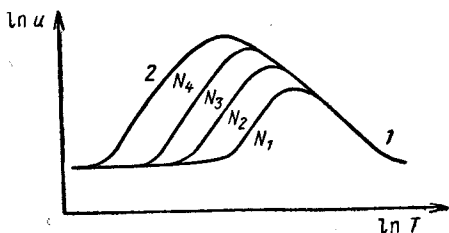


Рис. 5.5. Температурная зависимость подвижности для различного количества ионов примеси:

1 — область высоких температур (т. е. рассеяния на акустических тепловых колебаниях) $\mu \sim T^{-3/2}$; 2 — область низких температур рассеяния на ионах примеси.

ратуре, чем больше содержание ионов ($N_1 > N_2 > N_3 > N_4$). Наконец, при дальнейшем понижении температуры наступает вырождение: скорость электронов перестает зависеть от температуры и поэтому подвижность перестает зависеть от температуры и содержания примесей *).

Посмотрим теперь, что происходит в достаточно сильных полях, когда скорость электрона начинает увеличиваться с ростом поля. С этой точки зрения (т. е. в связи с ростом скорости электронов) увеличение поля эквивалентно росту температуры **). Поэтому эффективность рассеяния на ионах

*) Описанная выше (и представленная на рис. 5.5) картина, как уже упоминалось, схематична. В действительности по целому ряду причин наблюдаются отклонения от закона $l \sim e^2$, и все зависимости в связи с этим носят более сложный характер.

***) Когда $v \gg v_0$ и $\frac{1}{2}mv^2 \gg \frac{3}{2}kT$, т. е. нарушается тепловое равновесие между электронами и решеткой, вводят температуру электронного газа $T_э$, определяемую условием $\frac{1}{2}mv^2 = \frac{3}{2}kT_э$.

Из сказанного выше следует, что температура электронного газа всегда выше температуры решетки (или равна ей).

падает и подвижность растет до тех пор, пока не начинает сказываться рассеяние на тепловых колебаниях решетки, после этого она начинает падать по закону $\mu \sim E^{-\frac{1}{2}}$. Поэтому зависимость подвижности от поля должна иметь вид, аналогичный температурным зависимостям подвижности (рис. 5.5). Эти закономерности, как мы уже упоминали, качественно подтвердились на германии при низких температурах.

Рассеяние на оптических колебаниях ионной решетки. При температурах ниже температуры Дебая

$$r = 1$$

и

$$\mu \sim \frac{l}{v} \sim v \sim E^{\frac{1}{2}}, \quad (5.111)$$

т. е. подвижность должна расти с увеличением поля; при температуре выше температуры Дебая $r = 1/2$ и подвижность не должна, если исходить из этих качественных соображений, зависеть от поля:

$$\mu \sim \frac{l}{v} \sim \frac{\frac{1}{\varepsilon^2}}{\frac{1}{\varepsilon^2}} = \text{const.} \quad (5.112)$$

Однако более строгий анализ, проведенный Б. И. Давыдовым [9], показал, что в этом случае в сильных полях (т. е. при росте v) нарушаются предпосылки, необходимые для правомерности выражения для подвижности (5.3а), и все выводы надо пересмотреть.

В заключение этого раздела уточним определение сильного поля. В действительности полевые приращения скорости Δv , входящие в неравенство (5.100) и равенство (5.102), разные. Дело в том, что большинство столкновений, которые испытывает электрон, почти упругие, а следовательно, при них скорость почти не меняется по абсолютной величине и требуется большое число таких столкновений, чтобы электрон пришел в тепловое равновесие с решеткой; поэтому добавка Δv к скорости электрона в стационарном состоянии, входящая в неравенство (5.100), много больше добавки, приобретаемой на *одной* длине свободного

пробега, определяемой соотношением (5.102), эту добавку мы обозначим Δv_1 .

Оценим отношение $\Delta v/\Delta v_1$ на примере столкновений электронов с ионами. В этом случае при одном столкновении электрон теряет следующую часть своей энергии [см. (5.22)]:

$$\Delta \bar{\varepsilon} \approx \frac{m}{M} \bar{\varepsilon}. \quad (5.113)$$

Следовательно, энергия электрона в электрическом поле будет возрастать до тех пор, пока потеря энергии при одном столкновении не будет равна энергии, приобретаемой за время одного свободного пробега. Таким образом, условие стационарности будет иметь вид

$$eE\bar{l} - \frac{m}{M} \bar{\varepsilon} = 0,$$

откуда

$$\bar{\varepsilon} = eEl \frac{M}{m} = eEl\alpha, \quad (5.114)$$

где

$$\alpha = M/m$$

и

$$\Delta v = \Delta v_1 \sqrt{\frac{M}{m}} = \Delta v_1 \sqrt{\alpha}.$$

Полагая $M = 10^4 m$, получаем

$$\Delta v = 100 \Delta v_1. \quad (5.115)$$

Согласно (5.115) явления в сильных полях могут начинаться тогда, когда приращение энергии электрона на длине свободного пробега в 10^4 раз меньше его средней тепловой энергии. Аналогичные оценки параметра α для рассеяния на тепловых колебаниях дают $\alpha = 100$; эта цифра ближе к истине, так как в сделанном выше расчете мы не учли тепловых колебаний, столкновения с которыми являются менее упругими и роль которых будет расти по мере роста энергии электрона.

ВЛИЯНИЕ ПОЛЯ НА КОНЦЕНТРАЦИЮ НОСИТЕЛЕЙ

Существует несколько механизмов влияния поля на концентрацию носителей:

— электростатическая ионизация — за счет туннельного эффекта,

— термоэлектронная ионизация,

— ударная ионизация.

Рассмотрим каждое из этих явлений в отдельности.

Электростатическая ионизация. На рис. 5.6 представлена зонная структура беспримесного полупроводника при наличии сильного электрического поля; как видно

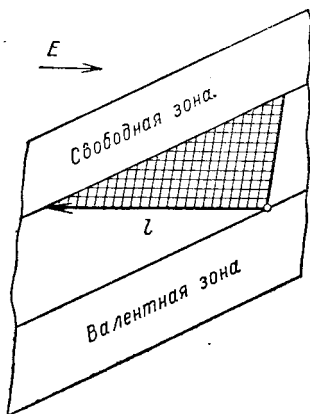


Рис. 5.6. Схема зонной диаграммы беспримесного полупроводника при наличии сильного поля, иллюстрирующая электростатическую ионизацию.

Заштрихована площадь барьера.

из рисунка, в этом случае возможен переход электрона из валентной зоны в свободную. Оценим вероятность такого перехода. Волновая функция зонного электрона $\psi = A e^{-2\pi i k x}$ может быть преобразована к виду, более удобному для оценки этой вероятности; для этого выразим волновое число k через энергию электрона:

$$k = \frac{p}{\hbar} = \frac{1}{\hbar} \sqrt{2Tm} = \frac{1}{\hbar} \sqrt{2m(\mathcal{E} - U)}, \quad (5.116)$$

и тогда

$$\psi = A \exp \left[\frac{2\pi i}{\hbar} \sqrt{2m(\mathcal{E} - U)} x \right], \quad (5.117)$$

где T — кинетическая; U — потенциальная и \mathcal{E} — полная энергия электрона.

Формула (5.117) показывает, что при попадании электрона в область отрицательной полной энергии, т. е. в область потенциального барьера, волновая функция начинает

затухать по закону

$$\psi(x) = A e^{-\alpha x},$$

где

$$\alpha = \frac{2\pi}{h} \sqrt{2\Delta\mathcal{E}m}.$$

Здесь $\Delta\mathcal{E} = (U - \mathcal{E})$ — высота барьера; вероятность просачивания через барьер определяется соотношением

$$\omega \approx e^{-\int_0^l \alpha^2 dx}. \quad (5.118)$$

В рассматриваемом случае $\Delta\mathcal{E}_0$ — ширина запрещенной зоны, барьер имеет форму треугольника, основание которого l , как видно из рис. 5.6, определяется условием $eEl = \Delta\mathcal{E}_0$, откуда

$$\omega \approx \exp\left(-\frac{\pi^2 ma}{h^2} \frac{\Delta\mathcal{E}_0^2}{lE}\right). \quad (5.119)$$

Оценка по формуле (5.119) показывает, что при $\Delta\mathcal{E}_0 \approx 1$ эВ вероятность туннельного эффекта становится заметной при полях порядка 10^7 в/см.

При электростатической ионизации доноров вероятность туннельного эффекта значительно возрастает, так как при этом и высота, и ширина барьера становятся меньше; в данном случае этот эффект может быть замечен при полях порядка 10^6 в/см.

Еще более вероятен туннельный эффект на контакте полупроводника и металла, если при этом ширина барьера не увеличивается областью объемного заряда. На этом эффекте основано действие туннельных диодов (см. гл. 8).

Особенностью электростатической ионизации является тот факт, что ее вероятность не зависит от температуры; по этому признаку такой механизм ионизации можно отличить от двух других, к рассмотрению которых мы сейчас перейдем.

Термоэлектронная ионизация. На рис. 5.7 представлен энергетический спектр донорного примесного атома при наличии сильного электрического поля. Как видно из рисунка, в этом случае действие поля проявляется в двух аспектах: 1) появляется возможность безактивационного просачивания в зону проводимости путем туннельного эффекта (этот механизм мы рассмотрели выше) и 2) умень-

шается на величину ΔW энергия активации, необходимая для теплового заброса электрона в зону проводимости. Так как величина ΔW (как видно из рисунка) растет с ростом поля, то соответственно будет расти число электронов в свободной зоне и электропроводность. Этот механизм иони-

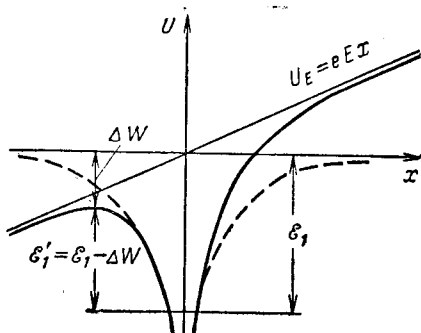


Рис. 5.7. Схема термоэлектронной ионизации.

зации называется термоэлектронным. Нетрудно оценить этот эффект количественно.

В отсутствие поля число электронов в зоне проводимости

$$n_0 = A e^{-\Delta \mathcal{E}_1 / 2kT}, \quad (5.120)$$

где A — коэффициент, слабо зависящий от температуры и $\Delta \mathcal{E}_1$ — энергия ионизации (активации) примесных атомов (см. рис. 5.7).

При наличии поля величина \mathcal{E}_1 снижается на ΔW и

$$\mathcal{E}'_1 = \mathcal{E}_1 - \Delta W, \quad (5.120a)$$

$$n = A e^{-\frac{\Delta \mathcal{E}'_1}{2kT}} = n_0 e^{\frac{\Delta W}{2kT}}. \quad (5.121)$$

Величина ΔW численно равна потенциальной энергии электрона в той точке x_0 , где энергия достигает максимума (см. рис. 5.7).

Согласно рис. 5.7 при наличии поля ход потенциальной энергии вблизи примесного атома в направлении поля (x) будет выражаться соотношением

$$U = -\frac{e^2}{\kappa x} - eEx. \quad (5.122)$$

Высота потенциального барьера будет максимальна в точке x_0 , определяемой уравнением

$$\left[\frac{\partial U}{\partial x} \right]_{x=x_0} = \frac{e^2}{x_0^2 \kappa} - eE = 0, \quad (5.123)$$

откуда

$$x_0 = \sqrt{\frac{e}{\kappa E}}. \quad (5.124)$$

Высота барьера в точке $x = x_0$

$$U(x)_{x=x_0} = \Delta W = -2e \sqrt{\frac{eE}{\kappa}}, \quad (5.125)$$

и, следовательно,

$$n = n_0 \exp \left(- \frac{e \sqrt{\frac{eE}{\kappa}}}{kT} \right). \quad (5.126)$$

Оценка n согласно формуле (5.126) показывает, что рост концентрации носителей за счет термоионной ионизации должен наблюдаться при относительно слабых полях, порядка 10^2 — 10^3 в/см, что в действительности не имеет места. Для того чтобы согласовать теорию Я. И. Френкеля (изложенную выше) с опытом, А. И. Губанов внес в нее два уточнения.

1. Если учесть экранировку заряда примесного атома свободными электронами, то ход потенциала вблизи него будет выражаться формулой

$$U = \frac{e^2}{\kappa x} e^{-\frac{x}{r_0}}, \quad (5.127)$$

где r_0 — радиус экранировки Дебая, равный

$$r_0 = \sqrt{\frac{\kappa kT}{8\pi n e^2}}. \quad (5.128)$$

Согласно (5.127) U быстрее стремится к нулю по мере роста r , чем без учета экранировки, и, следовательно, при наличии поля «горб» потенциала будет ближе к ядру и выше; поэтому эффект поля будет меньше.

2. Согласно развитой теории электрон каждый раз покидает примесный атом, когда его энергия достигает значения (5.120а).

В действительности это неточно, так как только в одной точке орбиты, x_0 , энергии электрона достаточно, чтобы

оторваться от примесного атома (во всех остальных точках орбиты высота потенциального барьера выше).

Для того чтобы электрон достиг этой точки, не растратив по пути свою энергию, приобретенную в поле, необходимо, чтобы длина орбиты была меньше длины свободного пробега электрона $2\pi x_0 < l$ или, более грубо,

$$l > x_0. \quad (5.129)$$

Учет этих двух поправок позволяет лучше согласовать изложенную выше теорию с опытом.

Согласно теории Я. И. Френкеля эффект поля должен экспоненциально расти с температурой.

Ударная ионизация. На рис. 5.8 схематически представлено движение электрона в зоне проводимости при наличии

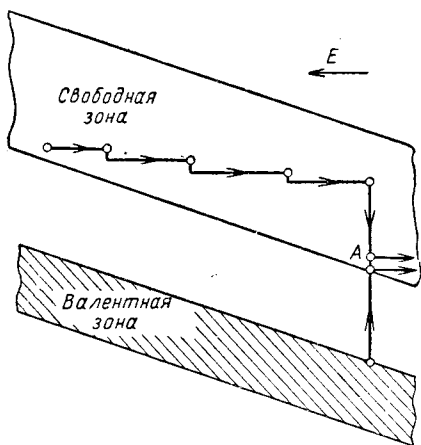


Рис. 5.8. Схема ударной ионизации.

сильного поля. Как видно из рис. 5.8, в точке А его кинетическая энергия становится достаточной для того, чтобы перебросить электрон из валентной зоны в свободную, после чего весь процесс может начаться сначала. Поля, при которых начинается ударная ионизация, можно грубо оценить согласно формуле (5.114)

$$aeEl = \Delta\mathcal{E}_0, \quad (5.130)$$

где $\Delta\mathcal{E}_0$ — ширина запрещенной зоны; недостатком формулы (5.130) является то, что в нее входит длина свободного пробега электрона, которая, как мы видели, сама зависит от поля.

Однако соотношение (5.130) правильно указывает, что вероятность ударной ионизации должна зависеть от l , т. е. от чистоты материала и от температуры, что действительно наблюдается на опыте. Теоретические оценки и опыт показывают, что ударная ионизация начинает играть существенную роль при полях 10^4 — 10^6 в/см.

Впервые предположение об ударной ионизации было высказано А. Ф. Иоффе для объяснения механизма пробоя твердых диэлектриков. Теория ударной ионизации была развита Фрейлихом, Гиппелем, Келдышем и др.

Существует два механизма ударной ионизации. Первый из них — *низковольтный*, т. е. такой, при котором электрон на одной длине свободного пробега приобретает лишь незначительную энергию, но все же превышающую tu , которую он теряет при соударении.

При этом накопление энергии электроном происходит постепенно и, так как при каждом или в результате небольшого числа столкновений электрон теряет направленную по полю составляющую скорости (v_x), она (т. е. направленная составляющая скорости) остается малой ($v_x \ll v$), а растет скорость хаотического движения, т. е. тепловая скорость. Таким образом, при этом нарушается тепловое равновесие между решеткой и электронным газом; последний приобретает свою собственную температуру, которая может значительно превосходить температуру решетки.

Так как в этом случае рост энергии электрона происходит постепенно, то такой пробой должен развиваться более эффективно в толстом слое (в противоположность туннельному эффекту). Эту закономерность подтвердил опыт: пробивное напряжение p - n перехода падает с ростом толщины запирающего слоя.

Разумеется, эта закономерность наблюдается до определенного предела: если толщина слоя становится настолько большой, что все электроны, созданные полем, не успевают из него выйти, а рекомбинируют в нем, то критическое поле перестает зависеть от L .

Второй механизм ударной ионизации, в отличие от первого, можно условно назвать *высоковольтным*; в этом случае электрон уже на одной длине свободного пробега приобретает скорость, достаточную для вырывания другого электрона.

В этом случае становится велика направленная составляющая скорости v_x , а хаотическая скорость может оставаться близкой к нормальной, если напряженность поля такова, что он (электрон) при неупругом столкновении всю приобретенную скорость теряет.

Интересно отметить, что, хотя поле в этом случае больше, чем в первом, температура электронов может быть значительно ниже, так как начинает более эффективно действовать новый (неупругий) механизм столкновений.

В данном случае может наблюдаться периодический рост температуры электронов с ростом поля (так как она определяется разностью между энергией, приобретенной на длине свободного пробега, и потерей энергии при одном акте ионизации).

Отличие этого типа ионизации от рассмотренного выше заключается в том, что материал при этом сохраняет свой тип проводимости: в полупроводнике n -типа электрон ионизирует донорные уровни и образец остается электронным, в полупроводнике p -типа дырки забрасывают электроны из валентной зоны на примесные уровни и сохраняется дырочная проводимость.

В p - n переходе ионизация производится, как правило, и дырками и электронами и при каждом акте создается пара электрон — дырка.

ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ И ТЕПЛОПРОВОДНОСТЬ

В этой главе мы рассмотрим группу явлений переноса, возникающих при наличии градиента температур, и явление Пельтье, вследствие которого создается разность температур. Как следует из названия главы, к ним относятся термоэлектрические явления и теплопроводность. Мы уже говорили коротко об этих явлениях в гл. I, но при этом, стремясь к максимальной простоте, допустили некоторые неточности и теперь должны внести соответствующие поправки.

Первым делом следует упомянуть о принятом нами в гл. I делении термо-э. д. с. на две составляющие — объемную и контактную. Если следовать той же классификации, то явление Пельтье следует считать целиком контактным явлением. При этом остается неясным, каким образом может выполняться соотношение Пельтье

$$P = \alpha T, \quad (6.1)$$

если левая часть его целиком зависит от условий на контакте, а правая — лишь наполовину.

В действительности и термо-э. д. с. и эффект Пельтье являются одновременно и объемными и контактными явлениями, но в другом смысле. Объемными потому, что оба они возникают в результате нарушения равновесия в потоке электронов в объеме полупроводника, т. е. потому, что средняя энергия электронов, участвующих в переносе электричества или тепла при наличии градиента потенциала или температуры, не равна энергии электронов на уровне Ферми. Контактными оба эффекта являются в том смысле, что оба они проявляются лишь в том случае, когда отступ-

ление от равновесия в потоке неодинаково в обеих ветвях термоэлемента.

В гл. I было показано, что явление Пельтье возникает вследствие того, что среднее значение энергии $\bar{\epsilon}$ электронов, участвующих в переносе тока, по разные стороны от контакта различно, так как различно среднее отступление от уровня электрохимического потенциала ($\epsilon - \mu$).

Аналогично обстоит дело и при наличии градиента температур. В этом случае по обе стороны от любого сечения полупроводника различна величина $(\epsilon - \mu)/kT$, определяющая заселенность электронами уровня с энергией ϵ . Однако функция Ферми сама по себе симметрична по отношению к уровню химического потенциала, и если бы подвижность электронов с различной энергией была также одинакова, то потоки электронов справа налево и слева направо были бы одинаковы и явление термо-э. д. с. вообще не возникло бы. В действительности электроны, находящиеся на примесных уровнях, практически совершенно неподвижны, а электроны, находящиеся на различных энергетических уровнях зоны проводимости, обладают различной подвижностью, поэтому возникает направленное движение электронов, которое будет продолжаться до тех пор, пока возникшая как его следствие разность потенциалов не создаст встречный поток электронов, компенсирующий это движение. Эта разность и называется термоэлектродвижущей силой E_T .

Выше мы рассмотрели разомкнутую цепь. Если цепь замкнута, то в ней будет протекать термоэлектрический ток, определяемый законом Ома $I = E_T/R$, где R — полное сопротивление цепи. После этих замечаний можно привести вывод, в котором термо-э. д. с. разделяет цепь на две части — объемную и контактную. В соответствии со сказанным выше он не точен, но тем не менее наглядно показывает, что в некоторых случаях термо-э. д. с. может менять свой знак (т. е. в электронном полупроводнике быть положительной, а в дырочном — отрицательной) *).

В § 1.6 мы привели основные сведения о термоэлектрических явлениях: их определения, термодинамические соотношения Томсона, качественную картину их возникно-

*) В соответствии со сказанным выше это происходит тогда, когда меняет свой знак $(\epsilon - \mu)$ в потоке.

вения и, наконец, нестрогий, но дающий правильный результат вывод коэффициента Пельтье — и затем, воспользовавшись соотношением Томсона, получили выражения для коэффициента термо-э. д. с. Такой путь вывода термоэлектрических коэффициентов называется π -подходом. Теперь покажем, что те же выражения можно получить, начиная непосредственно с вывода коэффициента термо-э. д. с.

6.1. ТЕРМОЭЛЕКТРОДВИЖУЩАЯ СИЛА

В гл. I мы предположили, что температура во всей цепи постоянна и существует электрический ток, и вычислили теплоту Пельтье, выделяющуюся (или поглощающуюся в зависимости от направления тока) на контактах. Теперь мы предположим, наоборот, что цепь разомкнута и ток в ней отсутствует, но существует градиент температуры. Вычислим возникающую в ней вследствие этого разность потенциалов (термоэлектродвижущую силу):

$$V = \alpha \Delta T. \quad (6.2)$$

Предположим, что эта разность потенциалов будет состоять из двух частей: контактной, возникающей из-за температурной зависимости контактной разности потенциалов, и, объемной, возникающей за счет диффузии носителей от горячего конца к холодному:

$$V = \left(\frac{\partial U_K}{\partial T} + \frac{\partial U_{об}}{\partial T} \right) \nabla T. \quad (6.3)$$

Следовательно,

$$\alpha = \frac{\partial U_K}{\partial T} + \frac{\partial U_{об}}{\partial T} = \alpha_K + \alpha_{об}. \quad (6.4)$$

Предположим также, что одну ветвь термоэлемента составляет металл, а вторую — электронный полупроводник, и вычислим термо-э. д. с., создаваемую полупроводниковой ветвью. Тогда 1) при вычислении $U_{об}$ мы должны учитывать только разность потенциалов внутри полупроводника и 2) в таком же приближении можем считать, что температурная зависимость контактной разности потенциалов целиком обусловлена температурной зависимостью положения уровня химического потенциала в полупроводнике:

$$\alpha_K = -\frac{1}{e} \frac{\partial \mu}{\partial T},$$

где согласно (4.35а)

$$\mu = kT \ln \frac{h^3 n}{2 (2\pi m_n kT)^{3/2}} .$$

Следовательно,

$$\alpha_K = \frac{k}{e} \left(\frac{3}{2} - \mu^* - T \frac{\partial \ln n}{\partial T} \right) , \quad (6.5)$$

где $\mu^* = \mu/kT$.

Для вычисления коэффициента объемной термо-э. д. с. $\alpha_{об}$ мы должны вычислять электрический ток при наличии градиента температуры и потенциала, приравнять этот ток к нулю и из этого условия найти градиент потенциала. Мы, однако, отвергаем пока этот строгий путь и воспользуемся более наглядными представлениями, которые сразу приведут к цели.

Электронный газ, как и всякий другой, создает давление $p = nkT$. При наличии градиента температуры существует, следовательно, и перепад давления, и для равновесия необходимо, чтобы электрическое поле действовало на электроны с силой, которая уравновесила бы этот перепад давления:

$$U_{об} en = \frac{\partial p}{\partial x} = \frac{\partial p}{\partial T} \frac{\partial T}{\partial x} ;$$

следовательно, объемный коэффициент термо-э. д. с. выразится формулой

$$\alpha_{об} = \frac{1}{en} \frac{\partial p}{\partial T} = \frac{k}{e} \left(1 + T \frac{\partial \ln n}{\partial T} \right) . \quad (6.6)$$

Более строгий вывод, учитывающий то, что в стационарном состоянии должны уравновешиваться не давления, а диффузионный и электрический потоки электронов, дает несколько отличное выражение:

$$\alpha_{об} = \frac{k}{e} \left(r + \frac{1}{2} - T \frac{\partial \ln n}{\partial T} \right) , \quad (6.7)$$

где r — показатель степени в выражении для зависимости длины свободного пробега электрона от его энергии.

Складывая (6.7) и (6.5), получаем

$$\alpha_{об} = \frac{k}{e} (r + 2 - \mu^*) . \quad (6.8)$$

Из (6.7) следует, что в случае $n = \text{const}$ и $r < -1/2$ объемное поле в полупроводнике изменит знак, т. е. будет «толкать» носители не от холодного спая к горячему, а в обратном направлении.

Этому парадоксальному на первый взгляд результату не следует удивляться. Действительно, $r < 0$ означает, что электроны с большей энергией обладают меньшей длиной свободного пробега и меньшей диффузионной способностью, чем медленные. Следовательно, при наличии градиента температуры электроны с холодного конца проводника (обладающие меньшими скоростями) будут быстрее диффундировать по направлению к горячему концу, чем в обратном направлении, и в объеме проводника создастся поле в таком направлении, чтобы скомпенсировать разность скоростей диффузии *).

6.2. ВЫВОД КОЭФФИЦИЕНТА ТЕРМО-Э. Д. С. ИЗ КИНЕТИЧЕСКОГО УРАВНЕНИЯ

Строгий вывод термо-э. д. с. следует из ее определения — это разность потенциалов, возникающая в цепи при наличии разности температур и при отсутствии тока (т. е. в разомкнутой цепи).

Таким образом, для того чтобы быть последовательными, мы должны подставить выражение для неравновесной функции распределения (5.67) при наличии градиента температуры в выражение для плотности тока (5.68), приравняв его (т. е. ток) к нулю и из этого условия найти электрическое поле E . В результате получаем

$$j = \int \tau \left[\frac{\mu - \epsilon}{T} \nabla T - \nabla (\mu - e\varphi) \right] \frac{\partial f_0}{\partial \epsilon} v_x^2 g(\epsilon) d\epsilon = 0; \quad (6.9)$$

учитывая, что при усреднении по шаровому слою фазового пространства $v_x^2 = 1/3 v^2 = 2\epsilon^2/3m$, получаем

$$\nabla \left(\frac{\mu}{e} - \varphi \right) = \frac{k}{e} \left(\frac{1}{T} \frac{\int \epsilon^2 \frac{\partial f_0}{\partial \epsilon} \tau g(\epsilon) d\epsilon}{\int \epsilon \frac{\partial f_0}{\partial \epsilon} \tau g(\epsilon) d\epsilon} - \mu^* \right) \nabla T. \quad (6.10)$$

На рис. 6.1 представлена цепь, состоящая из полупроводника и из металла (или из двух полупроводников).

*) Такой механизм рассеяния наблюдается, в частности, в ряде переходных металлов и их сплавов.

Чтобы получить термо-э. д. с. в этой цепи, мы должны вычислить разность потенциалов между точками 1 и 4:

$$V = - \int_1^4 \nabla \varphi dx; \quad (6.11)$$

вычисление этого интеграла представляет некоторые трудности, так как в точках 2 и 3 потенциал меняется скачком из-за контактной разности потенциалов V_K , причем скачок

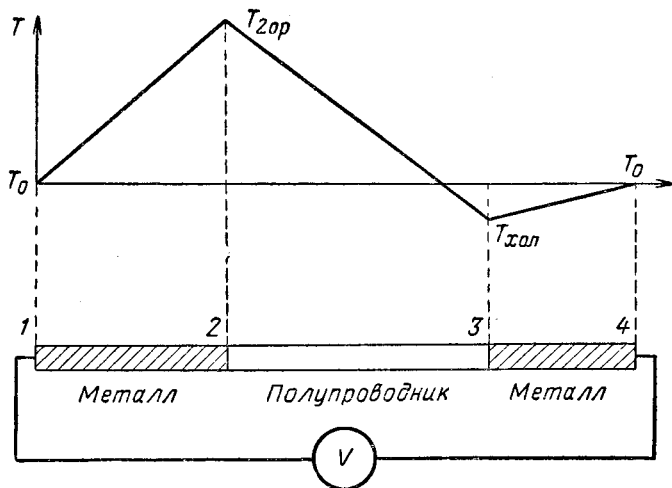


Рис. 6.1. Схема распределения температур в термоэлементе.

в этих точках неодинаков, так как U_K зависит от температуры. Чтобы обойти эту трудность, можно заменить интеграл (6.11) следующим интегралом:

$$V' = \frac{1}{e} \int_1^4 \nabla (\mu - e\varphi) dx. \quad (6.12)$$

Нетрудно убедиться, что (6.11) и (6.12) эквивалентны. Действительно,

$$V' = V - \frac{1}{e} (\mu_4 - \mu_1), \quad (6.13)$$

Однако $\mu_4 = \mu_1$, так как в точках 1 и 4 находится один и тот же материал при одной и той же (комнатной) темпера-

туре; следовательно, $V'' = V$. С другой стороны, величина $\nabla(\mu - e\varphi)$ не претерпевает скачок на контакте, так как условием контактного равновесия является равенство электрохимических потенциалов. С учетом всего сказанного выше и равенства (6.11) получаем

$$V = \int_1^4 \nabla \left(\frac{\mu}{e} - \varphi \right) dx = \int_1^2 + \int_2^3 + \int_3^4. \quad (6.14)$$

Заменяя согласно (6.10) интегрирование по координатам интегрированием по температуре и объединив вместе интегралы \int_1^2 и \int_3^4 , получаем окончательно

$$V = \int_{T_x}^{T_r} (\alpha_{\text{п}} - \alpha_{\text{м}}) dT, \quad (6.15)$$

где

$$\alpha_{\text{п, м}} = \frac{k}{e} \left[\frac{1}{T} \frac{\int \varepsilon^2 \frac{\partial f_0}{\partial \varepsilon} \tau g(\varepsilon) d(\varepsilon)}{\int \varepsilon \frac{\partial f_0}{\partial \varepsilon} \tau g(\varepsilon) d\varepsilon} - \mu^* \right]_{\text{п, м}}. \quad (6.16)$$

Выражение (6.16) показывает, что $\alpha_{\text{п}}$ и $\alpha_{\text{м}}$, по определению, являются относительными коэффициентами термо-э. д. с. ветвей термоэлемента.

Наиболее общее выражение для термо-э. д. с. металлов (т. е. сильно вырожденного электронного газа) получено Моттом:

$$\alpha = \frac{\pi^2}{3} \frac{k^2 T}{e} \left(\frac{\partial \ln \sigma(\varepsilon)}{\partial \varepsilon} \right)_{\varepsilon=\mu}, \quad (6.17)$$

где k — постоянная Больцмана; e и ε — соответственно заряд и энергия электрона и μ — химический потенциал (энергия электрона на уровне Ферми).

В случае простой параболической зонной структуры и степенной зависимости длины свободного пробега (l) электрона от его энергии (т. е. $l \sim \varepsilon^r$) выражение (6.17) упрощается и принимает вид

$$\alpha = \frac{\pi^2}{3} \frac{k}{e} (r+1) \frac{kT}{\mu}. \quad (6.18)$$

Выражение для термо-э. д. с. при тех же предположениях (один сорт носителей, параболическая зонная структура, степенная зависимость длины свободного пробега от энергии), но при произвольной степени вырождения имеет вид

$$\alpha = \frac{k}{e} \left[\frac{r+2}{r+1} \frac{F_{r+1}(\mu^*)}{F_r(\mu^*)} - \mu^* \right], \quad (6.19)$$

где $F_r(\mu^*)$, $F_{r+1}(\mu^*)$ — интегралы Ферми:

$$F_r(\mu^*) = \int_0^{\infty} \frac{x^r dx}{e^{x-\mu^*} + 1}. \quad (6.20)$$

В случае невырожденного электронного газа ($\mu^* < -2$) выражение (6.19) упрощается:

$$\alpha = \frac{k}{e} \left[r + 2 + \ln \frac{2(2\pi m k T)^{3/2}}{h^3 n} \right], \quad (6.21)$$

где m — эффективная масса плотности состояний; $m = N^{\frac{2}{3}} (m_1, m_2, m_3)^{\frac{1}{3}}$ (N — число эллипсоидов и m_1, m_2, m_3 — главные значения тензора эффективной массы)*.

В случае нескольких сортов носителей (электронов с различной эффективной массой или электронов и дырок) выражение для термо-э. д. с. принимает следующий вид:

$$\alpha = \frac{\sum \alpha_i \sigma_i}{\sum \sigma_i}, \quad (6.22)$$

где α_i — «парциальная» термо-э. д. с. для i -го сорта носителей, вычисленная согласно (6.19) или (6.21) с учетом знака носителей, и σ_i — соответствующая электропроводность.

6.3. УВЛЕЧЕНИЕ ЭЛЕКТРОНОВ ФОНОНАМИ

Относительно недавно был открыт еще один источник термо-э. д. с. — так называемый эффект увлечения электронов фононами. При низких температурах эта составляющая термо-э. д. с. может быть в десятки и сотни раз больше рассмотренных выше. Качественно это явление можно объяснить следующим образом. Если в твердом теле существует градиент температуры, то число фононов, движущихся от

*) Формула (6.21) была выведена впервые советским физиком Н. Л. Писаренко (1940 г.).

горячего конца к холодному, будет больше, чем в обратном направлении. При своем направленном движении фононы в результате столкновений с электронами будут увлекать за собой последние, и на холодном конце образца будет накапливаться отрицательный заряд (а на горячем — положительный) до тех пор, пока возникшая разность потенциалов не уравнивает эффект увлечения. Эта разность потенциалов и представляет собой дополнительную термо-э. д. с. Наоборот, если в теле существует электрический (а не тепловой) ток, то электроны увлекают за собой фононы, чем вызывается дополнительный перенос тепла (эффект Пельтье). Впервые возможность существования такого эффекта в металлах в 1945 г. предсказал Л. Э. Гуревич. Он же указал на возможность существования такого эффекта в полупроводниках. В 1951 г. Г. Е. Пикус рассмотрел это явление в полупроводниках количественно и вывел формулу для дополнительной (возникающей вследствие увлечения электронов фононами) термо-э. д. с.:

$$\alpha_{\phi} = \frac{1}{3} \frac{k}{e} \frac{m\omega_{\phi}^2}{k} \frac{1}{T} \frac{\tau_{\phi}}{\tau_{\text{э}}}, \quad (6.23)$$

где ω_{ϕ} — скорость звука; τ_{ϕ} и $\tau_{\text{э}}$ — время релаксации фононов и электронов соответственно.

Величина $m\omega_{\phi}^2/k$ имеет размерность температуры, а именно, это та температура, при которой тепловая скорость электронов равна скорости фононов. Нетрудно убедиться в том, что эта величина приблизительно равна 1°K , $k/e = 90 \text{ мкв/град}$, следовательно,

$$\alpha_{\phi} \approx \frac{30}{T} \frac{\bar{\tau}_{\phi}}{\tau_{\text{э}}}. \quad (6.24)$$

Значение $\bar{\tau}_{\phi}$ можно получить из выражения для теплопроводности кристаллической решетки, если воспользоваться формулой Дебая для $\kappa_{\text{р}}$ (1.64)

$$\kappa_{\text{р}} = \frac{1}{3} C \bar{l}_{\phi} \omega_{\phi} = \frac{1}{3} C \omega_{\phi}^2 \bar{\tau}_{\phi}. \quad (6.25)$$

Значение $\bar{\tau}_{\text{э}}$ можно оценить, зная подвижность электронов $u = (e/m)\tau_{\text{э}}$.

Проведя этот анализ и убедившись в том, что $\bar{\tau}_{\phi}$ обычно не больше чем на порядок отличается от $\bar{\tau}_{\text{э}}$, Пикус пришел к выводу, что α_{ϕ} не может играть существенной роли.

Однако в 1953 г. Фредерикс и Гебалле, измеряя термо-электродвижущую силу германия, обнаружили, что при низких температурах она возрастает до значений, во много раз превосходящих предсказанные обычной теорией. В 1955 г. Хуллом и Гебалле было найдено, что в p -кремнии термо-э. д. с. достигает еще больших значений, порядка 50 мв/град . Теоретическое объяснение этих аномально больших значений было дано в 1958 г. Херрингом и Фредериксом на основе эффекта увлечения электронов фононами и формулы (6.23), впервые выведенной Пикусом.

Дело в том, что в (6.23) должно входить среднее время релаксации не для всех фононов, а для тех, которые могут взаимодействовать с электронами. Как было показано в гл. 5, из законов сохранения энергии и импульса при столкновении электронов с фононами следует, что волновой вектор фонона q должен удовлетворять неравенству

$$q < 2k,$$

где k — волновой вектор электрона.

Следовательно, длина волны таких фононов $\lambda \approx 1/q$ должна быть одного порядка или больше, чем длина волны электронов. Последняя при комнатной температуре достигает порядка 10^{-7} см , а при низких — соответственно возрастает. Отсюда следует, что в (6.23) должно входить время релаксации для фононов, длина волны которых во много раз превышает постоянную решетки. Херринг показал, что время релаксации для таких фононов $\tau_{\text{ф}}$ на несколько порядков больше, чем среднее время релаксации $\bar{\tau}_{\text{ф}}$, получаемое из анализа формулы (6.25) для теплопроводности, и, следовательно, будет значительно больше, чем ожидал Пикус.

Покажем, как на основании простых наглядных соображений можно вывести формулу (6.23). Если температура всего тела постоянна, то фононы движутся во всех возможных направлениях; распределение скоростей фононов по направлениям изотропно. Если же существует градиент температуры, то, как мы уже упоминали, число фононов, движущихся от горячего конца к холодному, становится больше, чем в обратном; мы можем сказать, что на беспорядочное движение фононов накладывается дрейф в тепловом поле (аналогично тому, как при наличии градиента потенциала на хаотическое движение электронов накладывается дрейф в электрическом поле).

Через скорость этого дрейфа ω_d и количество энергии, заключенное в единице объема, $\epsilon = CT$ (где C — теплоемкость) можно выразить поток тепла, переносимый фононами:

$$Q = \epsilon \omega_d = CT \omega_d, \quad (6.26)$$

с другой стороны,

$$Q = \kappa \frac{\partial T}{\partial x} = \frac{1}{3} C \omega_\phi l_\phi \frac{dT}{dx}. \quad (6.27)$$

Сравнивая эти две формулы, получаем

$$\omega_d = \frac{1}{3} \frac{\omega_\phi l_\phi}{T} \frac{dT}{dx}, \quad (6.28)$$

Если электроны взаимодействуют только с фононами (а не рассеиваются на каких-либо дефектах решетки), то они тоже приобретут скорость ω_d и будут накапливаться на холодном конце проводника до тех пор, пока возникшее вследствие этого электрическое поле E_ϕ не создаст скорость uE_ϕ , равную и противоположно направленную ω_d . Следовательно, в равновесии $uE_\phi = \omega_\phi$, откуда

$$E_\phi = \frac{1}{3} \frac{\omega_\phi l_\phi}{Tu} \frac{dT}{dx} \quad (6.28a)$$

и

$$\alpha_\phi = \frac{1}{3} \frac{\omega_\phi l_\phi}{Tu}. \quad (6.29)$$

Подставив в (6.29) $l_\phi = \tau_\phi \omega_\phi$ и $u = (e/m) \tau_\phi$, получим

$$\alpha_\phi = \frac{1}{3} \frac{k}{e} \frac{m \omega_\phi^2}{kT} \frac{\tau_\phi}{\tau_\theta} \quad (6.30)$$

в полном согласии с (6.23).

Разумеется, все рассуждения, приведенные выше, относятся к длинноволновым фононам, и соответствующее значение τ_ϕ следует подставить в (6.30).

При рассмотрении термоэлектрических явлений, связанных с взаимодействием электронов с фононами, также возможны два подхода: α -подход (т. е. вывод, проведенный нами выше) и π -подход. Покажем, что формулу (6.30) можно получить и вторым путем. Для этого предположим, что градиент температуры отсутствует, но существует

градиент потенциала. В этом случае электроны и фононы меняются местами: электроны, двигаясь в направлении градиента потенциала, увлекают за собой взаимодействующие с ними длинноволновые фононы и таким образом создают поток тепла, пропорциональный электрическому току (дополнительный эффект Пельтье). Подсчитаем этот поток тепла и, поделив его на ток, узнаем добавочный коэффициент Пельтье:

$$Q = \sum_q \varepsilon_q \omega_\Phi(q) = \bar{\omega}_\Phi^2 \sum hq = \bar{\omega}_\Phi^2 P_q, \quad (6.31)$$

где ε_q — энергия фононов; $\varepsilon_q = h\nu = hq\omega_\Phi$, и $P_q = \sum hq$ — полный импульс фононов, взаимодействующих с электронами.

При вычислении $\bar{\omega}_\Phi$ мы должны провести усреднение по всевозможным направлениям, поэтому $\bar{\omega}_\Phi^2 = 1/3 \omega_\Phi^2$; нам остается вычислить P_q . Для этого рассмотрим более детально механизм взаимодействия электронов с длинноволновыми фононами и длинноволновых фононов с коротковолновыми.

Электроны, ускоряясь в электрическом поле, приобретают за единицу времени импульс $\Delta p_a = eEn$ и при столкновениях передают этот импульс длинноволновым фононам, которые, получив этот импульс от электронов, теряют его при столкновениях с коротковолновыми фононами *). Следовательно, в стационарных условиях

$$\frac{dP_q}{dt} = eEn - \frac{P_q}{\tau_\Phi} = 0, \quad (6.32)$$

откуда $P_q = \tau_\Phi eEn$ и, так как

$$E = \frac{j}{eun}, \quad (6.33)$$

то $P_q = (\tau_\Phi/u) j$. Согласно (6.31)

$$Q = \frac{1}{3} \frac{\bar{\omega}_\Phi^2 \tau_\Phi}{u} j, \quad (6.34)$$

откуда

$$\Pi_\Phi = \frac{Q}{j} = \frac{1}{3} \frac{\bar{\omega}_\Phi^2 \tau_\Phi}{u} = \frac{1}{3} \frac{\omega_\Phi l_\Phi}{u} \quad (6.35)$$

*) Коротковолновые фононы, в свою очередь, отдают этот импульс кристаллической решетке.

$$\alpha_{\Phi} = \frac{\Pi_{\Phi}}{T} = \frac{1}{3} \frac{\omega_{\Phi} l_{\Phi}}{uT} \quad (6.36)$$

в полном согласии с (6.23) и (6.29).

6.4. ЗАВИСИМОСТЬ ТЕРМО-Э. Д. С. ОТ ТЕМПЕРАТУРЫ И КОНЦЕНТРАЦИИ НОСИТЕЛЕЙ

Приведем несколько графиков, иллюстрирующих качественно концентрационную и температурную зависимости термо-э. д. с. На рис. 6.2 представлены качественно

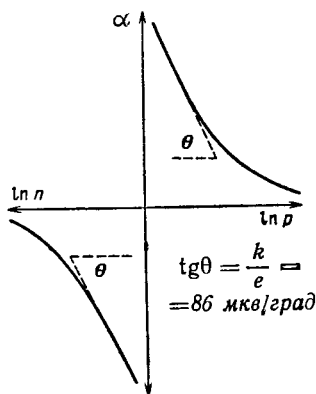


Рис. 6.2. Зависимость термо-э. д. с. от концентрации носителей.

в полулогарифмическом масштабе зависимости термо-э. д. с. от концентрации электронов (левая часть рисунка) или дырок (правая часть рисунка), соответствующие формуле (6.37):

$$\alpha_{p, n} = \pm \frac{k}{e} \left[\frac{r+2}{r+1} \frac{F_{r+1}(\mu^*)}{F_r(\mu^*)} - \mu^* \right], \quad (6.37)$$

где μ^* — приведенный химический потенциал — выражается через концентрацию носителей соотношением

$$n = \frac{4\pi(2mkT)^{3/2}}{h^3} F_{1/2}(\mu^*). \quad (6.38)$$

Рисунок построен для $r=0$ и $m=m_0$. При отсутствии вырождения ($\mu^* < -2$) оба выражения, (6.37) и (6.38),

упрощаются и принимают вид

$$\alpha_{p, n} = \pm \frac{k}{e} \left[r + 2 + \ln \frac{2(2\pi m k T)^{3/2}}{h^3 n} \right] \quad (6.39)$$

и

$$n = \frac{2(2\pi m k T)^{3/2}}{h^3} e^{\mu^*}. \quad (6.40)$$

Эта область ($\mu^* < -2$) соответствует прямолинейным участкам кривых на рис. 6.2 с постоянным наклоном:

$$\operatorname{tg} \theta = \frac{k}{e} = 86 \text{ мкв/град}. \quad (6.41)$$

Из (6.39) видно, что при $r \neq 0$ и $m \neq m_0$ линейная часть не меняет своего характера, но соответственно смещается вправо или влево. При наличии вырождения ($\mu^* > -2$), как видно из (6.37) и (6.38), при $m \neq m_0$ зависимость термо-э. д. с. от концентрации имеет тот же вид и лишь смещается вдоль оси абсцисс.

Зависимость термо-э. д. с. от механизма рассеяния (r) в этой области концентрации носит более сложный характер; однако частное от двух интегралов Ферми, входящее в (6.37), также в первом приближении линейно зависит от r . Это значит, что в этом приближении изменение r также смещает, но не деформирует зависимость термо-э. д. с. от концентрации носителей.

Все сказанное выше позволяет сделать вывод, что концентрационная зависимость термо-э. д. с. не позволяет судить об эффективной массе и механизме рассеяния в отдельности. Если же один из этих параметров определен в результате исследования каких-либо других явлений (например, гальвано- или термомагнитных), то из (6.37) или (6.39) можно определить второй параметр.

На рис. 6.3, а построена зависимость термо-э. д. с. от концентрации носителей для смешанной проводимости, когда надо учитывать наличие носителей обоих знаков, построенная согласно (6.22):

$$\alpha = \frac{\alpha_p \sigma_p - \alpha_n \sigma_n}{\sigma_p + \sigma_n}, \quad (6.42)$$

где α_p и α_n выражаются как (6.37) или (6.39). Точка, в которой эта кривая пересекает ось абсцисс, зависит от соотношения подвижностей и эффективных масс электронов и дырок и может также зависеть от механизмов рассеяния

электронов и дырок (r_n и r_p), если они неодинаковы, однако этот случай весьма редок.

Высота и острота «горбов», т. е. максимальных значений термо-э. д. с. в p - и n -области, определяется шириной запрещенной зоны: чем зона уже, тем меньше термо-э. д. с.

Следует, однако, учитывать, что неоднородности образцов (т. е. локальные отступления от средней концентра-



Рис. 6.3. Зависимость термо-э. д. с. от концентрации носителей в области смешанной проводимости для полупроводников с различной шириной запрещенной зоны $\Delta E_0' < \Delta E_0'' < \Delta E_0'''$ (а) и зависимость электропроводности от концентрации носителей (б).

ции доноров и акцепторов) также приводят к снижению термо-э. д. с. не только в области смешанной проводимости, но и при приближении к ней.

Из рис. 6.3, а видно, что зависимость концентрации носителей от термо-э. д. с. неоднозначна: каждому значению α соответствуют два значения n или p (n_1 и n_2 и p_1 и p_2).

На рис. 6.3, б приведена также зависимость электропроводности от концентрации носителей, знание которой помогает судить о том, в какой области концентраций (смешанной или примесной) мы находимся.

На рис. 6.4 представлена температурная зависимость термо-э. д. с. для ряда образцов полуметалла с различным содержанием доноров или акцепторов, при этом сделано предположение, что подвижность электронов больше, чем подвижность дырок *); при этих условиях электронные

*) При этом также сделано предположение, что эффективные массы и механизмы рассеяния не зависят от температуры. При этих условиях (см. рис. 6.4) $\text{tg } \theta = 3/2 (k/e)$.

образцы не меняют знака термо-э. д. с. в области собственной проводимости, а дырочные меняют *).

На рис. 6.5 представлены температурные зависимости термо-э. д. с. для ряда образцов полупроводника с различным содержанием доноров и акцепторов; в этом случае (т. е. при энергии активации примесей, отличной от нуля) термо-э. д. с. падает при низких температурах за

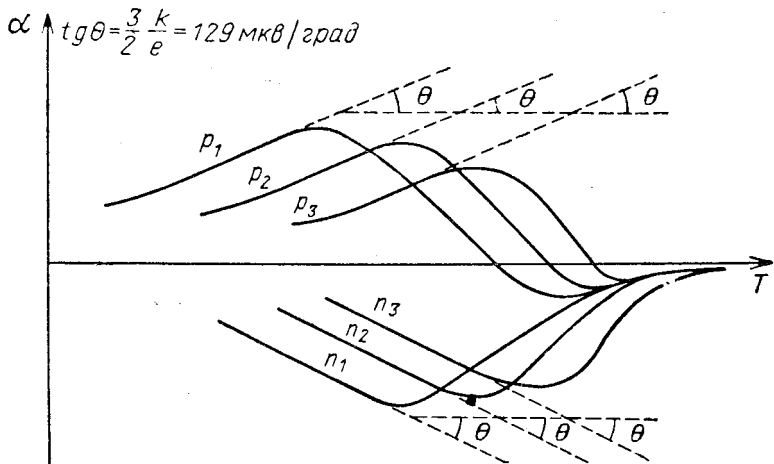


Рис. 6.4. Температурная зависимость термо-э. д. с. для ряда образцов полуметалла с различным содержанием примесей: $n_1 < n_2 < n_3$; $p_1 < p_2 < p_3$.

счет увеличения концентрации примесных электронов; затем наступает участок логарифмического роста термо-э. д. с., соответствующий истощению примесей, и затем более крутое падение в области собственной проводимости. При большой концентрации примесей область истощения может отсутствовать.

Выше были представлены концентрационные и температурные зависимости термо-э. д. с. для простой зонной структуры — одной электронной и одной дырочной зоны, и для постоянных эффективных масс дырок и электронов.

*) Более точно, согласно (6.42) термо-э. д. с. в области собственной проводимости определяется не только соотношением подвижностей дырок и электронов, но и соотношением эффективных масс.

В действительности этот случай очень редко встречается на практике.

Как уже упоминалось, в большинстве случаев энергетический спектр и электронов и дырок имеет значительно более сложный характер, и мы рассмотрим здесь, как это отражается на термоэлектродвижущей силе.

В ряде случаев (в некубических кристаллах и в кубических, если экстремум расположен не в центре зоны Бриллюэна)

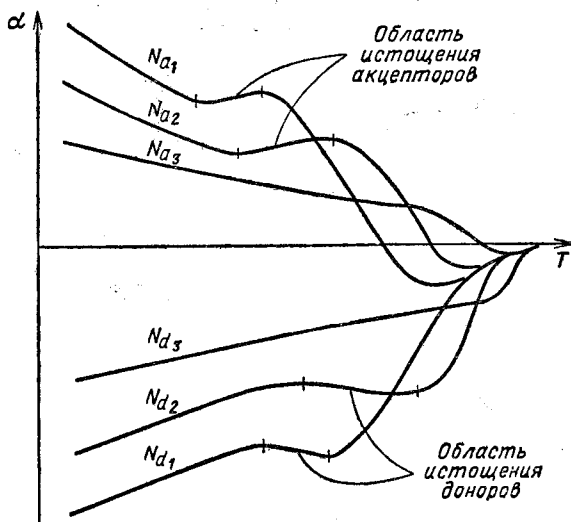


Рис. 6.5. Температурная зависимость термо-э. д. с. для ряда образцов полупроводника с различным содержанием доноров (N_d) и акцепторов (N_a): $N_{d1} < N_{d2} < N_{d3}$; $N_{a1} < N_{a2} < N_{a3}$.

люэна) эффективная масса анизотропна. При этом, однако, можно ввести эффективную массу плотности состояний

$$m_{\text{плс}} = \sqrt[3]{m_1 m_2 m_3}, \quad (6.43)$$

и с заменой m на $m_{\text{плс}}$ выражения для α (6.37) и (6.39) остаются прежними. Если имеется несколько (N) эквивалентных эллипсоидов, то выражение (6.43) принимает вид

$$m_{\text{плс}} = N^{2/3} \sqrt[3]{m_1 m_2 m_3}, \quad (6.44)$$

а (6.37) и (6.39) сохраняют тот же вид.

Таким образом, анизотропия и многоэллипсоидный характер спектра не сказываются непосредственно на термо-э. д. с.

Совершенно иная картина будет, если мы имеем дело с несколькими неэквивалентными экстремумами; в этом случае следует вычислять термо-э. д. с. по общей формуле

$$\alpha = \frac{\sum \alpha_i \sigma_i}{\sum \sigma_i}, \quad (6.45)$$

где α_i и σ_i — «парциальные» термо-э. д. с. и электропроводность и суммирование производится по всем экстрему-

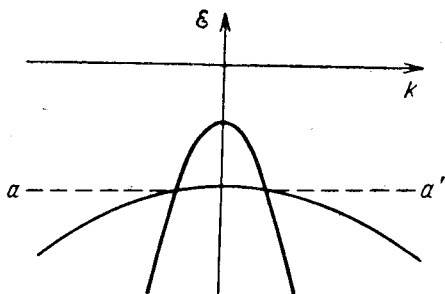


Рис. 6.6. Энергетический спектр дырок в теллуридах германия и олова.

мам (причем для дырок α_i берется с плюсом, а для электронов с минусом). В этом случае зависимость термо-э. д. с. от концентрации и от температуры может иметь значительно более сложный характер. Это проиллюстрировано рис. 6.6 и 6.7 на примере теллуридов германия и олова.

На рис. 6.6 представлен схематически энергетический спектр дырок в этих соединениях, а на рис. 6.7, *a* и *б* — зависимость термо-э. д. с. от концентрации носителей. Качественно эту зависимость можно объяснить следующим образом.

Когда концентрация дырок невелика и уровень химического потенциала проходит выше линии *a — a'*, зона тяжелых дырок целиком заполнена электронами и термо-э. д. с. описывается кривой *1* для легких дырок.

Напротив, при уровне химического потенциала, проходящем значительно ниже *a — a'*, термо-э. д. с. целиком

определяется тяжелыми дырками (так как число их много больше, чем легких) и описывается кривой 2 (термо-э. д. с. при этом значительно больше, так как $m_T \gg m_L$); в промежуточной $p_1 - p_2$ области термо-э. д. с. описывается дву-членной формулой

$$\alpha = \frac{\alpha_L \sigma_L + \alpha_T \sigma_T}{\sigma_L + \sigma_T} \quad (6.46)$$

и постепенно переходит с кривой 1 на кривую 2.

В теллуриде олова термо-э. д. с. в промежуточной области падает до низких значений, которые нельзя объяс-

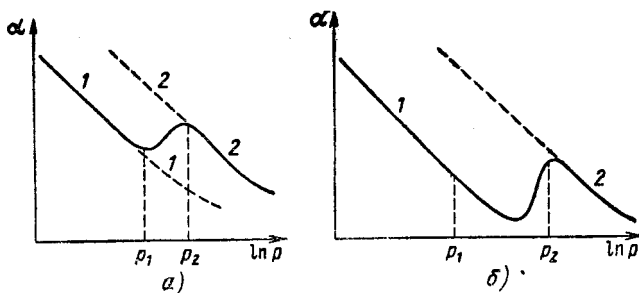


Рис. 6.7. Зависимость термо-э. д. с. от концентрации носителей в теллуриде германия (а) и в теллуриде олова (б).

нить только наличием вырождения; в этом случае имеет место межзонное рассеяние и длина свободного пробега перестает быть простой степенной функцией от энергии.

В переходной $p_1 - p_2$ области средняя энергия в потоке легких дырок очень сильно падает за счет их рассеяния при переходах в тяжелую зону. В результате этого термо-э. д. с. резко падает (а вообще говоря, может переменить знак). Еще более сложный вид могут иметь экспериментальные зависимости термо-э. д. с. в случае непараболической зонной структуры.

Теория показывает, что зависимость энергии от волнового вектора вблизи экстремума носит параболический характер лишь в том случае, когда данный экстремум достаточно удален в энергетической шкале от всех других. Если это не имеет места, то масса начинает расти с ростом энергии и вероятность рассеяния носителей становится весьма сложной функцией от их энергии. Температурные зависимости термо-э. д. с. также становятся

в рассмотренных выше случаях значительно сложнее. Иногда аномалии в термо-э. д. с. объясняются (по крайней мере частично) температурной зависимостью эффективной массы.

Посмотрим, в каких случаях может наблюдаться анизотропия термо-э. д. с. Как мы уже упоминали выше, анизотропия эффективной массы не приводит к анизотропии термо-э. д. с. Анизотропия электропроводности (и подвижности) также не приводит непосредственно к анизотропии термо-э. д. с. Действительно, если длина свободного пробега носителя зависит от направления его движения:

$$l = f(\varphi) \varepsilon^r, \quad (6.47)$$

но показатель степени r не зависит, то согласно (6.37) термо-э. д. с. также не будет зависеть от направления.

В действительности анизотропия термо-э. д. с. наблюдается в следующих двух случаях:

— если одновременно конкурируют несколько механизмов рассеяния (с различными r), причем их эффективность различна для разных направлений;

— если одновременно имеется несколько сортов носителей и при этом парциальные электропроводности по-разному анизотропны.

6.5. ЭЛЕКТРОННАЯ ТЕПЛОПРОВОДНОСТЬ

В § 1.4 мы коротко рассмотрели два механизма теплопроводности — электронный и фононный; однако все выводы, сделанные при этом, носили грубо качественный характер. Теперь мы рассмотрим эти же вопросы более последовательно. Остановимся также коротко на фотонной теплопроводности, которая сравнительно редко, но все же проявляется в полупроводниках. Начнем рассмотрение теплопроводности с ее электронной составляющей, воспользовавшись законом Видемана и Франца.

Напомним соображения, на основании которых он был получен; в гл. 1 для этого мы воспользовались выражением для теплопроводности

$$\kappa_{эл} = \frac{1}{3} C v_0 l \quad (6.48)$$

и электропроводности

$$\sigma = \frac{e^2 n \bar{l}}{2m v_0}. \quad (6.49)$$

Поделив одно на другое, сократив n и \bar{l} и заменив $1/2mv_0$ на $3/2kT$, мы, таким образом, получили выражение

$$\frac{\kappa}{\sigma} = 3 \left(\frac{k}{e} \right)^2 T. \quad (6.50)$$

В приведенном выводе есть целый ряд погрешностей:

1. Заменив в двух различных выражениях, (6.48) и (6.49), выражения длины свободного пробега электронов их средними значениями, мы поступили слишком опрометчиво, так как, по существу, приняли негласно допущение, что обе эти величины не зависят от энергии. Полагая $l \sim \varepsilon^r$, как это мы делали в предыдущей главе, и подставляя это выражение в (6.48) и (6.49), получаем $\kappa_{эл} \sim \varepsilon^{r+\frac{1}{2}}$, $\sigma \sim \varepsilon^{r-\frac{1}{2}}$, и, следовательно,

$$\frac{\kappa}{\sigma} = f \left(\frac{\varepsilon^{r+\frac{1}{2}}}{\varepsilon^{r-\frac{1}{2}}} \right). \quad (6.51)$$

Таким образом, выражение (6.51) показывает, что отношение κ/σ должно зависеть от механизма рассеяния носителей.

2. Из (6.48) и (6.49) очевидно также, что κ/σ должно зависеть от степени вырождения.

3. Мы не учли влияния термо-э. д. с. на тепловой поток, переносимый электронами.

4. Наконец, далеко не очевидно, что в (6.48) и (6.49) входит одна и та же средняя длина свободного пробега.

Чтобы понять это, рассмотрим более подробно, какой вид имеет функция распределения при наличии электрического и теплового поля.

В первом случае электрическое поле E_x смещает функцию распределения на величину

$$\Delta v_x = \omega_x \tau = \frac{eE_x}{m} \tau \quad (6.52)$$

и несколько деформирует ее (как показано на рис. 6.8), так как τ для электронов с различной энергией различно.

Во втором случае дело обстоит сложнее. На рис. 6.9 схематически представлена функция распределения f для трех точек образца при наличии градиента температуры,

в точку b приходят электроны из точек a и b , поэтому функция распределения в точке b сугубо несимметрична. Однако при таком виде функции распределения f в образце существует не только тепловой поток, но и термоэлектри-

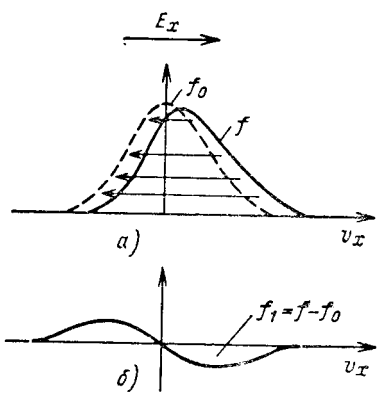


Рис. 6.8. Изменение функции распределения под действием электрического поля (а), неравновесная добавка к функции распределения (б).

ческий ток, теплопроводность же обычно измеряется при разомкнутой цепи; при этих условиях на холодном конце образца накапливаются горячие электроны и возникает

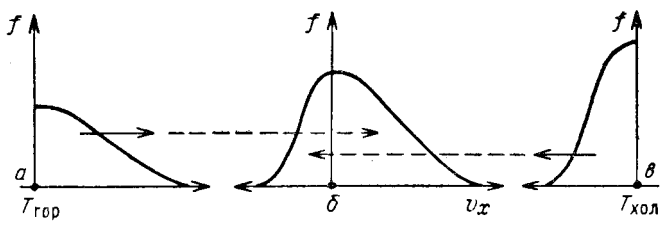


Рис. 6.9. Изменение функции распределения под действием градиента температуры.

термоэлектрическое поле, смещающее функцию распределения в обратном направлении. Таким образом, в окончательном виде она имеет вид, представленный на рис. 6.10, a и b .

Сопоставление рис. 6.9 и рис. 6.10 показывает, что для восстановления равновесия в первом и втором случае могут играть роль совершенно разные столкновения. В первом случае основную роль играют так называемые гори-

горизонтальные переходы *), т. е. упругие столкновения, связанные с отклонением электронов на большие углы (порядка 180°), эти переходы показаны на рис. 6.9 горизонтальными стрелками.

Во втором случае, напротив, основную роль должны играть почти вертикальные переходы (т. е. неупругие

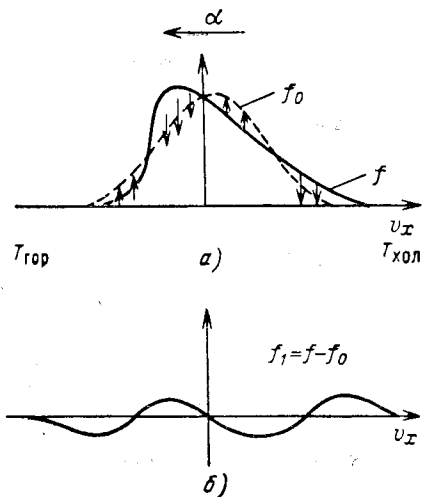


Рис. 6.10. Измененная функция распределения под действием градиента температуры и поля термо-э. д. с. (а) и неравновесная добавка к функции распределения (б).

столкновения); электроны, находящиеся за пределами поверхности Ферми, попадают после такого перехода внутрь ее и меняют при этом свою энергию.

Ясно, что при этом сокращать длину свободного пробега в (6.48) и (6.49), где она определяется различными процессами, в общем случае нельзя.

Положение облегчает тот факт, что при рассеянии на дефектах, как мы показали в гл. 5, все столкновения почти упруги, поэтому l в обоих выражениях оказывается одним и тем же. При рассеянии электронов на акустических тепловых колебаниях решетки при высоких температурах столкновения также почти упруги, при низких температурах и при рассеянии на оптических колебаниях заметную роль начинают играть неупругие столкновения, которые ограничивают теплопроводность и слабо влияют

*) Переходы внутри одного шарового слоя в пространстве импульсов.

на электропроводность; в этих последних случаях электронная теплопроводность оказывается меньше, чем это следует из закона Видемана и Франца. Этот случай требует особого рассмотрения, так как выражение для функции распределения (5.67), полученное в предположении, что все столкновения упругие, в данном случае неприменимо.

Неупругими являются также межэлектронные столкновения; как мы уже упоминали, на электропроводность они влияют лишь косвенно, так как не меняют полного импульса электронной подсистемы, а лишь перераспределяют его и приводят, таким образом, к изменению средней энергии. При наличии разности температур мы имеем встречные потоки электронов, и эти столкновения влияют непосредственно. Таким образом, и абсолютное число столкновений, и показатель степени r в κ (α), с одной стороны, и в σ с другой, могут быть различны.

Для того чтобы получить строгое выражение для отношения теплопроводности и электропроводности, можно воспользоваться выражением (5.68 а) и вывести аналогичное выражение для теплопроводности. Для этого воспользуемся выражением для функции распределения при наличии градиентов потенциала и температуры

$$f = \tau v_x \frac{\partial f_0}{\partial \epsilon} \left[\frac{\epsilon - \mu}{T} + \frac{\partial (\epsilon \phi - \mu)}{\partial x} \right] \quad (6.53)$$

и напомним снова выражения для электрического тока и теплового потока:

$$j = \int e f g(\epsilon) d\epsilon \quad (6.54)$$

и

$$Q = \int \epsilon f g(\epsilon) d\epsilon. \quad (6.55)$$

Приравняв ток к нулю, находим поле термо-э. д. с.:

$$\frac{\partial \left(\frac{\mu}{e} - \phi \right)}{\partial x} = \frac{k}{e} \left[\frac{1}{T} \frac{\int \epsilon^2 \frac{\partial f_0}{\partial \epsilon} \tau g(\epsilon) d\epsilon}{\int \epsilon \frac{\partial f_0}{\partial \epsilon} \tau g(\epsilon) d\epsilon} - \mu^* \right]. \quad (6.56)$$

Подставив (6.56) в (6.53) и затем в (6.55), можно получить выражение для теплового потока с учетом термо-э. д. с. и электронную теплопроводность.

Для случая параболической зоны $g(\varepsilon) = 4\pi/h^3 (2m)^{3/2} \times \varepsilon^{1/2} d\varepsilon$ и степенной зависимости длины свободного пробега электрона от его энергии $l = l_0(T) \varepsilon^r$, получим:

$$\kappa_{эл} = \frac{16\pi m l_0(T) k}{3h^3} (kT)^{r+2} \left[(r+3) F_{r+2}(\mu^*) - \frac{(r+2)^2}{(r+1)} \frac{F_{r+1}^2(\mu^*)}{F_r(\mu^*)} \right]. \quad (6.57)$$

Согласно (5.68а) и (6.57) $\kappa/\sigma = A (k/e)^2 T$, где A — безразмерный коэффициент, равный

$$A = \left[\frac{r+3}{r+1} \frac{F_{r+2}(\mu^*)}{F_r(\mu^*)} - \frac{(r+2)^2}{(r+1)^2} \frac{F_{r+1}^2(\mu^*)}{F_r^2(\mu^*)} \right]. \quad (6.58)$$

В случае невырожденного электронного газа $A = r + 2$, в случае полного вырождения $A = \pi^2/3$, в промежуточных

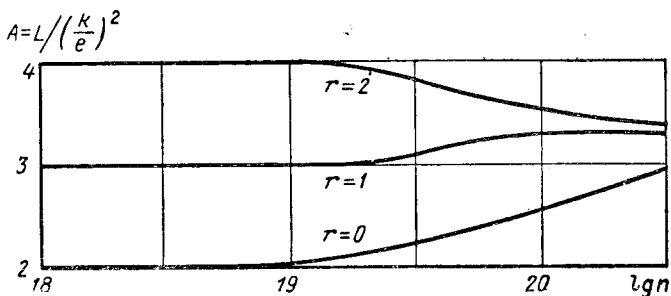


Рис. 6.11. Зависимость числа Лоренца (т. е. постоянной в законе Видемана — Франца) от механизма рассеяния и степени вырождения электронов проводимости.

случаях A , как это видно из (6.58), является функцией от r и μ^* . На рис. 6.11 для иллюстрации приведена зависимость A от концентрации носителей для трех случаев: $r = 0$ (рассеяние на акустических колебаниях), $r = 1$ (рассеяние на оптических колебаниях) и $r = 2$ (рассеяние на ионах примеси). Как мы уже упоминали, при наличии неупругого рассеяния (межэлектронные столкновения и рассеяние на оптических колебаниях) полученные выше выражения для A будут неточны.

Если исходить из фононных представлений, т. е. рассматривать распространение упругих волн как хаотическое движение фононного газа, то можно получить для теплопроводности решетки выражение, формально совершенно идентичное (6.48):

$$\kappa_p = \frac{1}{3} Cl\omega_{гр} = \frac{1}{3} C\tau\omega_{гр}^2, \quad (6.59)$$

где C — теплоемкость решетки; l и τ — длина и время свободного пробега фононов и $\omega_{гр}$ — их скорость.

Это же выражение для κ_p получил в 1914 г. Дебай из волновой картины, рассматривая кристалл как упругий континуум и учитывая рассеяние тепловых волн на изменениях плотности, вызванных ангармоничностью колебаний.

Однако выражение (6.59) было бы удобно для анализа, если бы все величины, входящие в его правую часть, были одинаковы для всех колебаний. В действительности это, разумеется, не так, и вычисление средних величин связано в данном случае с такими трудностями, что формула (6.59) в таком виде, как она написана, даст немногим больше, чем представление о размерности κ_p .

Более того, недостаточно осторожное применение этой формулы может привести к серьезным ошибкам. В гл. 2 мы установили, что за перенос энергии «ответственна» групповая скорость фононов $\omega_{гр}$, которая колеблется в весьма широких пределах. Групповая скорость для всех оптических ветвей (представляющих, по сути дела, внутримолекулярные колебания) относительно мала, поэтому в (6.59) можно исключить часть теплоемкости, обусловленную этими колебаниями. Но как мы видели, для коротковолновых акустических колебаний $\omega_{гр} = 0$, и их, следовательно, тоже надо исключить из (6.59); таким образом, значение ω и C в (6.59) становится весьма неопределенным.

Вычислить среднюю длину свободного пробега еще сложнее, и расчет и опыт показывают, что при этом усреднении легко ошибиться в сотни раз, так как длина свободного пробега для фононов различной частоты и поляризации отличается на много порядков; так, для продольных длинноволновых колебаний акустической ветви она дости-

гает макроскопических размеров *), в то же время для коротковолновых колебаний l бывает ничтожно мало — порядка межатомных расстояний. Таким образом само представление о волнах, а следовательно и фонах, становится сомнительным.

Впрочем, как это часто бывает, исключительные трудности приводят в конце к упрощению ситуации. В данном случае чрезвычайно широкий диапазон длин свободного пробега приводит к тому, что в теплопроводности необходимо учитывать в первую очередь перенос тепла длинноволновыми продольными акустическими волнами, что значительно упрощает задачу. После этих вводных замечаний перейдем непосредственно к вычислению κ_p . Будем исходить из общего выражения для теплового потока

$$Q = \sum N_q \hbar \omega_q \bar{v}_q \quad (6.60)$$

(аналогичного выражению $Q = \varepsilon n \bar{v}_x$, где $\varepsilon = \hbar \omega$, $n = \sum N_q$ и $\bar{v}_x = \bar{\omega}_q$).

Суммирование, вообще говоря, надо проводить по всем ветвям, по всем длинам волн и по всем (трем) видам поляризации. В таком виде подсчет (6.60) был бы неразрешимой задачей, однако с учетом замечаний, сделанных выше, можно ограничиться суммированием по длинноволновым продольным акустическим фононам. Для этих колебаний кристалл можно рассматривать как континуум и сумму заменить интегралом; таким образом, задача сильно упрощается.

Основная трудность остается в вычислении N_q — неравновесной функции распределения, при этом мы не можем заменить $N_q \rightarrow N_{q0}$ равновесным распределением Эйнштейна, так как поток обратится в нуль. Для того чтобы определить N_q , мы должны решить кинетическое уравнение для фононов:

$$\left(\frac{\partial N_q}{\partial t} \right)_{\text{дрейф}} = \left(\frac{\partial N_q}{\partial t} \right)_{\text{столкн}} \quad (6.61)$$

Вычисление дрейфового члена в данном случае еще более просто, чем в случае электронов (так как фононы —

*) Непосредственное измерение этой величины на кремнии при комнатной температуре дало «астрономическое значение» — десятки микрон.

незаряженные частицы и в данном случае отсутствует член, содержащий электрическое поле)*):

$$\left(\frac{\partial N_q}{\partial t}\right)_{\text{дрейф}} = \frac{\partial N_q}{\partial x} \frac{\partial x}{\partial t} = \frac{\partial N_q}{\partial T} \frac{\partial T}{\partial x} \frac{\partial x}{\partial t} = \frac{\partial N_q}{\partial T} \omega_{\text{гр}} \nabla T. \quad (6.62)$$

Значительно сложнее обстоит дело с членом столкновений. Рассмотрим, какими причинами может вызываться рассеяние фононов. В идеальном кристалле с упругими межатомными силами гармонические волны распространяются независимо друг от друга, не взаимодействуют и не затухают, точно так же, как электромагнитные колебания в пустоте. Это видно из того, что общее решение уравнения (2.146) представляет собой сумму независимых незатухающих колебаний. Иными словами, для идеального кристалла гармонические волны являются стационарными состояниями; их вероятность рассеяния — вероятность перехода одного в другое, равна нулю.

Эта вероятность становится отличной от нуля в результате возмущений — нарушения периодичности потенциала. Такими возмущениями могут быть:

1) чужеродные атомы, атомы в междуузлиях и пустые узлы (так называемые «точечные дефекты»**).

2) линейные дефекты — дислокации;

3) плоские дефекты — границы кристалла и зерен в поликристаллическом образце;

4) объемные дефекты — трещины, полости, поры, включения другой фазы;

5) ангармоничность колебаний, обуславливающая их взаимодействие, или, на корпускулярном языке, рассеяние фононов на фононах.

Если первые четыре причины рассеяния могут (по крайней мере в принципе) отсутствовать, то последняя всегда есть и в этом смысле является наиболее общей и наиболее важной, поэтому мы начнем с ее рассмотрения. Взаимодействие нормальных колебаний можно рассматривать и на волновом и на корпускулярном языке. На волновом языке

*) Будем считать, что градиент температуры направлен по оси x , и рассматривать одномерную задачу.

**) Все перечисленные выше дефекты можно называть точечными лишь в том случае, когда размер вносимого ими искажения много меньше длины волны. В действительности это не всегда имеет место (не для всех перечисленных дефектов и не для всех длин волн).

это означает следующее. Если по кристаллу распространяется волна $\Phi_1 = A_1 e^{2\pi i(q_1 x - \omega_1 t)}$, то из-за ангармоничности она вызывает периодические колебания (модуляцию) упругих постоянных f и плотности ρ , т. е. создает таким образом объемную решетку, движущуюся с фазовой скоростью $\omega_\Phi = \omega/q$ по кристаллу*). Другие волны будут дифрагировать на этой решетке, причем нетрудно показать, что если волна Φ_2 взаимодействует с волной, то при этом появится новая волна, амплитуда которой зависит от взаимодействия волн Φ_1 и Φ_2 , т. е. от степени ангармоничности колебаний, а частота и волновой вектор удовлетворяют условиям

$$\omega_3 = \omega_1 \pm \omega_2, \quad (6.63)$$

$$q_3 = q_1 \pm q_2. \quad (6.64)$$

На корпускулярном языке это означает, что в результате столкновений два фонана (ω_1, q_1) и (ω_2, q_2) превратились в фонon третьего сорта (ω_3, q_3) . С этой корпускулярной точки зрения соотношения (6.63) и (6.64) означают соблюдение при таких (так называемых трехфононных) столкновениях законов сохранения энергии и импульса. Приведенные выше рассуждения относятся к знаку плюс в формулах (6.63) и (6.64). Знак минус перед ω_2 и q_2 означает, что фонon (ω_1, q_1) «распался» на два (ω_2, q_2) и (ω_3, q_3) .

На первый взгляд может показаться, что наряду с трехфононными могут существовать более простые, двухфононные процессы рассеяния. Но в действительности, если написать законы сохранения для этого случая, то мы получим

$$\omega_1 = \omega_2 \text{ и } q_1 = q_2 \quad (6.65)$$

и убедимся, что они означают неизменность частоты и волнового вектора, т. е. отсутствие рассеяния. В действительности могут существовать более сложные, четырехфононные процессы, но вероятность их относительно мала.

Возможность существования двухфононных процессов появляется при наличии третьего тела — какого-либо дефекта решетки, которому фонon может отдать часть

*) При гармонических колебаниях $\rho/f = \text{const}$, поэтому такие колебания не вызывают изменения скорости распространения упругих волн и коэффициента преломления, а следовательно, и не рассеивают друг друга.

энергии (при неупругом рассеянии) или импульса (при упругом).

Таким образом, соотношения (6.63) и (6.64) описывают простейшие и наиболее вероятные процессы рассеяния фононов на фононах.

Однако, как показал еще в 1934 г. Паейрлс, такие процессы не могут быть «ответственными» за конечную теплопроводность, т. е. если бы не существовало никаких других столкновений, кроме этих, теплопроводность кристалла была бы бесконечна. Действительно, напишем еще раз выражение для теплового потока:

$$Q = \sum h\omega_q \omega_{гр} N_q. \quad (6.66)$$

Заменяя в (6.66) $\omega_q = \omega_\phi q$ и пренебрегая дисперсией, т. е. считая $\omega_{гр} = \omega_\phi = \text{const}$, получим

$$Q = \bar{\omega}^2 \sum h_q N_q = \bar{\omega}^2 P, \quad (6.67)$$

где P — полный импульс фононной системы. Но соотношение (6.64) показывает, что при рассмотренных выше столкновениях суммарный импульс фононов не меняется и такие столкновения не будут ограничивать теплопроводность кристалла. Такие процессы [т. е. описываемые соотношениями (6.63) и (6.64)] Паейрлс назвал нормальными N -процессами. Но он также показал, что наряду с N -процессами (которые являются единственно разрешенными процессами для непрерывного континуума) в периодической среде (кристаллической решетке) могут происходить процессы, при которых соотношение (6.63) сохраняется, а соотношение (6.64) заменяется более общим:

$$q_3 = q_1 \pm q_2 \pm g, \quad (6.68)$$

где g — вектор обратной решетки.

Такие столкновения были названы umclarr-процессами, или процессами переброса. Это означает, что при таких столкновениях сумма волновых векторов не сохраняется, а меняется скачком на величину g . Так как при этом согласно (6.68) полный импульс фононной системы не сохраняется, то такие столкновения дают конечное тепловое сопротивление кристалла. Качественно процессы переброса можно интерпретировать как комбинированную одновременную дифракцию волны ($q_1 \omega_1$) на движущейся со скоростью $\omega_2 = \omega_2/q_2$ решетке с постоянной решетки,

равной $1/q_2^*$), и на неподвижной кристаллической решетке с постоянной, равной $1/g^{**}$).

С математической точки зрения появление процессов переброса можно понять из следующих соображений. В гл. 4 мы пришли к выводу, что волновой вектор q в кристалле может меняться в пределах $1/2L < q < 1/2a$, так как $q > 2\bar{L}$ не удовлетворяет граничным условиям, а $q > 1/2a$ всегда можно представить как $q = q' + n/2a$, где n подобрано таким образом, что $q' < 1/2a$; при этом оказывается, что волна q' описывает точно те же колебания атомов, что и q , а волна q добавляет к этому не имеющие физического смысла колебания пустоты между атомами.

Представим теперь, что в (6.64) $q_1 > 1/4a$ и $q_2 > 1/4a$, тогда $q_3 > 1/2a$, т. е. q_3 попадает в область «запрещенных» значений. Заметив в соответствии со сказанным выше, что $q_3 = q'_3 + g$, где g — вектор обратной решетки, мы получим уравнение (6.68), описывающее процессы переброса.

Из сказанного не следует делать вывод, что N -процессы можно вообще исключить из рассмотрения (так как они не вносят конечного теплового сопротивления). В действительности их весьма существенное влияние проявляется в том, что они приводят к перераспределению импульса между фононами. (С подобными явлениями мы уже сталкивались при анализе влияния межэлектронных столкновений, см. гл. 4 и 5.)

Как мы уже много раз упоминали, обычные столкновения (т. е. столкновения с потерей импульса) стремятся установить равновесное распределение по скоростям и энергиям; для электронов — это распределение Ферми

$$f_0 = \frac{1}{e^{\frac{\varepsilon - \mu}{kT}} + 1},$$

для фононов — это распределение Бозе — Эйнштейна

$$N_q^* = \frac{1}{e^{\frac{\hbar\omega_q}{kT}} - 1}. \quad (6.69)$$

*) Эта решетка создается колебаниями плотности, обусловленными второй волной.

**) Так как кристаллическая решетка неподвижна, то ее участие в соударении не вносит никаких дополнительных членов в уравнение (6.63).

N -процессы не могут установить нормального распределения, так как общий импульс, а следовательно, и средняя скорость дрейфа всей системы фононов при них сохраняется; можно показать, что они приводят к *нормальному распределению в потоке*:

$$N_q(\lambda) = [\exp \hbar (\omega_q + \lambda q) / kT - 1]^{-1}, \quad (6.70)$$

где λ — параметр, характеризующий скорость потока. Наличие в (6.70) члена λq приводит к тому, что количество фононов с одинаковым по абсолютным значениям q , а следовательно, и $\omega_q = \omega_{\phi} q$, но с противоположным направлением движения, ($-q$) и ($+q$), становится неодинаковым. Как видно из (6.70), при положительном λ

$$N_{+q}(\lambda) < N_{-q}(\lambda). \quad (6.71)$$

Можно показать, что (6.70) представляет собой распределение, движущееся в направлении $-q$ со скоростью 2λ . Заменяя в (6.70) $\omega_q = \omega q$, можно переписать его в виде

$$N_{\pm q}(\lambda) = [\exp \hbar [q] (\omega \pm \lambda) / kT - 1]^{-1}, \quad (6.72)$$

где уже явно видно, что (6.72) характеризует распределение, движущееся со скоростью 2λ .

Учитывая лишь косвенное влияние N -процессов на тепловые сопротивления, рассмотрим интеграл столкновений, учитывая пока лишь процессы переброса.

Будем искать N_q в виде

$$N_{q1} = N_{q1}^{\circ} + n_{q1}, \quad (6.73)$$

где N_{q1}° — равновесная функция распределения и n_{q1} — небольшая неравновесная добавка.

Как следует из (6.63) и (6.64), столкновения фононов (в отличие от электронных столкновений) отнюдь не упруги: $\omega_{q1} \neq \omega_{q3} \neq \omega_{q2}$; это значит, что скорость восстановления равновесия для фононов с волновым вектором q_1 зависит от чисел заполнения во всем фазовом пространстве. Следовательно, если мы попытаемся ввести время релаксации $\tau(q_1)$, то оно окажется весьма сложной функцией от времени и от всех чисел заполнения (q_2 и q_3), т. е. по существу потеряет всякий смысл.

Однако положение облегчают следующие обстоятельства. Как мы уже упоминали, за теплопроводность в основном «ответственны» длинноволновые фононы, для которых

время релаксации и длина свободного пробега много больше, чем для всех остальных. В то же время длинноволновые фононы теряют импульс в основном при столкновениях с коротковолновыми фононами. Это следует из того, что, во-первых, коротковолновых фононов во много раз больше и, во-вторых, как видно из (6.68), длинноволновые фононы не могут непосредственно принимать

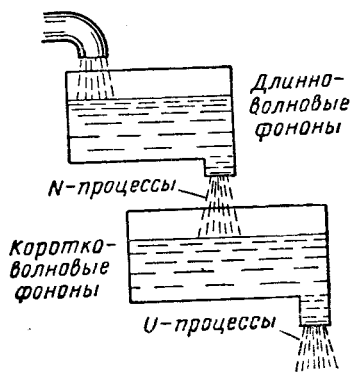


Рис. 6.12. Иллюстрация механизма рассеяния длинноволновых и коротковолновых фононов.

(так как для коротковолновых фононов время релаксации намного меньше, то и их числа заполнения куда ближе к равновесным). После этого правая часть (6.61) становится функцией от $N_{q_2}^{\circ}$, $N_{q_3}^{\circ}$ (которые есть не что иное, как распределение Бозе — Эйнштейна для соответствующих групп фононов) и от единственной неизвестной величины

$$n_{q_1} = N_{q_1} - N_{q_1}^{\circ}, \quad (6.74)$$

которую и нужно найти. При этих условиях правая часть (6.61) может быть приведена к виду

$$\left(\frac{\partial N_{q_1}}{\partial t} \right)_{\text{столкн}} = \frac{N_{q_1} - N_{q_1}^{\circ} *}{\tau_q}, \quad (6.75)$$

т. е. это значит, что уже может быть введено время релаксации.

*) В дальнейшем индекс «1» при q мы будем опускать.

Заменяя в левой части (6.61) N_q на N_q^* , получаем*)

$$\omega_q \frac{\partial N_q^*}{\partial T} \nabla T = \frac{N_q - N_q^*}{\tau(q)} = \frac{n_q}{\tau(q)}, \quad (6.76)$$

и, следовательно,

$$n_q = \tau(q) \omega_q \nabla T \frac{\partial N_q^*}{\partial T}. \quad (6.77)$$

Подставив полученное выражение (6.77) для n_q — неравновесной части функции распределения фононов, вместо N_q в (6.66), можно получить выражение для фононной теплопроводности. Таким образом, при сделанных выше упрощениях вопрос сводится к определению времени релаксации для фононов.

Однако эта задача весьма трудна. Преимуществом выражения (6.77) является то, что оно написано в таком общем виде, что в него могут быть включены все виды рассеяния:

$$\tau_q^{-1} = \tau_L^{-1} + \tau_I^{-1} + \tau_p^{-1} + \tau_d^{-1} + \tau_\phi^{-1}, \quad (6.78)$$

где τ_L — время релаксации для рассеяния на границах зерен; τ_I — время релаксации для рассеяния за счет так называемого изотопического эффекта, т. е. разности масс примесных атомов и основных атомов решетки; τ_p — время релаксации, обусловленное изменением атомных объемов и коэффициента упругой связи, вносимых примесными атомами; τ_d — время релаксации при рассеянии на дислокациях и τ_ϕ — время релаксации при рассеянии фононов на фононах.

При этом для каждого процесса рассеяния должно быть в отдельности вычислено соответствующее время релаксации, что было сделано рядом авторов с большей или меньшей точностью (см., например, [21]). Основным результатом этих вычислений являются дисперсионные зависимости, т. е. зависимости времени релаксации τ от частоты ω , которые мы приведем ниже:

1) при «изотопическом» рассеянии (т. е. рассеянии из-за разности масс примесного и основных атомов $\tau_I^{-1} = I\omega^4$, где I — коэффициент, пропорциональный числу примесей N_I и разности масс в квадрате: $I \sim N_I (\Delta M_I)^2$; в ряде

*) Подробно о методах решения кинетического уравнения для фононов см. [12] и [21].

случаев этот коэффициент может быть вычислен в явном виде [см. 21];

2) при рассеянии на границах кристалла $\tau_L^{-1} = c/L$, где c — коэффициент порядка единицы, а L — характерная длина — приблизительно радиус круга, по сечению равного наименьшему сечению кристалла (следует отметить, что это единственный тип рассеяния, при котором время релаксации не зависит от частоты);

3) при рассеянии на деформациях решетки $\tau_p^{-1} \sim q^4$;

4) при рассеянии на дислокациях $\tau_d^{-1} \sim q$;

5) наибольшие трудности представляет вычисление времени релаксации для рассеяния фононов на фононах τ_ϕ . Первым делом мы должны отметить тот факт, что непосредственно в выражение (6.78) для τ_q^{-1} можно включить только U -процессы, приводящие к установлению равновесного распределения (N_q°), но нельзя включить N -процессы, так как они приводят к распределению N_q^λ , а не N_q° .

По аналогии с (6.75) можно написать для N столкновений

$$\left(\frac{dN}{dt} \right)_{\text{столкн}} = \frac{N_q^\lambda - N_q}{\tau_N} \quad (6.79)$$

и с учетом N -процессов и всех остальных видов рассеяний, входящих в τ_q^{-1} , переписать (6.79) в виде

$$\left(\frac{\partial N}{\partial t} \right)_{\text{столкн}} = \frac{N_q^\circ - N_q}{\tau_q} + \frac{N_q^\lambda - N_q}{\tau_N}. \quad (6.80)$$

Параметр λ может быть вычислен из того условия, что нормальные процессы не изменяют полного импульса системы фононов. Уравнение (6.80) после этого приводится к виду, аналогичному (6.75), при этом для τ_ϕ получается следующее выражение (см. [30]):

$$\tau_\phi^{-1} = (B_N + B_U) T^3 \omega^2, \quad (6.81)$$

где коэффициенты B_N и B_U характеризуют интенсивность нормальных процессов и процессов переброса. Таким образом, более или менее точно могут быть вычислены все слагаемые, входящие в (6.78).

Зная выражение для τ_q , можно воспользоваться соотношением (6.60) для определения потока тепла и теплопроводности:

$$Q = \sum n_q \hbar \omega_q v_q. \quad (6.82)$$

Так как в теплопроводности принимают участие только длинноволновые фононы, то сумма (6.82) может быть заменена интегралом; таким образом получаем

$$\kappa_p = \frac{4\pi}{3(2\pi)^3} \int \tau(q) C(q) \omega_q^2 q^2 dq, \quad (6.83)$$

где

$$C(q) = \frac{d}{dT} (\hbar\omega_q N_q^*) = \frac{(\hbar\omega)^2}{kT^2} \frac{e}{(e^{\frac{\hbar\omega}{kT}} - 1)^2}. \quad (6.84)$$

Следует подчеркнуть, что из (6.78) отнюдь не следует, что тепловое сопротивление решетки $1/\kappa_p$ складывается аддитивно из сопротивлений, обусловленных различными видами дефектов и межфононным взаимодействием:

$$\frac{1}{\kappa_p} \neq \sum_i \frac{1}{\kappa_i}. \quad (6.85)$$

Объясняется это тем, что различные виды рассеяния имеют различные зависимости от ω (а следовательно, и от q) и от температуры. Поэтому при наличии нескольких видов рассеяния интегрирование выражения (6.83) дает сложную зависимость от температуры и концентрации различных видов дефектов. Правило Маттисона (т. е. аддитивность различных видов сопротивлений) имеет место тогда, когда один тип рассеяния превалирует над всеми остальными. Отметим некоторые закономерности, которые вытекают из выражения (6.83).

Так как при достаточно низких температурах все виды колебаний «вымерзают», за исключением очень длинных волн, а для последних играет роль рассеяние только на границах зерен и при этом l не зависит от температуры, то поэтому при $T \rightarrow 0$ теплопроводность должна изменяться аналогично изменению теплоемкости, т. е. $\kappa \sim T^3$ при $T \rightarrow 0$. По мере роста температуры начинают «бороться» два фактора: 1) рост числа колебаний, переносящих тепло и 2) рост межфононного рассеяния. Короткие волны и оптические фононы почти не переносят тепло, но играют существенную роль в рассеянии длинноволновых фононов, поэтому теплопроводность переходит через максимум при температуре, значительно ниже дебаевской, и вначале после максимума падает быстрее, чем T^{-1} (рис. 6.13).

В кристаллах, не содержащих дефектов, теплопроводность при температурах выше дебаевской обратно про-

порциональна температуре при небольшом количестве дефектов:

$$\frac{1}{\kappa_p} = aT + b \quad (6.86)$$

(т. е. сопротивление, вносимое фононным и дефектным рассеяниями, аддитивно). В промежуточной области зависимость κ от температуры, числа и характера дефектов может носить весьма сложный характер.

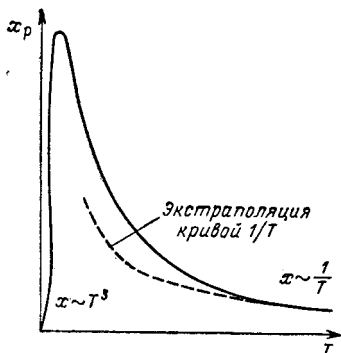


Рис. 6.13. Температурная зависимость теплопроводности решетки для идеального кристалла конечных размеров.

В заключение заметим, что для решения кинетического уравнения (6.61) для фононов наряду с методом времени релаксации широко применяются вариационные методы. Сущность их заключается в следующем.

Как упоминалось в гл. 4, в адиабатически изолированной системе в равновесии устанавливается такое распределение частиц по энергиям и пространству, которое соответствует максимуму энтропии. Это и есть распределение

Ферми для электронов, Бозе — Эйнштейна для фотонов и фононов, Максвелла — Больцмана для молекул газа и т. д.

При этом экспоненциальная зависимость вероятности заполнения данного состояния от его энергии является следствием того, что данная частица, приобретая какую-то энергию ϵ , отнимает ее у остальных $(N - 1)$ частиц, составляющих систему, и тем самым уменьшает объем, который могут занять эти частицы в многомерном фазовом пространстве, пропорционально множителю $((\mathcal{E} - \epsilon)/\mathcal{E})^{N-1}$, где \mathcal{E} — энергия всей системы. Это выражение при $N \rightarrow \infty$ и дает экспоненциальную зависимость от ϵ и от температуры $T \sim \mathcal{E}/N$.

В термодинамике необратимых процессов доказывается, что если система выведена из состояния равновесия, то в ней устанавливается такое распределение, которое обеспечивает максимальную скорость приближения к состоянию равновесия, т. е. максимальную скорость возрастания (генерации) энтропии. Это положение и положено

в основу вариационного метода определения неравновесной функции распределения фононов.

Энтропия системы определяется выражением $S = Q/T$. Следовательно, скорость возрастания энтропии \dot{S} при наличии градиента температуры и потока энергии

$$Q = \int h\omega_q w_{\text{гр}} N_q g(q) dq \quad (6.87)$$

будет выражаться следующим образом:

$$\dot{S} = Q \text{grad} \frac{1}{T} = \int h\omega_q w_{\text{гр}} N_q g(q) \text{grad} \frac{1}{T} dq. \quad (6.88)$$

Неравновесная функция распределения N_q или $n_q = N_q - N_q^0$ может быть найдена из условия, чтобы \dot{S} было максимально. Сначала из физических соображений находится вид пробной функции распределения N_q , а затем путем вариации выражения (6.88) находят значения параметров [12]. При этом точность решения зависит от того, насколько удачно выбрана пробная функция. Недостатком вариационного метода является то, что степень точности в общем не может быть оценена — всегда остается открытым вопрос, нельзя ли при другом выборе пробной функции обеспечить большую скорость роста энтропии.

6.7. ФОТОННАЯ ТЕПЛОПРОВОДНОСТЬ

В тех случаях, когда коэффициент поглощения в области теплового излучения достаточно мал (т. е. велика длина свободного пробега фотонов), может играть существенную роль перенос тепла электромагнитным излучением.

Формулу для соответствующей теплопроводности $\kappa_{\text{ф}}$ можно вывести, если исходить из корпускулярных представлений и применить к фотонному газу формулу, выведенную для обычного классического газа:

$$\kappa_{\text{ф}} = \frac{1}{3} C_{\text{ф}} l v. \quad (6.89)$$

Подставив в (6.89) фотонную теплоемкость $C_{\text{ф}} = \frac{16h^3 T^3 \sigma_0}{c}$, скорость фотонов $v = c/n$, где c — скорость света в пустоте, n — показатель преломления и σ_0 — постоянная в законе Стефана — Больцмана, и выразив длину свободного пробега фотонов через коэффициент поглощения ($l = 1/K$),

получим

$$\kappa_{\Phi} = \frac{16}{3} \frac{n^2 \sigma_0 T^3}{K} . \quad (6.90)$$

Формула (6.90) была впервые получена Генцелем для анализа излучательной теплопроводности стекол. Б. Я. Мойжес применил ее к полупроводникам, обобщив при этом на случай анизотропного кристалла и образца, размеры которого сравнимы с обратным коэффициентом поглощения. При этом удалось объяснить аномалии в теплопроводности теллура и получить удовлетворительное согласие с опытными данными.

ГАЛЬВАНОМАГНИТНЫЕ И ТЕРМОМАГНИТНЫЕ ЯВЛЕНИЯ

7.1. ОБЩИЕ СВЕДЕНИЯ

Если проводник (или полупроводник), по которому протекает электрический ток или тепловой поток, поместить в магнитное поле, то в нем возникает ряд эффектов: уменьшается его электро- и теплопроводность, в направлении, перпендикулярном магнитному полю и току, возникает разность потенциалов и разность температур. Каждое из этих явлений имеет свое название (большой частью связанное с именем открывшего его ученого), а все они называются соответственно гальваномагнитными и термомагнитными явлениями.

Иными словами, гальваномагнитными явлениями называют явления, возникающие в проводниках первого рода *) при одновременном воздействии электрического и магнитного полей; термомагнитными явлениями называют эффекты, вызванные одновременным воздействием на проводник и магнитного, и температурного поля.

Рассмотрим коротко характер движения электрона при наличии электрического и магнитного поля, вначале каждого в отдельности, а затем при их одновременном действии на электрон. При этом мы будем рассматривать свободный электрон, т. е. влияние периодического поля учтем приближенно введением изотропной эффективной массы, а столкновениями на первых порах пренебрежем.

*) Т. е. в материалах, в которых ток переносится электронами. В проводниках второго рода, в которых ток переносится ионами, гальваномагнитные явления не обнаруживаются, по-видимому, потому, что скорости ионов и вследствие этого и магнитная составляющая силы Лоренца малы.

В постоянном электрическом поле на электрон действует сила

$$\mathbf{F} = e\mathbf{E} \quad (7.1)$$

и ускорение

$$\mathbf{w}_E = \frac{e}{m} \mathbf{E}. \quad (7.2)$$

Таким образом, скорость электрона изменяется по закону

$$\mathbf{v} = \mathbf{v}_0 + \mathbf{w}_E t,$$

и в общем случае (если направление начальной скорости не совпадает с направлением поля) он движется по параболе

$$\mathbf{r} = \mathbf{v}_0 t + \frac{\mathbf{w}_E t^2}{2}. \quad (7.3)$$

В частном случае, если $v_0 = 0$ или $\mathbf{v}_0 \parallel \mathbf{E}$, эта парабола вырождается в прямую.

В магнитном поле на электрон действует сила

$$\mathbf{F}_H = \frac{e[\mathbf{vH}]}{c} \quad (7.4)$$

и ускорение

$$\mathbf{w}_H = \frac{e}{mc} [\mathbf{vH}]. \quad (7.5)$$

По абсолютной величине согласно (7.4)

$$|\mathbf{F}_H| = \frac{1}{c} evH \sin(\mathbf{vH}), \quad (7.6)$$

т. е. в том случае, когда $\mathbf{v} \parallel \mathbf{H}$, то $\mathbf{F}_H = 0$; если $\mathbf{v} \perp \mathbf{H}$, то $\mathbf{F}_H = evH/c$. В общем случае скорость электрона можно разложить на две составляющие: параллельную и перпендикулярную полю \mathbf{H} :

$$\mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp}; \quad (7.7)$$

магнитная сила будет равна

$$|\mathbf{F}_H| = \frac{ev_{\perp}H}{c}. \quad (7.8)$$

По направлению сила \mathbf{F}_H будет перпендикулярна v_{\perp} и v_{\parallel} и, меняя все время направление v_{\perp} (но не воздействуя при этом на v_{\parallel}), заставит электрон двигаться по винтовой линии вдоль магнитного поля.

Рассмотрим случай, когда $v_{\parallel} = 0$. При этом электрон будет двигаться по окружности, радиус которой можно

найти из условия равенства центробежной и центростремительной сил:

$$\frac{mv^2}{r} = \frac{evH}{c}, \quad (7.9)$$

откуда

$$r = \frac{mvc}{eH}. \quad (7.10)$$

Так как все входящие в (7.10) величины постоянны, радиус траектории будет тоже постоянен; это и подтверждает, что электрон будет двигаться по окружности. Период обращения электронов согласно (7.10) равен

$$T = \frac{2\pi r}{v} = \frac{2\pi mc}{eH}, \quad (7.11)$$

угловая скорость

$$\omega = \frac{2\pi}{T} = \frac{eH}{mc} \quad (7.12)$$

и число оборотов в 1 сек

$$\nu = \frac{\omega}{2\pi} = \frac{eH}{2\pi mc}. \quad (7.13)$$

Как видно из (7.12) и (7.13), угловая скорость и число оборотов не будут зависеть от скорости электрона до тех пор, пока скорость не достигнет релятивистских значений, при которых масса начинает зависеть от скорости.

Перейдем к наиболее важному для нас случаю — к одновременному воздействию на электрон электрического и магнитного полей. Общая сила, действующая на электрический заряд e в электрическом поле E и магнитном поле H , выражается формулой

$$\mathbf{F} = \mathbf{F}_E + \mathbf{F}_H = e \left[\mathbf{E} + \frac{1}{c} [\mathbf{vH}] \right]. \quad (7.14)$$

Как мы уже упоминали, соотношение (7.14) было выведено немецким ученым Лоренцом, и \mathbf{F} поэтому носит название силы Лоренца. Формула (7.14) лежит в основе расчета всех современных ускорителей (циклотрона, бетатрона, синхротрона), масс-спектрографов (т. е. приборов для разделения потока заряженных частиц по их массам), β -спектрометров (приборов для измерения энергий β -частиц). При этом особую роль играет то, что частота ω , входящая в (7.13), зависит только от магнитного поля и отношения e/m ; на этой закономерности, в частности, основан метод

исследования энергетического спектра электронов, названный методом циклотронного резонанса.

При одновременном воздействии электрического и магнитного полей скорость электрона уже не будет постоянной, и в общем случае он будет двигаться по весьма сложной траектории; поэтому мы остановимся лишь на относительно простых примерах, имеющих для нас наибольшее значение. Если электрическое и магнитное поля параллельны, то электрон движется так же, как при отсутствии электрического поля — по винтовой линии, с той лишь разницей, что его скорость вращения, как и в первом случае ($E = 0$), остается постоянной ($v_{\perp} = \text{const}$), а скорость поступательного движения (v_{\parallel}) все время возрастает, поэтому соответственно возрастает и шаг винта. (Можно представить траекторию электрона как параболу, навитую на цилиндр.)

Если начальная скорость электрона равна нулю или параллельна магнитному полю, то эта винтовая линия (как и при $E = 0$) вырождается в прямую.

Этот сравнительно простой случай ($E \parallel H$) играет роль в гальвано- и термомагнитных явлениях лишь при наличии анизотропии эффективной массы или времени релаксации. При этих условиях он становится весьма сложным и выходит за пределы нашего рассмотрения. Поэтому мы перейдем к наиболее важному для нас случаю скрещенных (перпендикулярных) электрических и магнитных полей.

Рассмотрим эту ситуацию сначала качественно, причем положим для простоты начальную скорость электрона v_0 равной нулю, следовательно, в начальный момент и $F_H = 0$. Представим себе, что электрическое поле направлено по оси y (рис. 7.1), а магнитное — перпендикулярно плоскости чертежа. Тогда под действием электрического поля будет создаваться ускорение

$$\omega_y = eE_y/m,$$

и в первый момент электрон начнет двигаться равноускоренно вдоль оси y . Однако сразу же при этом возникнет магнитная сила

$$F_H = evH/c,$$

которая начнет отклонять электрон от первоначального направления движения.

В первый момент магнитная сила $\mathbf{F}_H \perp \mathbf{F}_E$ и поэтому не будет влиять на ускорение электрона, однако по мере возрастания скорости электрона будет возрастать и по мере искривления траектории, оставаясь перпендикулярной к скорости, будет раскладываться на две составляющие, одна из которых $(F_H)_y = ev_x H/c$ будет создавать ускорение, вычитающееся из ускорения электрического поля. Поэтому с некоторого момента T составляющая

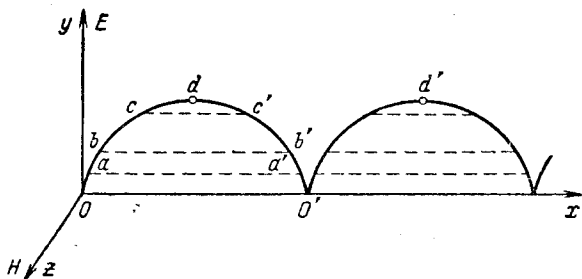


Рис. 7.1. Траектория движения электрона в скрещенных электрическом и магнитном полях.

скорости электрона по оси y начнет уменьшаться и в какой-то момент обратится в нуль (точка d на рис. 7.1). После этой точки электрон начнет двигаться по оси y в обратном направлении, приближаясь к оси x . При этом надо иметь в виду, что полная магнитная сила все время остается перпендикулярной полной скорости электрона и поэтому не может менять эту скорость, равную $v = \sqrt{v_x^2 + v_y^2}$, а уменьшая одну составляющую, одновременно соответствующим образом увеличивает другую. Поэтому в каждый данный момент скорость электрона, как это следует из закона сохранения энергии, будет определяться работой сил электрического поля

$$\frac{mv^2}{2} - \frac{mv_0^2}{2} = e(\varphi_0 - \varphi) \quad (7.15)$$

или в рассматриваемом нами случае

$$\frac{mv^2}{2} = eEy. \quad (7.16)$$

Из (7.16) непосредственно следует, что скорости электрона в точках a и a' , b и b' и c и c' (т. е. в точках, одина-

ково удаленных от оси x) будут одинаковы; следовательно, в точке o' скорость опять обратится в нуль, и после этого весь участок траектории $oabb'a'o'$ повторится, снова сдвинутый на величину oo' по оси x .

Для того чтобы найти траекторию электрона в аналитической форме, проинтегрируем его уравнение движения:

$$m\mathbf{w} = e\mathbf{E} + \frac{e[\mathbf{vH}]}{c} \quad (7.17)$$

или в координатной форме

$$m\ddot{x} = \frac{eyH}{c}, \quad (7.18)$$

$$m\ddot{y} = eE - \frac{eH\dot{x}}{c}. \quad (7.19)$$

Обозначим $eE/m = \omega$ и $eH/cm = \omega = \frac{1}{T}$. Тогда (7.18), (7.19) перепишутся в виде

$$\ddot{x} = \omega\dot{y}, \quad (7.20)$$

$$\ddot{y} = \omega - \omega\dot{x}. \quad (7.21)$$

Интегрируя эти уравнения при начальных условиях $t=0$, $\dot{x}=0$, $\dot{y}=0$, получаем:

$$\dot{x} = \omega T \left(1 - \cos \frac{t}{T}\right), \quad (7.22)$$

$$\dot{y} = \omega T \sin \frac{t}{T}, \quad (7.23)$$

$$x = \omega T^2 \left(\frac{t}{T} - \sin \frac{t}{T}\right), \quad (7.24)$$

$$y = \omega T^2 \left(1 - \cos \frac{t}{T}\right). \quad (7.25)$$

Уравнения (7.24) и (7.25) есть уравнения циклоиды (кривой, которую описывает точка, находящаяся на ободе катящегося колеса) в параметрической форме. Согласно (7.22) и (7.23) средняя скорость поступательного движения $v_y = 0$, $v_x = \omega T$, угловая скорость вращения $\omega = eH/mc$ и радиус

$$r = \frac{m\omega c}{eH}. \quad (7.26)$$

Мы предположили, что $v_0 = 0$; если бы мы не делали этого ограничения, то в зависимости от направления начальной скорости получили бы укороченную или удлиненную циклоиду.

Таким образом, полученное решение подтверждает тот парадоксальный на первый взгляд вывод, к которому мы пришли при качественном рассмотрении движения электрона: при наличии электрического поля на вращательное движение электрона накладывается равномерное поступательное, но в направлении, перпендикулярном к электрическому полю.

Прежде чем перейти непосредственно к изучению гальвано- и термомагнитных явлений, напомним классификацию слабых и сильных магнитных полей. Как уже упоминалось в гл. 1, слабыми называют такие магнитные поля, в которых радиус кривизны r [см. (7.26)] траектории электрона много больше его длины свободного пробега l :

$$r \gg l. \quad (7.27)$$

При этих условиях действие магнитного поля на электрон за время его свободного пробега будет относительно слабым; оно выразится в том, что траектория электрона незначительно искривится и направление его движения отклонится от первоначального на небольшой угол φ :

$$\varphi = \frac{l}{r}. \quad (7.28)$$

Но подвижность электрона u выражается через его длину свободного пробега формулой

$$u = \frac{e}{m} \frac{l}{v}.$$

Подставляя в (7.28) значение r из (7.26), получаем

$$\varphi = \frac{e}{m} \frac{l}{v} \frac{H}{c} \quad (7.29)$$

или

$$\varphi = \frac{uH}{c}. \quad (7.30)$$

Следовательно, неравенство (7.27), положенное нами в основу определения слабого поля, может быть заменено

эквивалентным $\varphi \ll 1$ или

$$\frac{uH}{c} \ll 1. \quad (7.31)$$

Последнее определение слабого поля (7.31) является наиболее удобным, так как в него входят величины u и H , непосредственно измеряемые на опыте.

Сильными магнитными полями мы называем такие поля, в которых выполняется неравенство, противоположное (7.31) или (7.27):

$$r \ll l \quad (7.32)$$

или

$$\frac{uH}{c} \gg 1. \quad (7.33)$$

В этом случае характер движения электрона существенно изменяется: в промежутке между столкновениями он уже движется не по почти прямолинейной траектории (как в случае слабых магнитных полей), а проходит ряд циклов либо винтовой линии, либо циклоиды, либо еще более сложной траектории.

Из сказанного выше следует, что деление на слабые и сильные поля не является абсолютным: при одном и том же значении поля для одного материала (с низкой подвижностью) может выполняться неравенство (7.31), а для другого (с большой подвижностью) — противоположное неравенство (7.33). Более того, для одного и того же материала при высоких температурах данное поле может быть слабым, а при низких температурах — сильным. Так, например, для чистого германия подвижность электронов составляет при комнатной температуре $u \approx \approx 3000 \text{ см}^2/\text{в} \cdot \text{сек} \cdot 300 = 9 \cdot 10^5 \text{ CGS (E)}$. Поле напряженностью 10^4 эрст будет приблизительно удовлетворять неравенству (7.31):

$$\frac{uH}{c} = \frac{9 \cdot 10^5 \cdot 10^4}{3 \cdot 10^{10}} < 1. \quad (7.34)$$

При температуре 10° К подвижность для того же германия будет $10^4 - 10^5 \text{ см}^2/\text{в} \cdot \text{сек}$, и то же самое магнитное поле в 10^4 эрст мы должны будем считать сильным.

В дырочном германии имеются две зоны — зона легких и тяжелых дырок, подвижности которых отличаются при комнатной температуре примерно в 10 раз. Поэтому

одно и то же поле может быть слабым для одного сорта дырок и сильным для другого.

Наряду с приведенными выше определениями сильных магнитных полей существует другое, основанное на изменениях, вносимых магнитным полем в спектр электрона (в то время как приведенное выше определение основано на кинетических эффектах).

Можно доказать строго (а мы здесь ограничимся качественными соображениями), что при наличии магнитного поля энергетический спектр электрона становится частично дискретным. Действительно, для свободного электрона кинетическая энергия

$$\epsilon = \frac{\hbar^2}{2m_0} (k_x^2 + k_y^2 + k_z^2), \quad (7.35)$$

где волновые векторы k_x , k_y и k_z могут принимать любые значения.

Представим теперь, что электрон находится в магнитном поле, направленном по оси z . Тогда согласно сказанному выше электрон будет в плоскости xy совершать вращательное движение с частотой $\omega = eH/mc$, которое можно представить как сумму двух колебательных движений. Но, как мы видели выше, осциллятор может иметь не любые значения энергии, а лишь квантованные, кратные $\hbar\omega$, это же следует отнести и к движению электрона в плоскости xy . Следовательно, при наличии магнитного поля энергия электрона может принимать значения

$$\epsilon = \hbar\omega \left(n + \frac{1}{2} \right) + \frac{\hbar^2 k_z^2}{2m}. \quad (7.36)$$

(Составляющая кинетической энергии по оси z может принимать по-прежнему любые значения.)

Такое радикальное изменение спектра сказывается на всех свойствах электронного газа, особенно четко это проявляется при вырождении. Так, если мы будем постепенно увеличивать магнитное поле (а следовательно, и ω) так, что через уровень Ферми будут последовательно проходить сгущения и разрежения плотности состояний, все свойства электронного газа, обусловленные плотностью состояний на поверхности Ферми, будут меняться периодически. Точно так же будут осциллировать и другие кинетические коэффициенты.

Целый ряд этих явлений получил свое название по имени открывших их ученых. Так, осцилляция магнитной восприимчивости названа эффектом де Гааза — Ван-Альфена, осцилляция электропроводности — эффектом Шубникова — де Гааза.

Теория для таких так называемых «квантовых полей» должна быть также радикально перестроена, так как кинетическое уравнение Больцмана в данном случае неприменимо. В самом деле, кинетическое уравнение (точнее, его левая часть) подразумевает непрерывное перемещение (дрейф) частиц в фазовом пространстве, а это невозможно при дискретных уровнях энергии и ограниченной (окружностью) траектории в обычном пространстве*).

Энергетический спектр электрона имеет дискретный вид для движения в плоскости xy до тех пор, пока средняя тепловая энергия меньше, чем расстояние $h\omega$ между дискретными уровнями, называемыми уровнями Ландау, т. е. до тех пор, пока

$$kT \ll h\omega = 2\mu H,$$

где $\mu = eh/2mc$ — магнетон Бора для электрона с эффективной массой m .

Если имеет место обратное неравенство $kT \gg h\omega$, то тепловое движение размывает всю картину и спектр электрона становится таким же, как в отсутствие магнитного поля.

Таким образом, мы имеем два критерия сильного поля: классический (7.33)

$$\frac{uH}{c} \gg 1 \quad (7.37)$$

и квантовый (7.37)

$$\frac{h\omega}{kT} \gg 1. \quad (7.37a)$$

Эти критерии неэквивалентны друг другу и по смыслу и по величине критического поля; первый (7.33) лишь означает, что траектория электрона претерпевает радикальные изменения, а второй (7.37) указывает на радикаль-

*) В действительности дискретное изменение энергии есть следствие ограниченности траектории. Условие стабильности любой замкнутой орбиты требует, чтобы на ней поместилось целое число волн: $2\pi r = n\lambda$, отсюда вытекают дискретные значения импульса и энергии.

ное изменение энергетического спектра *). Обычно критерию (7.37а) соответствуют более сильные поля или низкие температуры.

Объем и характер этой книги не позволяют остановиться более подробно на явлениях в квантующих магнитных полях; следующие разделы будут посвящены гальваномагнитным явлениям в слабых магнитных полях, а затем мы коротко остановимся на явлениях в сильных полях в классическом смысле.

Гальваномагнитные явления и соответствующие коэффициенты различают также по условиям теплообмена с окружающей средой; если теплообмен отсутствует, то коэффициенты называют адиабатическими, если образец поддерживается при постоянной температуре, то изотермическими. Мы здесь рассмотрим изотермические эффекты, так как выражения для соответствующих коэффициентов в этом случае значительно проще.

Все гальвано- и термомагнитные явления в слабых полях могут быть проанализированы строго и последовательно на основе решения кинетического уравнения с учетом магнитных членов (которые мы отбросили вначале). Мы начнем, однако, с качественного рассмотрения картины, а затем коротко воспроизведем математический анализ.

7.2. ЭФФЕКТ ХОЛЛА И ИЗМЕНЕНИЕ СОПРОТИВЛЕНИЯ В МАГНИТНОМ ПОЛЕ

В первой главе мы получили приближенное выражение для э. д. с. Холла, исходя из тех соображений, что в стационарных условиях «холловская сила» eE_y должна уравновешивать магнитную силу $e\mathbf{v}H/c$:

$$eE_y = \frac{ev_x H}{c}. \quad (7.38)$$

В действительности этот ход рассуждений неправилен по крайней мере по двум причинам:

1) равенство (7.38) не может выполняться одновременно для всех электронов, имеющих скорости, различные и по величине и по направлению;

*) Точнее, для того чтобы спектр стал дискретным, необходимо выполнение обоих условий, так как если условие (7.37) не будет выполняться, то квантованный уровень будет размываться за счет столкновений.

2) в действительности стационарное состояние наступает не тогда, когда магнитная сила уравнивает электрическую для каждого электрона, что вообще не может быть, а тогда, когда перестает накапливаться заряд на боковых гранях образца, т. е. когда ток, создаваемый холловским полем, компенсирует ток на боковую грань, создаваемый магнитным полем. Иными словами, условием стационарности является не равенство сил, а равенство токов.

Если все электроны обладают одной и той же подвижностью, то эти условия эквивалентны*), но чем больше различие в подвижности различных групп электронов, тем сильнее различаются эти два условия. Мы рассмотрим это на простейшем примере двух групп электронов с концентрациями n_1 и n_2 и подвижностями u_1 и u_2 .

Тогда ток, создаваемый первой группой электронов, $j_x^1 = en_1 u_1 E_x$ под действием магнитного поля отклоняется на угол $\varphi_1 = u_1 H/c$ и создает составляющую по оси y :

$$j_y^1 = j_x^1 \varphi_1 = \frac{e u_1^2 n_1 H}{c} E_x; \quad (7.39)$$

точно так же можно вычислить ток вдоль оси y , создаваемый второй группой электронов:

$$j_y^2 = j_x^2 \varphi_2 = \frac{e u_2^2 n_2 H}{c} E_x, \quad (7.40)$$

и полный ток, создаваемый вдоль оси y холловским полем (E_y):

$$j_x = E_y e (n_1 u_1 + n_2 u_2). \quad (7.41)$$

В соответствии со сказанным выше, для того чтобы заряды не накапливались, необходимо выполнение условия

$$j_x = j_y^1 + j_y^2$$

*) Это имеет место в двух случаях: а) при простой зонной структуре и полном вырождении, при этом скорости и подвижности всех электронов на поверхности Ферми одинаковы; б) при рассеянии носителей на оптических колебаниях ионной решетки при низких температурах, при этом $r = 1/2$ и подвижность $u = (e/m)$, $v/l \sim \sim \epsilon^{1/2}/\epsilon^{1/2}$ не зависит от скорости электрона. В этих двух случаях угол φ отклонения в магнитном поле для всех электронов одинаков и холловское поле может одновременно «выправить» все траектории.

или согласно (7.39), (7.40) и (7.41)

$$E_y (n_1 u_1 + n_2 u_2) = \frac{H}{c} (u_1^2 n_1 + u_2^2 n_2) E_x. \quad (7.42)$$

В правой и левой части выражения (7.42) могут играть основную роль различные группы электронов: «магнитный ток» [т. е. правая часть (7.42)] создается электронами, подвижность которых больше, а компенсируется теми, концентрация которых больше. Учитывая, что

$$E_x = \frac{j_x}{e (n_1 u_1 + n_2 u_2)}, \quad (7.43)$$

из (7.42) получаем

$$E_y = \frac{j_x H}{e c} \frac{n_1 u_1^2 + n_2 u_2^2}{(n_1 u_1 + n_2 u_2)^2}. \quad (7.44) *$$

При $u_1 = u_2$ выражение (7.44) совпадает с полученным без учета дисперсии:

$$E_y = j_x H \frac{1}{e (n_1 + n_2) c}. \quad (7.45)$$

Чем больше различие подвижностей, тем сильнее (7.44) отличается от (7.45). При наличии двух групп носителей с сильно различающейся массой это различие (даже при полном вырождении) может быть очень велико. Если же считать, что массы всех носителей одинаковы, то в (7.44) отношение подвижностей можно заменить отношением времен релаксации:

$$E_y = \frac{j_x H}{e n c} \frac{\overline{\tau^2}}{(\overline{\tau})^2}, \quad (7.46)$$

где введены обозначения

$$\overline{\tau^2} = \frac{\sum \tau_i^2 n_i}{\sum n_i},$$

и

$$\overline{\tau} = \frac{\sum \tau_i n_i}{\sum n_i}. \quad (7.47)$$

Выражение (7.47) без труда обобщается на любое число групп носителей с различной подвижностью. Его также нетрудно обобщить на случай непрерывной зависимости времени релаксации от энергии, заменив суммы интегралами. Мы, однако, сделаем это более строго, решив

*) Если подвижность одной группы электронов много больше (например, $u_1 \gg u_2$), а концентрации одного порядка, то согласно (7.44) $E_y = j_x H / e n_1 c$, т. е. эффект Холла будет давать концентрацию электронов, подвижность которых больше.

для этой цели кинетическое уравнение при наличии магнитного поля.

Как мы уже упоминали, слабое магнитное поле не действует на равновесную функцию распределения. Качественно это следует из того, что равновесная функция обладает шаровой симметрией, магнитная же сила перпендикулярна скорости электрона и в этом случае не меняет распределение электронов по скоростям. Количественно это выражается в том, что если мы подставим в магнитный член кинетического уравнения (5.49а) вместо f величину f_0 , то он обратится в нуль:

$$\left(\frac{\partial f_0}{\partial t} \right)_{\text{дрейф магн}} = e \frac{\partial f_0}{\partial \epsilon} m v \frac{[\mathbf{vH}]}{c} = 0.$$

Иначе обстоит дело при одновременном воздействии электрического и магнитного полей; как уже много раз упоминалось, электрическое поле смещает и деформирует функцию распределения (см. гл. 5).

Магнитное поле поворачивает эту деформированную функцию на угол

$$\varphi = \omega \tau = \frac{eH}{mc} \tau \quad (7.48)$$

и опять несколько деформирует, так как угол φ зависит от времени τ , различного для электронов с различной энергией. Мы не можем теперь ограничиться рассмотрением одномерного случая, так как у функции $f_1 = f - f_0$ будут составляющие по оси v_x и v_y ; простейший вид, в котором мы ее можем искать, следовательно, будет

$$f_1 = \chi_1 v_x + \chi_2 v_y, \quad (7.49)$$

где χ_1 и χ_2 — неизвестные функции энергии.

В соответствии с (5.66)

$$-f_1 = \tau \left(\frac{\partial f}{\partial x} v_x + \frac{\partial f}{\partial y} v_y + \frac{\partial f}{\partial v_x} \omega_x + \frac{\partial f}{\partial v_y} \omega_y \right). \quad (7.50)$$

Ограничимся рассмотрением изотермических эффектов в однородных образцах, в этом случае

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0, \quad (7.51)$$

$$\omega_x = \frac{1}{m} \left\{ e\mathbf{E} + \frac{e}{c} [\mathbf{vH}] \right\}_x = \frac{e}{m} \left[E_x + \frac{1}{c} v_y H \right], \quad (7.52)$$

$$\omega_y = \frac{1}{m} \left\{ e\mathbf{E} + \frac{e}{c} [\mathbf{vH}] \right\}_y = \frac{e}{m} \left[E_y - \frac{1}{c} v_x H \right].$$

Подставив выражения (7.49), (7.51) и (7.52) в (7.50), заменим, как мы это делали раньше, во всех членах, не содержащих магнитного поля, f на f_0 и учтем также, что $\partial f_0 / \partial v_x = (\partial f_0 / \partial \varepsilon) m v_x$, $\partial f_0 / \partial v_y = (\partial f_0 / \partial \varepsilon) m v_y$. После этого выражение (7.52) примет вид

$$-\frac{1}{\tau} (v_x \chi_1 + v_y \chi_2) = \frac{\partial f_0}{\partial \varepsilon} e E_x v_x + \frac{\partial f_0}{\partial \varepsilon} e E_y v_y - \frac{e}{m} \frac{H}{c} \chi_2 v_x + \frac{eH}{mc} \chi_1 v_y. \quad (7.53)$$

Равенство (7.53) должно быть тождественным, т. е. удовлетворяться при любых значениях скоростей*). Для этого коэффициенты при v_x и v_y в левой и правой части (7.53) должны быть одинаковы. Из этого условия находим

$$\chi_1 - \omega \tau \chi_2 = -\tau e E_x \frac{\partial f_0}{\partial \varepsilon}, \quad (7.54)$$

$$\chi_2 + \omega \tau \chi_1 = -\tau e E_y \frac{\partial f_0}{\partial \varepsilon}. \quad (7.55)$$

Решая совместно (7.54) и (7.55) относительно χ_1 и χ_2 , после несложных преобразований получаем

$$\chi_1 = -\frac{\partial f_0}{\partial \varepsilon} e \tau \frac{E_x + \varphi E_y}{1 + \varphi^2}, \quad (7.56)$$

$$\chi_2 = -\frac{\partial f_0}{\partial \varepsilon} e \tau \frac{E_y - \varphi E_x}{1 + \varphi^2}, \quad (7.57)$$

где $\varphi = \omega \tau$.

Выражения (7.56) и (7.57) наглядно показывают, что магнитное поле поворачивает неравновесную добавку к функции распределения на угол φ и укорачивает ее в $(1 + \varphi^2)$ раз (за счет сокращения эффективной длины свободного пробега).

Зная χ_1 и χ_2 , можно получить f_1 и выражения для плотности токов j_x и j_y :

$$j_x = \frac{2}{h^3} \int e v_x f_1 dg, \quad j_y = \frac{2}{h^3} \int e v_y f_1 dg, \quad (7.58)$$

где dg — элемент объема в пространстве импульсов, и интегрирование должно быть проведено по всему пространству.

*) Это следует также из того, что члены, содержащие в виде множителя v_x и v_y , перпендикулярны друг другу в пространстве скоростей.

Учитывая это, можно сразу же отбросить все члены, содержащие v_x или v_y в первой степени, так как они исчезнут при интегрировании по соответствующим импульсам в пределах от $-\infty$ до $+\infty$; учтем также, что усреднение по углам дает

$$\overline{v_x^2} = \overline{v_y^2} = \frac{1}{3} \overline{v^2}. \quad (7.59)$$

На основании сделанных выше замечаний получим

$$j_x = \frac{2e}{h^3} \int v_x (v_x \chi_1 + v_y \chi_2) dg = \frac{2e}{3h^3} \int v^2 \chi_1 dg, \quad (7.60)$$

$$j_y = \frac{2e}{h^3} \int v_y (v_x \chi_1 + v_y \chi_2) dg = \frac{2e}{h^3} \int v^2 \chi_2 dg \quad (7.60a)$$

и, подставляя в (7.60) и (7.60a) выражения для χ_1 и χ_2 , получаем

$$\begin{aligned} j_x &= \frac{2}{3} \frac{e^2}{h^3} \int \frac{\tau v^2 \frac{\partial f_0}{\partial \varepsilon} (E_x + \tau \omega E_y)}{1 + \omega^2 \tau^2} dg = \\ &= \frac{1}{3} e^2 [L_1 E_x + \omega L_0 E_y], \end{aligned} \quad (7.61)$$

где введены обозначения:

$$L_1 = \frac{2}{h^3} \int \frac{\tau v^2}{1 + \omega^2 \tau^2} \frac{\partial f_0}{\partial \varepsilon} dg = \frac{2}{h^3} \int \frac{v l}{1 + \omega^2 \tau^2} \frac{\partial f_0}{\partial \varepsilon} dg \quad (7.62)$$

и

$$L_0 = \frac{2}{h^3} \int \frac{\tau^2 v^2}{1 + \omega^2 \tau^2} \frac{\partial f_0}{\partial \varepsilon} dg = \frac{2}{h^3} \int \frac{l^2}{1 + \omega^2 \tau^2} \frac{\partial f_0}{\partial \varepsilon} dg. \quad (7.63)$$

Точно так же можно получить выражение для j_y :

$$j_y = \frac{1}{3} e^2 [L_1 E_y - \omega L_0 E_x]. \quad (7.64)$$

Полагая теперь $j_y = 0$, находим холловское поле:

$$E_y = \frac{\omega L_0}{L_1} E_x, \quad (7.65)$$

и, подставив E_y из (7.65) в (7.61), найдем выражение для плотности тока

$$j_x = \frac{e^2}{3} \left[\frac{L_1^2 + \omega^2 L_0^2}{L_1} \right] E_x, \quad (7.66)$$

откуда электропроводность в магнитном поле

$$\sigma(H) = \frac{j_x}{E_x} = \frac{e^2}{3} \left[\frac{L_1^2 + \omega^2 L_0^2}{L_1} \right] \quad (7.67)$$

постоянная Холла

$$R_x = \frac{E_y}{i_x H} = \frac{\omega L_0}{L_1} \frac{E_x}{H} \frac{3}{e^2} \left[\frac{L_1}{L_1^2 + \omega^2 L_0^2} \right] \frac{1}{E_x} = \frac{3}{enc} \frac{L_0}{L_1^2 + \omega^2 L_0^2}. \quad (7.67a)$$

Рассмотрим на основе полученных выражений частный случай: слабые поля $\varphi \ll 1$, параболическая зона и применимость статистики Максвелла.

Пренебрегая $\varphi = \omega\tau$ в знаменателе (7.62) и (7.63), полагая, как мы делали раньше,

$$l = a\epsilon^r \quad \text{и} \quad f_0 = e^{\frac{\mu - \epsilon}{kT}}$$

и подставляя в (7.62) и (7.63) $dg = \frac{4\pi p^2 dp}{h^3} = \frac{4\pi (2m)^{3/2} e^{1/2} d\epsilon}{h^3}$ получаем

$$L_1 = \frac{4a^r \pi (2m)^{3/2}}{h^3} e^{\mu^*} (kT)^{r+1} \Gamma(r+2) \quad (7.68)$$

и

$$L_0 = \frac{4a^{2r} \pi (2m)^{3/2}}{h^3} (kT)^{2r+1/2} e^{\mu^*} \Gamma\left(2r + \frac{3}{2}\right), \quad (7.69)$$

где $\Gamma(r+2)$ и $\Gamma\left(2r + \frac{3}{2}\right)$ — гамма-функции.

Напомним, что

$$\int_0^\infty x^n e^{-\alpha x^2} dx = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2\alpha \frac{n+1}{2}} \quad \text{при } \alpha > 0 \text{ и } n > -1, \quad (7.70)$$

при n целом и четном ($n = 2k$):

$$\Gamma(n) = (n-1)!, \quad (7.71)$$

при n нечетном ($n = 2k + 1$): $\Gamma(n) = (n-1)! \sqrt{\pi}$,

при любом n

$$\Gamma(n) = (n-1) \Gamma(n-1) \quad (7.72)$$

и, наконец,

$$\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}. \quad (7.73)$$

Согласно (7.67a), (7.68) и (7.69)

$$R_x = \frac{3 \sqrt{\pi}}{4enc} \frac{\Gamma\left(2r + \frac{3}{2}\right)}{\Gamma^2(r+2)}. \quad (7.74)$$

С помощью соотношений (7.71), (7.72), (7.73) и (7.74) нетрудно получить выражения для постоянной Холла при конкретных механизмах рассеяния, приведенных в гл. 1.

Совершенно аналогично могут быть получены из (7.66) выражения для изменения электропроводности в магнитном поле:

$$\frac{\Delta\sigma(H)}{\sigma_0} = \frac{\sigma(H) - \sigma_0}{\sigma_0} = FH^2, \quad (7.75)$$

где

$$F = R^2\sigma_0^2 \left[\frac{B_r^2}{A_r^2} - 1 \right]; \quad (7.76)$$

$$A_r = \frac{\Gamma\left(2r + \frac{3}{2}\right)}{\Gamma^2(r+2)}; \quad B_r = \frac{9}{16\pi} \frac{\Gamma(3r+1)}{\Gamma^3(r+2)}. \quad (7.77)$$

Можно легко получить выражения, совершенно идентичные (7.57) и (7.56) для функции распределения дырок, вычислить дырочные токи j_x и j_y и затем рассмотреть собственную проводимость. При этом мы должны будем приравнять к нулю суммарный дырочный и электронный ток вдоль оси y и отсюда получить связь между E_x и E_y . Мы здесь приведем вывод, аналогичный сделанному выше, но менее строгий и более наглядный.

Ток дырок по оси x $j_p^x = ev_p = e\mu_p E_x$ создает соответствующий ток по оси y :

$$j_p^y = j_p^x \frac{u_p H}{c} = e\mu_p^2 \frac{H}{c} E_x; \quad (7.78)$$

аналогично можно вычислять электронный ток:

$$j_n^y = enu_n^2 \frac{H}{c} E_x. \quad (7.79)$$

Так как j_p^y и j_n^y текут в одном направлении, то результирующий ток по оси y равен их разности:

$$j_1^y = j_p^y - j_n^y = \frac{eHE_x}{c} (u_p^2 \rho - u_n^2 n). \quad (7.80)$$

Для того чтобы на одной из боковых граней не накапливались заряды, холловское поле E_y должно создать ток, равный и противоположно направленный j_1^y . Таким образом, условие стационарности будет иметь вид

$$j_1^y = E_y \sigma = E_y e (u_p \rho + u_n n), \quad (7.81)$$

откуда

$$E_y = \frac{H}{c} E_x \frac{(u_p^2 \rho - u_n^2 n)}{u_p \rho + u_n n}. \quad (7.81a)$$

При таких условиях полный ток на боковой поверхности образца равен нулю, так как ток дырок равен току электронов по величине, но противоположен по знаку. Нетрудно подсчитать каждый из этих токов в отдельности:

$$j_p^y = j_n^y = \frac{1}{2} \frac{e H E_x}{c} (u_p^2 \rho + u_n^2 n). \quad (7.82)$$

В этом случае также на одной грани будет повышенная концентрация носителей и рекомбинация будет преобладать над диссоциацией, что будет сопровождаться выделением тепла, на противоположной грани будет пониженная концентрация носителей и диссоциация будет преобладать над рекомбинацией, что вызовет поглощение тепла. Этот процесс в известной мере аналогичен биполярной диффузии при наличии градиента температур, но кардинально отличается тем, что сам создает разность температур, которая получила название эффекта Эттингсгаузена (см. следующий параграф).

Выразив E_x через j_x , получим выражение постоянной Холла:

$$R_x = \frac{E_y}{j_x H} = \frac{1}{ec} \frac{u_p^2 \rho - u_n^2 n}{(u_p \rho + u_n n)^2}. \quad (7.83)$$

Более строгий расчет ~~дает выражение~~

$$R_x = \frac{A}{ec} \frac{u_p^2 \rho - u_n^2 n}{(u_p \rho + u_n n)^2}. \quad (7.84)$$

Постоянная A , так же как и в случае одного знака носителей, определяется механизмом рассеяния, $A = 3\pi/8$ в атомной решетке и т. д.

Нетрудно убедиться, что: а) в случае, если $n = 0$ или $p = 0$, (7.84) переходит в (7.74) и б) если $n = p$ и $u_p = u_n$, постоянная Холла равна нулю, а следовательно, и э. д. с. Холла равна нулю.

Согласно (7.84) в области собственной проводимости знак э. д. с. Холла соответствует знаку носителей, подвижность которых больше, т. е. обычно электронов, поэтому в примесном дырочном проводнике э. д. с. Холла при переходе к собственной проводимости обычно проходит через нуль и меняет знак.

В случае смешанной проводимости выражение для электропроводности имеет вид

$$\sigma = e(u_p p + u_n n). \quad (7.85)$$

Как видно из (7.84) и (7.85), в этом случае одновременное измерение постоянной Холла и электропроводности не дает достаточных данных для определения подвижности и концентрации электронов и дырок, так как мы имеем два уравнения с четырьмя неизвестными. В этом случае можно выйти из положения, найдя значения подвижности дырок или электронов экстраполяцией из области примесной проводимости, а также воспользовавшись соотношением $n - p = N_d$. Задача может быть также решена, если к (7.84) и (7.85) добавить еще независимые уравнения, т. е. одновременно измерить другие эффекты.

За последние годы эффект Холла нашел широкое применение в технике: для измерения постоянных и переменных магнитных полей и «бесконтактного» измерения токов по их полям, в счетно-решающих устройствах (для умножения, деления, сложения, и вычитания, запоминающих, считывающих и записывающих устройств), для детектирования, усиления, модуляции, измерения мощности, измерения разности фаз, анализа спектров и решения множества других задач радиотехники и электротехники [29].

7.3. ЭФФЕКТ ЭТТИНГСГАУЗЕНА

Примесная проводимость. Этот случай мы рассмотрели достаточно подробно на стр. 91—92 (гл. 1). Здесь мы напомним лишь выражение для поперечного градиента температуры:

$$\frac{\partial T}{\partial y} = \frac{kT}{e} \frac{2r-1}{2} \frac{R_X \sigma}{\kappa_p + \kappa_{эл}} j_x H, \quad (7.86)$$

где κ_p и $\kappa_{эл}$ — соответственно теплопроводность решетки и электронного газа.

Собственная проводимость. Этот случай мы частично уже рассмотрели при анализе эффекта Холла в собственном полупроводнике. В случае собственной или смешанной проводимости, т. е. в переходной области от примесной к собственной, холловское поле вообще не образуется, если потоки дырок и электронов на боковую грань одинаковы, или образуется таким образом, чтобы уравнивать

эти потоки. В обоих случаях в направлении оси y , т. е. с одной грани на другую, текут равные электронные и дырочные токи, определяемые соотношением (7.82); весь этот ток генерируется на одной грани и рекомбинирует на другой, и, таким образом, можно подсчитать тепло, переносимое электронами и дырками в этом случае. Эффект Эттингсгаузена может, так же как и эффект Пельтье, быть использован для перекачки тепла, т. е. охлаждения, термостатирования, кондиционирования воздуха и т. д.

Эффекты Холла и Эттингсгаузена называются поперечными гальваномагнитными эффектами: разность потенциалов и разность температуры появляются в направлении, перпендикулярном направлению электрического тока.

7.4. ГАЛЬВАНОМАГНИТНЫЕ ЯВЛЕНИЯ В СИЛЬНЫХ МАГНИТНЫХ ПОЛЯХ

В заключение рассмотрим качественно эффект Холла и изменение сопротивления в сильных, в классическом смысле, магнитных полях, т. е. при

$$\frac{uH}{c} \gg 1. \quad (7.87)$$

Согласно (1.124) в этом случае холловское поле будет много больше исходного, и в первом приближении мы можем считать, что все поле направлено по «оси Холла», которую мы примем за ось x . Так как период прецессии электрона много меньше времени свободного пробега, то в первом приближении столкновениями мы также можем пренебречь и считать, что электроны движутся по циклоидам в направлении, перпендикулярном электрическому полю, т. е. в направлении y^*). Таким образом, вырисовывается очень своеобразная картина — поле повернулось на 90° по отношению к первоначальному, но ток повернут на 90° по отношению к полю и поэтому течет в нужном направлении. Согласно (7.86) $v_y = (E_x/H) c$ и

$$j_y = ev_y n = \frac{eE_x cn}{H}. \quad (7.88)$$

По определению, постоянная Холла равна отношению э. д. с. Холла к плотности тока и магнитному полю.

*) Таким образом ось y будет теперь осью тока, а ось x — осью холловского поля; т. е. оси будут повернуты на 90° по отношению к их положению, рассмотренному ранее.

Согласно (7.88)

$$R_x = \frac{E_x}{Hj_y} = \frac{1}{enc}, \quad (7.89)$$

т. е. несмотря на то, что макроскопическая картина явлений кардинально изменилась, выражение для постоянной Холла осталось прежним.

В случае двух знаков носителей дырочный ток

$$j_p = \frac{epc}{H} E_x, \quad (7.90)$$

электронный ток

$$j_n = \frac{enc}{H} E_x. \quad (7.91)$$

При этом нетрудно убедиться, что оба тока направлены в одну и ту же сторону. Следовательно, результирующий ток

$$j_y = j_p - j_n = \frac{ec}{H} E_x (p - n) \quad (7.92)$$

и постоянная Холла

$$R_x = \frac{E_x}{j_y H} = \frac{1}{ec} \frac{1}{p - n}. \quad (7.93)$$

До сих пор мы совершенно не учитывали столкновений, считая, что все поле направлено по оси x , $E_y = 0$, а следовательно, удельная электропроводность $\sigma = j_y/E_y = \infty$. Для того чтобы получить конечную электропроводность, мы должны учесть столкновения. Сделаем это приближенно и получим поэтому приближенные формулы. Учтем, что поле по оси y не равно нулю, а согласно (1.124) $E_x = E_y \frac{uH}{c}$ (напомним, что оси теперь

повернуты на 90°), откуда

$$E_y = E_x \frac{c}{uH} = E_x \frac{T}{\tau}. \quad (7.94)^*$$

Следовательно,

$$\sigma_y = \frac{j_y}{E_y} = \frac{encE_x uH}{HE_x c} = enuH, \quad (7.95)$$

*) Соотношение (7.94) можно просто интерпретировать, если предположить, что после каждого столкновения электрон перекакивает на соседнюю циклоиду и $\Delta y = \Delta x \frac{T}{\tau}$, где $T = cm/eH$; все рассуждения здесь справедливы с точностью до коэффициента, поэтому множитель 2π в T мы опустим.

т. е. мы пришли к обычному выражению для электропроводности, и нам остается лишь выяснить, насколько подвижность в сильном магнитном поле меньше, чем в отсутствие поля. Если мы предположим, что подвижность электронов падает во столько же раз, во сколько увеличивается длина их пути, т. е. во сколько раз длина циклоиды больше отрезка oo' , то получим

$$\frac{\sigma_H}{\sigma_0} = \frac{\pi}{4}, \quad (7.96)$$

где σ_H обозначена электропроводность в сильном магнитном поле. Строгая теория дает

$$\frac{\sigma_H}{\sigma_0} = \frac{9\pi}{32}. \quad (7.97)$$

При наличии двух знаков носителей картина значительно усложняется. Как мы уже упоминали, токи дырок и электронов текут в одну и ту же сторону и результирующий ток по оси y

$$j_y = j_p - j_n = \frac{ec}{H} E_x (p - n). \quad (7.98)$$

Однако если учесть столкновения, то электронный и дырочный токи отклоняются в противоположные стороны и (так как заряды электронов и дырок противоположны) по оси будет протекать суммарный ток

$$j'_x = j_p^y \frac{T_p}{\tau_p} + j_n^y \frac{T_n}{\tau_n} = \frac{ec}{H} \left(\frac{pT_p}{\tau_p} + \frac{nT_n}{\tau_n} \right) E_x. \quad (7.99)$$

Поле, направленное вдоль оси y (E_y), будет создавать ток вдоль оси x :

$$j''_x = \frac{ec}{H} E_y (p - n). \quad (7.100)$$

В стационарных условиях ток j''_x должен компенсировать ток j'_x :

$$\frac{ec}{H} \left(\frac{pT_p}{\tau_p} + \frac{nT_n}{\tau_n} \right) E_x = \frac{ec}{H} E_y (p - n), \quad (7.100a)$$

откуда

$$E_y = \frac{1}{p - n} \left(\frac{pT_p}{\tau_p} + \frac{nT_n}{\tau_n} \right) E_x$$

или

$$E_y = \frac{cE_x}{eH(p - n)} \left[\frac{pm_p}{\tau_p} + \frac{nm_n}{\tau_n} \right]; \quad (7.101)$$

с другой стороны,

$$j_y = \frac{ec}{H} E_x (\rho - n). \quad (7.102)$$

Согласно (7.101) и (7.102) находим удельное сопротивление ρ :

$$\rho = \frac{E_y}{j_y} = \frac{1}{e^2 (\rho - n)^2} \left[\frac{\rho m_p}{\tau_p} + \frac{n m_n}{\tau_n} \right]. \quad (7.103)$$

Выражение (7.103) можно также преобразовать к виду

$$\rho = \frac{1}{(\rho - n)^2} \left[\frac{\rho^2}{\sigma_p} + \frac{n^2}{\sigma_n} \right]. \quad (7.104)$$

При наличии одного знака носителей ($\rho = 0$ или $n = 0$) выражение (7.104) переходит в (7.95), при наличии двух знаков носителей сопротивление значительно возрастает; это является следствием уменьшения тока вдоль оси y (который в этом случае является разностью двух токов) и увеличения напряжения E_y .

Однако и в этом случае, как и при наличии одного знака носителей, сопротивление в магнитном поле не возрастает безгранично, а асимптотически приближается к предельному значению, выражаемому (7.104).

Иначе обстоит дело в области собственной проводимости. Очевидно, что выражения (7.103) и (7.104) в этом случае неприменимы, так как при $\rho = n$ они дают нулевой ток и бесконечное сопротивление вне зависимости от магнитного поля.

Все дело в том, что холловское поле в этом случае не образуется, так как дырочный ток вдоль оси

$$j_p^x = \frac{ec}{H} \rho E_y \quad (7.105)$$

и электронный

$$j_n^x = \frac{ec}{H} n E_y \quad (7.106)$$

текут в одну сторону и компенсируют друг друга.

Однако под действием столкновений они отклоняются в противоположные стороны и создают суммарный ток вдоль оси y :

$$j_y = j_p^x \frac{T_p}{\tau_p} + j_n^x \frac{T_n}{\tau_n} = E_y \frac{c^2}{H^2} n \left(\frac{m_p}{\tau_p} + \frac{m_n}{\tau_n} \right) \quad (7.107)$$

и

$$\sigma = \frac{j_y}{E_y} = \frac{c^2}{H^2} n \left(\frac{m_p}{\tau_p} + \frac{m_n}{\tau_n} \right). \quad (7.108)$$

Таким образом, в этом случае электропроводность неограниченно убывает ($\sim H^{-2}$) и, как это на первый взгляд ни парадоксально, обратно пропорциональна времени релаксации.

В действительности в этом нет ничего удивительного, так как ток вдоль оси y в этом случае возникает только в результате столкновений.

7.5. ТЕРМОМАГНИТНЫЕ ЯВЛЕНИЯ

Рассмотрим примесный полупроводниковый образец, вдоль которого существует перепад температур (пока в отсутствие магнитного поля). Концентрация и скорости электронов на горячем конце больше, чем на холодном; поэтому поток электронов от горячего конца к холодному $j_{гор}$ больше, чем в обратном направлении $j_{хол}$, в результате чего возникает поле термо-э. д. с., тормозящее горячие электроны, ускоряющее холодные и уравнивающее эти потоки. Таким образом, в стационарном состоянии в разомкнутой цепи благодаря термо-э. д. с.

$$j_{гор} = j_{хол} = j_0. \quad (7.109)$$

Уменьшая энергию и поток горячих электронов и увеличивая энергию и поток холодных электронов, термо-э. д. с. уменьшает теплопроводность электронного газа [16], но, разумеется, не обращает ее в нуль — это противоречило бы второму началу термодинамики.

Таким образом, энергия электронов, идущих от горячего конца, остается больше и поток тепла вдоль образца будет равен

$$Q = \frac{1}{e} (j_{гор} \bar{\epsilon}_{гор} - j_{хол} \bar{\epsilon}_{хол}) = \frac{j_0}{e} (\bar{\epsilon}_{гор} - \bar{\epsilon}_{хол}) > 0. \quad (7.110)$$

Посмотрим теперь, что произойдет, если мы включим магнитное поле. Тогда электроны, движущиеся от горячего конца, будут отклоняться на одну боковую грань, и она будет нагреваться, а электроны, движущиеся от холодного конца, будут отклоняться на противоположную грань, и она будет охлаждаться. Это явление носит название эффекта Риги — Ледюка. Теория дает следующее выражение для поперечного градиента температуры в примесном невырожденном полупроводнике при простой

параболической зонной структуре:

$$\frac{\partial T}{\partial y} = \frac{\left(\frac{2r+1}{2}\right)^2 + \frac{3}{2}}{r+2} \frac{\kappa_{эл}}{\kappa_p + \kappa_{эл}} R_x \sigma \frac{\partial T}{\partial x}. \quad (7.111)$$

Как видно из (7.111), даже в этом простейшем случае выражение для поперечной разности температур весьма сложно. Поэтому, а также из-за трудности тепловых измерений исследования эффекта Риги — Ледюка не нашли широкого применения.

Перейдем теперь ко второму поперечному эффекту, так называемому эффекту Нернста — Эттингсгаузена, имеющему сейчас весьма большое значение в исследовании и применении полупроводников.

Как мы уже упоминали, потоки горячих и холодных электронов $\bar{j}_{гор}$ и $\bar{j}_{хол}$ одинаковы, но средние энергии в потоке $\bar{\epsilon}_{гор}$ и $\bar{\epsilon}_{хол}$ различны. Следовательно, горячие и холодные электроны будут отклоняться магнитным полем на разные углы, в результате чего в образце возникнет поперечный ток

$$j'_y = j_{гор} \bar{\Phi}_{гор} - j_{хол} \bar{\Phi}_{хол} = \frac{j_0}{T} [\tau(\bar{\epsilon}_{гор}) - \tau(\bar{\epsilon}_{хол})] \quad (7.112) *$$

и поперечная разность потенциалов E_y , которая будет возрастать до тех пор, пока обусловленный этой разностью потенциалов ток $j''_y = \sigma E_y$ не скомпенсирует j'_y .

Для простейшего случая, рассмотренного выше (примесный невырожденный полупроводник, простая зонная структура), выражение для поперечной э. д. с. Нернста — Эттингсгаузена имеет вид

$$E_y = \frac{2r-1}{2} \frac{A(r)}{e} \frac{k}{c} u H \frac{\partial T}{\partial x}, \quad (7.113)$$

где $A(r)$ — коэффициент, входящий в выражение для постоянной Холла.

Как видно из (7.113), эффект Нернста — Эттингсгаузена можно использовать для непосредственного измерения подвижности u (если известно r , или для определения r , если известно u). Его преимущество перед эффектом Холла заключается в том, что в последнем случае мы получаем подвижность косвенно через постоянную Холла и электро-

* Из (7.112) видно, что если время релаксации не зависит от энергии, то поперечная разность потенциалов не возникает.

проводность ($u \approx R\sigma/A$), и если электропроводность оказывается заниженной за счет плохо проводящих прослоек между зернами поликристалла, то и для подвижности получаются неправильные значения. В эффекте Нернста — Эттингсгаузена значение прослоек значительно меньше.

В области собственной проводимости поперечная термо-э. д. с. Нернста — Эттингсгаузена приобретает значительно большие значения. Действительно, в этом случае поле термо-э. д. с. либо совсем не образуется, либо образуется лишь для того, чтобы выравнять потоки электронов и дырок, и после этого равные потоки электронов и дырок генерируются на горячем конце и рекомбинируют на холодном (см. гл. 6 — теплопроводность за счет биполярной диффузии).

Магнитное поле разворачивает эти потоки на противоположные грани образца, и на одной накапливается положительный заряд, а на другой — отрицательный.

Таким образом, в то время как в примесной проводимости э. д. с. Нернста — Эттингсгаузена возникала как эффект второго порядка за счет зависимости времени свободного пробега от энергии, здесь эта э. д. с. образуется всегда и достигает весьма больших значений. В этом случае выражение для э. д. с. Нернста — Эттингсгаузена примет вид

$$E_y = \frac{k}{2ec} \frac{A}{\sigma^2} \left[(1 - 2r) (\sigma_n^2 u_n + \sigma_p^2 u_p) - \sigma_p \sigma_n (u_n + u_p) \left(6r + 7 + \frac{2\Delta \mathcal{E}_0}{kT} \right) \right] H \frac{\partial T}{\partial x}. \quad (7.114)$$

Еще больших значений достигает э. д. с. Нернста — Эттингсгаузена в области смешанной проводимости, когда неосновные носители только появляются. Действительно, термо-э. д. с. в примесном полупроводнике уменьшает разность скоростей основных носителей на горячем и холодном конце и, наоборот, увеличивает эту разность для неосновных носителей, поэтому э. д. с. Нернста — Эттингсгаузена, создаваемая неосновными носителями, бывает очень велика.

В заключение этой главы приведем выражения для основных кинетических коэффициентов в слабом и нулевом магнитных полях, для одной параболической зоны (табл. 3 была любезно предоставлена нам В. И. Кайда-новым).

Выражения для основных кинетических коэффициентов в слабом и нулевом магнитном полях в полупроводниках со стандартной зоной и носителями одного типа

| Название коэффициента и определение | Условие измерения | Формула | Коэффициенты, зависящие от степени вырождения и механизма рассеяния |
|-----------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Удельная электропроводность $\sigma_0 = j_x / E_x$ | $\frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} =$ $= j_x = H_z = 0$ | $\sigma_0 = \frac{2}{3} q^2 4\pi (2m^*)^{3/2} \tau_0 (k_0 T)^{r+1} A$ (3.1) | $A = (r+1) F_r$ |
| Электронная составляющая коэффициента теплопроводности $\kappa_0^a = \left(W_x / \frac{\partial T}{\partial x} \right) - \chi^{\text{ф}}$ | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = H_z = 0$ | $\kappa_0^a = \left(\frac{k_0}{e} \right)^2 T \sigma_0 L$ (3.2) | $L = \frac{r+3}{r+1} \frac{F_{r+2}}{F_r} - \frac{(r+2)^2}{(r+1)^2} \frac{F_{r+1}^2}{F_r^2}$ |
| Коэффициент термоэ. д. с. $\alpha_0 = E_x / \frac{\partial T}{\partial x}$ | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = H_z = 0$ | $\alpha_0 = \frac{k_0}{q} C$ (3.3) | $C = \frac{r+2}{r+1} \frac{F_{r+1}}{F_r} - \mu^*$ |
| Постоянная Холла $R = E_y / (j_x H_z)$ | $j_y =$ $= \frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = 0$ | $R = \frac{3}{2} \frac{\hbar^3}{4\pi (2m^*)^{3/2}} \frac{(k_0 T)^{-3/2}}{q} B$ (3.4) | $B = \frac{2r+1/2}{(r+1)^2} \frac{F_{2r-1/2}}{F_r^2}$ |
| Коэффициент Эттингсгаузена $P = \left(-\frac{\partial T}{\partial y} \right) / (j_x H_z)$ | $\frac{\partial T}{\partial x} = j_y =$ $= W_y = 0$ | $P = \frac{k_0 T}{q} \frac{R \sigma_0}{\kappa_0} D$ (3.5) | $D = \frac{2r+3/2}{2r+1/2} \frac{F_{2r+1/2}}{F_{2r-1/2}} \frac{r+2}{r+1} \frac{F_{r+1}}{F_r}$ |
| Коэффициент поперечного эффекта Нернста-Эттингсгаузена $Q = E_y / \left(-\frac{\partial T}{\partial x} H_z \right)$ | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = 0$ | $Q = \frac{k_0}{q} R \sigma_0 D$ (3.6) | $G = \frac{1}{AL} \left\{ \frac{(2r+5/2)(r+1)}{2r+1/2} \times \right.$ $\times \frac{F_{2r+3/2} F_r}{F_{2r-1/2}} - 2 \frac{(r+2)(2r+3/2)}{2r+1/2} \times$ $\left. \times \frac{F_{2r+1/2} F_{r+1}}{F_{2r-1/2}} + \frac{(r+2)^2}{r+1} \frac{F_{r+1}^2}{F_r} \right\}$ |
| Коэффициент Риги-Ледюка $S = -$ $-\frac{\partial T}{\partial y} / \left(-\frac{\partial T}{\partial x} H_z \right)$ | $i_x = i_y = W_y = 0$ | $S = -R \sigma_0 - \frac{\kappa_0^a}{\kappa_0} G$ (3.7) | |

Выражения для основных кинетических коэффициентов в слабом и нулевом магнитном полях в полупроводниках со стандартной зоной и носителями одного типа

| Название коэффициента и определение | Условие измерения | Формула | Коэффициенты, зависящие от степени вырождения и механизма рассеяния |
|-----------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Удельная электропроводность $\sigma_0 = j_x / E_x$ | $\frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} =$ $= j_x = H_z = 0$ | $\sigma_0 = \frac{2}{3} q^2 4\pi (2m^*)^{3/2} \tau_0 (k_0 T)^{r+1} A$ (3.1) | $A = (r+1) F_r$ |
| Электронная составляющая коэффициента теплопроводности $\kappa_0^a = \left(W_x / \frac{\partial T}{\partial x} \right) - \chi^{\text{ф}}$ | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = H_z = 0$ | $\kappa_0^a = \left(\frac{k_0}{e} \right)^2 T \sigma_0 L$ (3.2) | $L = \frac{r+3}{r+1} \frac{F_{r+2}}{F_r} - \frac{(r+2)^2}{(r+1)^2} \frac{F_{r+1}^2}{F_r^2}$ |
| Коэффициент термоэ. д. с. $\alpha_0 = E_x / \frac{\partial T}{\partial x}$ | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = H_z = 0$ | $\alpha_0 = \frac{k_0}{q} C$ (3.3) | $C = \frac{r+2}{r+1} \frac{F_{r+1}}{F_r} - \mu^*$ |
| Постоянная Холла $R = E_y / (j_x H_z)$ | $j_y =$ $= \frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = 0$ | $R = \frac{3}{2} \frac{\hbar^3}{4\pi (2m^*)^{3/2}} \frac{(k_0 T)^{-3/2}}{q} B$ (3.4) | $B = \frac{2r+1/2}{(r+1)^2} \frac{F_{2r-1/2}}{F_r^2}$ |
| Коэффициент Эттингсгаузена $P = \left(-\frac{\partial T}{\partial y} \right) / (j_x H_z)$ | $\frac{\partial T}{\partial x} = j_y =$ $= W_y = 0$ | $P = \frac{k_0 T}{q} \frac{R \sigma_0}{\kappa_0} D$ (3.5) | $D = \frac{2r+3/2}{2r+1/2} \frac{F_{2r+1/2}}{F_{2r-1/2}} \frac{r+2}{r+1} \frac{F_{r+1}}{F_r}$ |
| Коэффициент поперечного эффекта Нернста-Эттингсгаузена $Q = E_y / \left(-\frac{\partial T}{\partial x} H_z \right)$ | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = 0$ | $Q = \frac{k_0}{q} R \sigma_0 D$ (3.6) | $G = \frac{1}{AL} \left\{ \frac{(2r+5/2)(r+1)}{2r+1/2} \times \right.$ $\times \frac{F_{2r+3/2} F_r}{F_{2r-1/2}} - 2 \frac{(r+2)(2r+3/2)}{2r+1/2} \times$ $\left. \times \frac{F_{2r+1/2} F_{r+1}}{F_{2r-1/2}} + \frac{(r+2)^2}{r+1} \frac{F_{r+1}^2}{F_r} \right\}$ |
| Коэффициент Риги-Ледюка $S = -$ $-\frac{\partial T}{\partial y} / \left(-\frac{\partial T}{\partial x} H_z \right)$ | $i_x = i_y = W_y = 0$ | $S = -R \sigma_0 - \frac{\kappa_0^a}{\kappa_0} G$ (3.7) | |

| Название коэффициента и определение | Условие измерения | Формула | Коэффициенты, зависящие от степени вырождения и механизма рассеяния |
|-----------------------------------------------------------|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| Магнетосопротивление | $i_y = \frac{\partial T}{\partial x} =$ $= \frac{\partial T}{\partial y} = 0$ | $\frac{\Delta \rho}{\rho_0 H_z^2} = (R\sigma_0)^2 M \quad (3.8)$ | $M = \frac{3r(r+1)}{(2r+1/2)^2} \frac{F_{3r-1} F_r}{F_{2r-1/2}^2} - 1$ |
| Магнитотермо-э. д. с. | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = 0$ | $\frac{\Delta \alpha}{\alpha_0 H_z^2} = (R\sigma_0) \frac{N}{C} \quad (3.9)$ | $N = D - M + \frac{(r+2)3r}{(2r+1/2)^2} \frac{F_{3r-1} F_{r+1}}{F_{2r-1/2}^2} + 1$ |
| Относительное изменение теплопроводности в магнитном поле | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = 0$ | $\frac{\Delta \kappa(H)}{\kappa_0 H_z^2} =$ $= \left(\frac{k_0}{e}\right)^2 \frac{\sigma_0 T}{\kappa_0} (R\sigma_0)^2 (D^2 - V) \quad (3.10)$ | $V = \frac{(3r+2)(r+1)}{(2r+1/2)^2} \frac{F_{3r+1} F_r}{F_{2r-1/2}^2} -$ $- 2 \frac{(3r+1)(r+2)}{(2r+1/2)^2} \frac{F_{3r} F_{r+1}}{F_{2r-1/2}^2} +$ |
| | $i_x = i_y = W_y = 0$ | $\frac{\Delta \kappa(H)}{\kappa_0 H_z^2} = \left(\frac{k_0}{e}\right)^2 \frac{\sigma_0 T}{\kappa_0} (R\sigma_0) \times$ $\times \left[D^2 - V + \left(\frac{k_0}{e}\right)^2 \frac{\sigma_0 T}{\kappa_0} G^2 \right] \quad (3.11)$ | $+ \frac{3r(r+2)^2}{(2r+1/2)^2 (r+1)} \frac{F_{3r-1} F_{r+1}^2}{F_{2r-1/2}^2 F_r}$ |

Обозначения: k_0 — постоянная Больцмана; q — заряд носителя тока; H — напряженность магнитного поля; E — напряженность электрического поля; j — плотность тока;

$$F_n \equiv F_0(\mu^*) \equiv \int_0^{\infty} \frac{x^n}{1 + e^{x - \mu^*}} dx \text{ — интеграл Ферми;}$$

μ^* — приведенный уровень химического потенциала; W — плотность потока энергии.

Примечание. Образцы имеют форму прямоугольных параллелепипедов; H направлено по оси z ; $E, \nabla T$ — в плоскости (x, y) .

| Название коэффициента и определение | Условие измерения | Формула | Коэффициенты, зависящие от степени вырождения и механизма рассеяния |
|-----------------------------------------------------------|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| Магнетосопротивление | $i_y = \frac{\partial T}{\partial x} =$ $= \frac{\partial T}{\partial y} = 0$ | $\frac{\Delta \rho}{\rho_0 H_z^2} = (R\sigma_0)^2 M \quad (3.8)$ | $M = \frac{3r(r+1)}{(2r+1/2)^2} \frac{F_{3r-1} F_r}{F_{2r-1/2}^2} - 1$ |
| Магнитотермо-э. д. с. | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = 0$ | $\frac{\Delta \alpha}{\alpha_0 H_z^2} = (R\sigma_0) \frac{N}{C} \quad (3.9)$ | $N = D - M + \frac{(r+2)3r}{(2r+1/2)^2} \frac{F_{3r-1} F_{r+1}}{F_{2r-1/2}^2} + 1$ |
| Относительное изменение теплопроводности в магнитном поле | $i_x = i_y =$ $= \frac{\partial T}{\partial y} = 0$ | $\frac{\Delta \kappa(H)}{\kappa_0 H_z^2} =$ $= \left(\frac{k_0}{e}\right)^2 \frac{\sigma_0 T}{\kappa_0} (R\sigma_0)^2 (D^2 - V) \quad (3.10)$ | $V = \frac{(3r+2)(r+1)}{(2r+1/2)^2} \frac{F_{3r+1} F_r}{F_{2r-1/2}^2} -$ $- 2 \frac{(3r+1)(r+2)}{(2r+1/2)^2} \frac{F_{3r} F_{r+1}}{F_{2r-1/2}^2} +$ |
| | $i_x = i_y = W_y = 0$ | $\frac{\Delta \kappa(H)}{\kappa_0 H_z^2} = \left(\frac{k_0}{e}\right)^2 \frac{\sigma_0 T}{\kappa_0} (R\sigma_0) \times$ $\times \left[D^2 - V + \left(\frac{k_0}{e}\right)^2 \frac{\sigma_0 T}{\kappa_0} G^2 \right] \quad (3.11)$ | $+ \frac{3r(r+2)^2}{(2r+1/2)^2 (r+1)} \frac{F_{3r-1} F_{r+1}^2}{F_{2r-1/2}^2 F_r}$ |

Обозначения: k_0 — постоянная Больцмана; q — заряд носителя тока; H — напряженность магнитного поля; E — напряженность электрического поля; j — плотность тока;

$$F_n \equiv F_0(\mu^*) \equiv \int_0^{\infty} \frac{x^n}{1 + e^{x - \mu^*}} dx \text{ — интеграл Ферми;}$$

μ^* — приведенный уровень химического потенциала; W — плотность потока энергии.

Примечание. Образцы имеют форму прямоугольных параллелепипедов; H направлено по оси z ; $E, \nabla T$ — в плоскости (x, y) .

КОНТАКТНЫЕ ЯВЛЕНИЯ

8.1. ОСОБЕННОСТИ КОНТАКТНЫХ ЯВЛЕНИЙ

Во многих случаях требуется, чтобы контакт обладал минимальным сопротивлением и был омическим: такой контакт не искажает форму переменного синусоидального тока, не вносит гармоник и пропускает без искажений импульсы (рис. 8.1, а). Идеализированная характеристика выпрямляющего контакта должна иметь вид, представленный на рис. 8.1, б, в пропускном направлении его сопротивление должно быть равным нулю, в запиорном — бесконечным.

Требования нелинейности контакта возникают всякий раз, когда требуется преобразование сигнала, для выпрямления и детектирования, модуляции и умножения частоты, ограничения сигналов и стабилизации напряжений, для усиления и генерации и т. д.

Кроме этого, как мы видели в гл. 1, контакт дырочного и электронного полупроводника обладает чрезвычайно важным свойством, на котором основана вся транзисторная электроника, — способностью инжектировать неосновные носители.

В зависимости от конкретной задачи, выполняемой нелинейным контактом, к нему предъявляются различные требования.

Как уже упоминалось, для технических выпрямителей желательно, чтобы сопротивление и допустимое напряжение в запиорном направлении были бы максимальными. Для этого в контакте не должны происходить так называемые явления в сильных полях (электростатическая

и ударная ионизация), p - n переход не должен быть слишком толстым, чтобы в нем не мог развиваться лавинный эффект. В то же время контакт должен быть достаточно толстым, чтобы исключить туннельный эффект, и длина свободного пробега в нем не должна быть велика, чтобы исключить ударную ионизацию.

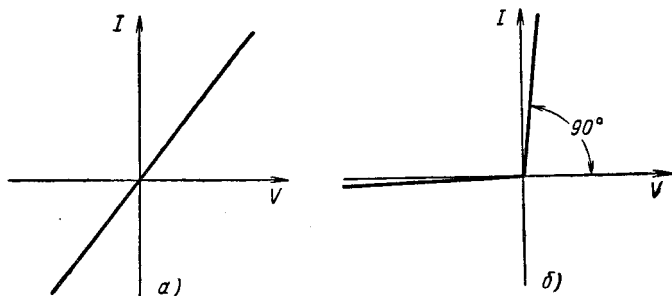


Рис. 8.1. Вольтамперная характеристика идеального омического (а) и идеального выпрямляющего (б) контактов.

Однако прямая характеристика толстого запирающего слоя хуже, чем тонкого, поэтому здесь приходится идти на известный компромисс *).

При пропускном направлении тока электроны диффундируют в p -область и там рекомбинируют с дырками,

*) Здесь следует отметить, что слишком толстый контакт вообще теряет выпрямляющие свойства. Критерием в данном случае является отношение двух характерных длин: дебаевского радиуса экранировки $r_0 = \sqrt{\kappa kT/4\pi e^2 n}$ (где n — число подвижных носителей заряда и κ — диэлектрическая постоянная) и расстояния x_0 , на котором значительно меняется концентрация N активных примесей (доноров или акцепторов),

$$x_0 = N/(dN/dx).$$

Как известно, радиус Дебая характеризует расстояние, на котором в данной среде восстанавливается нейтральность, т. е. любой локальный неподвижный заряд нейтрализуется окружающими подвижными зарядами.

Если $x_0 \gg r_0$, то это означает, что нейтральность практически вообще не нарушается и слой объемного заряда не образуется; такой плавный p - n переход (или контакт полупроводника с металлом) не будет обладать выпрямляющими свойствами.

дырки диффундируют в n -область и рекомбинируют с электронами. Поэтому ток в пропускном направлении будет тем больше, чем больше коэффициент диффузии (подвижность) носителей и чем больше коэффициент рекомбинации, так как тем больше будет градиент концентрации неосновных носителей и скорость их диффузии.

Совершенно иные требования предъявляются к p - n переходу детектора малых сигналов; в данном случае высокое обратное напряжение не играет существенной роли, к величине прямого и обратного сопротивления также предъявляются обычно более низкие требования. На первый план здесь выступает низкий уровень шумов и крутой подъем пропускной характеристики.

Для детектирования без искажений сигналов в широком диапазоне частот должны быть малы также электростатическая и диффузионная емкости (см. ниже) и время пролета через p - n переход неосновных носителей.

В ряде случаев характер детектирования ставит также вполне определенные требования перед формой прямой характеристики p - n перехода (линейное детектирование, квадратичное и др.).

В рассмотренных выше случаях и электростатическая, и барьерная емкости p - n переходов играли отрицательную роль. В то же время имеется ряд приборов, так называемые варикапы, действие которых основано на зависимости барьерной емкости от приложенного напряжения. Варикапы в настоящее время находят ряд важных применений: в параметрических усилителях, для стабилизации напряжения, для преобразования сигналов, в счетно-решающих устройствах и др. Каждое из этих применений предъявляет свои конкретные требования к величине емкости p - n перехода, к ее частотной зависимости и зависимости от приложенного напряжения, а также ко всем остальным характеристикам p - n перехода.

Во всех рассмотренных выше случаях туннельный эффект является отрицательным побочным явлением. Однако за последнее время были разработаны новые приборы, так называемые обращенные диоды и туннельные диоды, основанные на этом эффекте и нашедшие широкое применение в технике высоких частот (см. ниже).

Совершенно иные требования предъявляются к p - n переходам в транзисторах. Выше говорилось, что для силового выпрямителя желательны большие коэффициенты

рекомбинации и в p - и n -области. Для хорошей работы транзистора необходимо, чтобы эмиттерный p - n переход обладал высоким коэффициентом инжекции, а для этого он должен быть асимметричным; чтобы число инжектируемых в базу носителей было велико, должна быть велика концентрация основных носителей в эмиттерной области; чтобы неосновные носители проходили без потерь базовую область, их диффузионная длина в базовой области должна быть велика, т. е. велики коэффициент диффузии и время жизни. Следует заметить, что так как напряжение на базе обычно очень мало, ее емкость не играет существенной роли и поэтому для уменьшения времени пролета неосновных носителей базу делают достаточно тонкой. Время пролета носителей через базу можно также уменьшить созданием специального «тянущего» поля. За последнее время был разработан новый тип транзисторов — дрейфовые. В дрейфовых транзисторах примеси в базе распределены не равномерно, а таким образом, что создается поле, тянущее носители по направлению к коллектору, что во много раз уменьшает время их пролета через базу и повышает высокочастотные характеристики прибора.

Упомянем еще об одном типе транзисторов, разработанных за последнее время, так называемых лавинных триодах, в которых дополнительное усиление тока достигается за счет ударной ионизации и развития лавинного эффекта в коллекторной области. Для того чтобы развился лавинный эффект, толщина коллекторной области и длина свободного пробега носителей в ней должны быть достаточно велики.

Наконец, за последнее время расширились неудавшиеся в 30-х годах опыты по использованию выпрямления на контакте двух металлов, разделенных изоляционной прослойкой, и на контакте металла с полупроводником.

Приведенные выше примеры в какой-то мере показывают, насколько сложна и многообразна картина контактных явлений.

Следует отметить, что простой механический контакт двух металлов, металла и полупроводника и двух полупроводников никогда не дает ожидаемых свойств, так как при этом неизбежно сохраняется слой адсорбированного воздуха, глубже слой поверхностных пленок, еще глубже слой деформированного, богатого примесями и дефектами

основного материала. При этом сама поверхность имеет характер ступенчатых пирамид, высота которых может колебаться в зависимости от характера обработки в пределах от 10^{-6} до 10^{-3} см. Поверхность этих пирамид также испещрена впадинами, каналами и наростами. Все перечисленные выше дефекты и загрязнения поверхности смазывают всю картину, и поэтому для получения хорошо выпрямляющего контакта с заданной характеристикой требуется сложный технологический процесс.

Переход типа p - n получается, как правило, в единой пластинке полупроводника, в котором граница с заданными свойствами между p - и n -слоями создается путем сплавления или диффузии p -примеси в монокристалл n -типа или сплавления или диффузии примеси n -типа в монокристалл p -типа.

Для успешной разработки технологии контакта, обладающего необходимыми свойствами, необходимо детальное знание всех физических процессов, происходящих в области контакта. В этой главе читатель сможет начать знакомство с теорией контактных явлений. Более детально этот вопрос можно изучить, пользуясь литературой [4, 31].

8.2. КОНТАКТ ПОЛУПРОВОДНИКА И МЕТАЛЛА

На рис. 1.17 было представлено распределение заряда, поля и потенциала на границе полупроводника и металла *).

В зазоре между металлом и полупроводником поле постоянно и равно полю в плоском конденсаторе:

$$E_0 = 4\pi q_0 \quad (8.1)$$

и разность потенциалов

$$V_0 = dE_0, \quad (8.2)$$

где q_0 — заряд на единицу поверхности полупроводника и металла и d — толщина зазора.

*) Мы при этом предполагаем, что все примеси в слое d полупроводника ионизованы, весь расчет будет вестись для единицы поверхности, рис. 1.17 построен в предположении, что диэлектрическая постоянная $\epsilon = 1$.

В слое l объемного заряда полупроводника поле убывает по линейному закону:

$$E(x) = E_0 \frac{x}{l\kappa} \quad (8.3)$$

и разность потенциалов

$$V_K = \int_0^l E dx = \frac{E_0}{l\kappa} \int_0^l x dx = \frac{E_0 l^2}{2l\kappa} = \frac{2\pi q_0 l}{\kappa}. \quad (8.4)$$

Следовательно, полная контактная разность потенциалов будет распределена между слоем объемного заряда и зазором:

$$V_K = 4\pi q_0 \left(d + \frac{l}{2\kappa} \right). \quad (8.5)$$

Подставляя в (8.5) q_0 , выраженное через l

$$q_0 = N_d e l, \quad (8.6)$$

где N_d — число ионизированных донорных примесей в единице объема, получаем квадратное уравнение относительно l :

$$l^2 + 2dl\kappa - \frac{V_K \kappa}{2\pi e N_d} = 0,$$

откуда

$$l = -\kappa d + \sqrt{\varepsilon^2 d^2 + \frac{V_K \kappa}{2\pi e N_d}} \quad (8.7)$$

или

$$l = \kappa d \left(\sqrt{1 + \frac{V_K \kappa}{2\pi e N_d \kappa^2 d^2}} - 1 \right). \quad (8.8)$$

Интересно рассмотреть два предельных случая:

1. Случай большого зазора d :

$$\frac{\kappa V_K}{2\pi N_d e \kappa^2 d^2} \ll 1; \quad l = \frac{V_K}{4\pi N_d e d}. \quad (8.9)$$

2. Случай малого зазора d :

$$\frac{\kappa V_K}{2\pi N_d e \kappa^2 d^2} \gg 1 \quad (8.10)$$

и

$$l = \sqrt{\frac{\kappa V_K}{2\pi N_d e}} - \kappa d. \quad (8.11)$$

Оценка показывает, что критерий (8.10) не является очевидным. На практике в зависимости от концентрации доноров мы можем встретиться и с тем и с другим случаем. При $N = 10^{20}$ слой объемного заряда становится чрезвычайно тонким и основную роль начинает играть туннельный эффект. Этот случай мы рассмотрим при анализе работы обращенных и туннельных диодов, в ближайших параграфах мы будем считать запиорный слой достаточно толстым, чтобы туннельным эффектом можно было пренебречь.

8.3. ДИФФУЗИОННАЯ ТЕОРИЯ ВЫПРЯМЛЕНИЯ МОТТА (ХИМИЧЕСКИЙ БАРЬЕРНЫЙ СЛОЙ НА ГРАНИЦЕ МЕТАЛЛА С ПОЛУПРОВОДНИКОМ)

Если в результате технологической обработки на границе между металлом и полупроводником создан слой вещества, лишенный примесей *), то распределение потенциала на таком контакте будет иметь вид, представленный на рис. 8.2, а, или (полагая, что слой достаточно толст, а зазор тонок) вид, схематически представленный на рис. 8.2, б.

Моттом была создана теория выпрямления на таком контакте при следующих предположениях:

- контактная разность потенциалов (или, что то же самое, падение потенциала) в зазоре много больше энергии ионизации примесных центров;
- концентрация носителей и ионизованных примесей в запиорном слое настолько мала, что поле можно считать постоянным;
- падением напряжения в объеме полупроводника можно пренебречь;
- поверхностными уровнями и сопротивлением воздушного зазора можно также пренебречь;
- средняя длина свободного пробега носителей много меньше толщины слоя.

Последнее предположение означает, что в данном случае к движению носителей в слое можно применить диффузионное уравнение

$$I = neuE - De \frac{dn}{dx}. \quad (8.12)$$

*) Или каким-либо специальным приемом в промежутке между металлом и полупроводником создан слой диэлектрика.

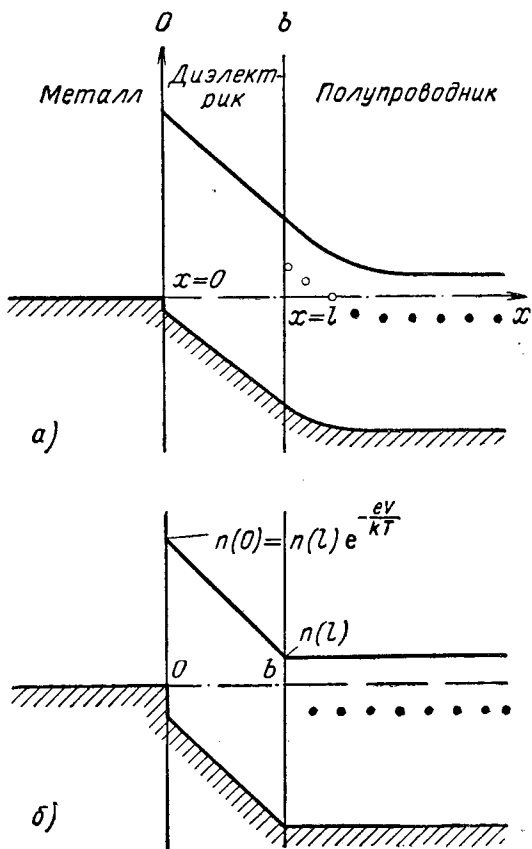


Рис. 8.2. Зонная диаграмма химического запорного слоя:

- а) с учетом объемного заряда в полупроводниках;
 б) в упрощенном виде.

Учитывая, что ток (в стационарном состоянии) во всем зазоре постоянен, можно в (8.13) разделить переменные:

$$\frac{dn}{an-l} = \frac{dx}{b}, \text{ где } a = euE \text{ и } b = eD, \quad (8.13)$$

и произвести интегрирование в пределах запорного слоя:

$$\ln \frac{an(l)-l}{an(0)-l} = \frac{a}{b} l = \frac{eEl}{kT}. \quad (8.14)$$

При выводе (8.14) использовано соотношение Эйнштейна

$$eD = kTu. \quad (8.15)$$

Потенцируя левую часть (8.14) и решая полученное выражение относительно тока, находим

$$I = euE \frac{n(0) e^{\frac{eEl}{kT}} - n(l)}{e^{\frac{eEl}{kT}} - 1}. \quad (8.16)$$

Полагая, что падение напряжения в запорном слое много больше kT , и пренебрегая единицей в знаменателе (8.16), получаем

$$I = euE [n(0) - n(l) e^{-\frac{eEl}{kT}}]. \quad (8.17)$$

Заменяя

$$n(0) = n(l) e^{-\frac{eV_K}{kT}}, \quad (8.18) *$$

где V_K — контактная разность потенциалов или диффузионное падение напряжения в запорном слое, и

$$E = \frac{V_K + V}{l}, \quad (8.19)$$

где V — приложенное внешнее напряжение, получаем окончательно

$$I = \frac{eun}{l} (V_K + V) e^{-\frac{eV_K}{kT}} (1 - e^{-\frac{eV}{kT}}), \quad (8.20)$$

где V считается отрицательным для пропускного направления и положительным для запорного.

Формуле (8.20) можно придать более наглядный вид:

$$I = ev_d n(0) (1 - e^{-\frac{eV}{kT}}), \quad (8.21)$$

*) В дальнейшем $n(l)$ — концентрацию носителей в объеме полупроводников — мы будем обозначать просто n .

где $v_d = uE$ — дрейфовая скорость носителей в запиорном слое; $n_0 = ne^{-eV_k/kT}$ — концентрация носителей в металле, имеющих достаточную энергию, чтобы преодолеть потенциалный барьер.

Формулы (8.19) и (8.21) верны до тех пор, пока для пропускного направления приложенное напряжение много меньше высоты потенциалного барьера.

С увеличением обратного напряжения ток из металла растет линейно (за счет v_d), а ток из полупроводника стремится к нулю. Сопротивление барьера растет и приближается к предельному значению

$$\frac{1}{\rho_\infty} = \frac{I}{V} = \frac{eun_0}{l}. \quad (8.22)$$

По мнению ряда авторов, в селеновых и меднозакисных выпрямителях осуществляются условия, близкие к предположениям, положенным в основу теории Мотта *), поэтому наиболее естественно сравнить результаты теории с вольтамперными характеристиками таких выпрямителей.

Проведение этого анализа показывает, что пропускная характеристика удовлетворительно согласуется с теорией, а для запиорного направления сопротивление не следует (8.21), а падает начиная с некоторого предельного значения.

Измерение емкости запиорного слоя в этих выпрямителях показывает, что толщина его очень велика, порядка 10^{-3} — 10^{-4} см, и, следовательно, это падение сопротивления нельзя объяснить туннельным эффектом.

В действительности причина падения сопротивления в том, что с увеличением напряжения на запиорном слое уменьшается работа выхода из металла, а следовательно, и высота потенциалного барьера.

Мы уже упоминали (см. гл. 1), что одним из компонентов работы выхода является работа против так называемых сил зеркального изображения, сущность этого явления заключается в следующем.

Если к поверхности проводящего тела приближается (предположим для определенности) отрицательный заряд, то под влиянием его поля в этом теле также начинают перемещаться заряды. Положительные заряды притягиваются

*) Ряд других исследователей считают, что в обоих случаях в результате технологических операций в выпрямителях образуется p - n переход; в меднозакисном — между закисью и окисью меди, в селеновом — между селеном и селенистым кадмием.

к поверхности, а отрицательные — уходят от нее. Это перемещение зарядов продолжается до тех пор, пока наведенный таким образом на поверхности заряд не нейтрализует полностью в объеме проводящего тела поле стороннего заряда.

Для того чтобы вычислить, с какой силой притягивается к поверхности заряд e (рис. 8.3), рассуждают следующим образом. Результирующее тангенциальное поле, возникшее в любой точке поверхности в результате действия

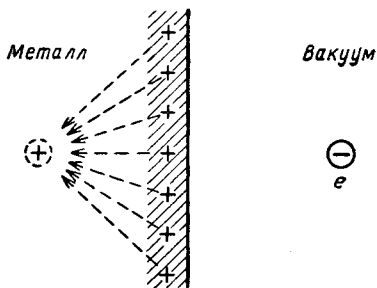


Рис. 8.3. Схема возникновения сил зеркального изображения.

двух зарядов, наведенного и стороннего, должно быть равно нулю *); это значит, что индуцированный заряд в каждой точке поверхности создает поле, равное и противоположное по направлению полю основного заряда. Такое поле создавало бы его «зеркальное изображение», т. е. заряд, равный ему, противоположный по знаку и находящийся на равном расстоянии по противоположную

сторону поверхности. Можно показать, что такое же поле этот заряд будет создавать в любой точке вне проводника. Следовательно, дополнительную силу, с которой притягивается основной заряд к поверхности, можно выразить как

$$F_{\text{зерк}} = \frac{e^2}{4x^2}, \quad (8.23)$$

где x — расстояние заряда от поверхности.

Чтобы определить соответствующий вклад в работу выхода, надо проинтегрировать выражение (8.23) от бесконечности до некоторого расстояния a , порядка межатомного расстояния, на котором эта сила перестает действовать:

$$\Delta W_{\text{зерк}} = - \int_a^{\infty} F_{\text{зерк}} dx = - \frac{e^2}{4a}. \quad (8.24)$$

*) В противном случае заряды перемещались бы до тех пор, пока не выполнится это условие; нормальная составляющая поля не должна быть равна нулю, так как поверхность ограничивает перемещение зарядов в этом направлении.

При наличии поля в запиорном слое

$$E = \frac{V_{\kappa} + V}{l}, \quad (8.25)$$

потенциал будет описываться уравнением

$$U = -\frac{e}{4x} - Ex \quad (8.26)$$

и проходить через максимум, величина которого будет зависеть от величины и знака приложенного внешнего напряжения V .

Расчет в данном случае совершенно аналогичен вычислению понижения энергии активации примесных атомов при термоэлектронной ионизации (см. стр. 287), и мы не станем его повторять, а приведем результат — снижение высоты потенциального барьера ΔV_{κ} как функцию от приложенного внешнего напряжения V :

$$\Delta V_{\kappa} = \frac{e^{3/2}}{\varepsilon^{1/2} l^{1/2}} [(V_{\kappa} + V)^{1/2} - V_{\kappa}^{1/2}]. \quad (8.27)$$

Если разложить ΔV_{κ} в ряд по степеням V/V_{κ} и ограничиться в разложении первым членом, то можно получить следующее выражение для тока:

$$I' = I_0 e^{\frac{-\alpha V}{kT}} (1 - e^{-\frac{eV}{kT}}), \quad (8.28)$$

где

$$I_0 = ev_d n_0 \quad (8.29)$$

и α — коэффициент порядка 0,05.

Чтобы согласовать теорию с опытом, надо брать α значительно больше. Качественно тот же результат должен дать туннельный эффект, однако толщина запиорного слоя в меднозакисных и селеновых выпрямителях слишком велика. Можно предположить, что в запиорном слое есть небольшие очень тонкие участки, влияние которых не сказывается заметно на емкости, но эти участки обуславливают электростатической пробой выпрямителя.

8.4. ДИОДНАЯ ТЕОРИЯ БЕТЕ

Если имеет место обратное соотношение между длиной свободного пробега и толщиной запиорного слоя по сравнению с рассмотренным выше, т. е. длина свободного пробега

много больше толщины запирающего слоя, то последний «простреливается» электронами как вакуумный промежуток в диоде. Отсюда и название теории выпрямления, рассматривающей этот случай.

Если ось x направлена перпендикулярно поверхности полупроводника, то через запирающий слой будут за 1 сек проскакивать все электроны, находящиеся вблизи контакта (т. е. в слое v_x) и имеющие кинетическую энергию, достаточную, чтобы преодолеть потенциальный барьер, т. е. удовлетворяющую условию

$$\frac{1}{2} m v_x^2 = \frac{p_x^2}{2m} \geq e(V_K + V), \quad (8.30)$$

где V — внешнее напряжение, приложенное к запирающему слою, или

$$p_x > \sqrt{2e(V_K + V)m}. \quad (8.31)$$

Ток из полупроводника в металл, создаваемый такими электронами, будет равен

$$\begin{aligned} I_1 &= \frac{2e}{h^3} \int f v_x dg = \frac{2}{h^3} e \frac{\mu}{kT} \times \\ &\times \int_{-\infty}^{\infty} e^{-\frac{p_y^2}{2mkT}} dp_y \int_{-\infty}^{\infty} e^{-\frac{p_z^2}{2mkT}} dp_z \times \\ &\times \int_{p_x = \sqrt{2e(V_K + V)m}}^{\infty} v_x e^{-\frac{p_x^2}{2mkT}} dp_x. \end{aligned} \quad (8.32)$$

Интегрирование (8.32) дает следующий результат:

$$I_1 = \frac{1}{4} n v_0 e e^{-\frac{e(V_K + V)}{kT}}, \quad (8.33)$$

где введено обозначение

$$v_0 = \sqrt{\frac{8kT}{\pi m}} \quad (8.34)$$

(v_0 — средняя тепловая скорость).

Ток, текущий из металла в полупроводник, не зависит от приложенного напряжения, следовательно, он должен быть равен току из полупроводника при $V=0$ (так как

при этом результирующий ток равен нулю):

$$I_2 = I_1(0) = \frac{1}{4} n v_0 e e^{-\frac{eV_K}{kT}}. \quad (8.35)$$

Таким образом, результирующий ток

$$I = I_2 - I_1 = \frac{1}{4} e n v_0 e^{-\frac{eV_K}{kT}} \left(1 - e^{-\frac{eV}{kT}}\right). \quad (8.36)$$

Итак, отличие вольтамперной характеристики тонкого запирающего слоя заключается в следующем:

1) ток насыщения в этом случае

$$I_s = \frac{1}{4} e n v_0 e^{-\frac{eV_K}{kT}} \quad (8.37)$$

не зависит от приложенного напряжения, в то время как в толстом запирающем слое он растет для запирающего направления и падает для пропускного;

2) ток насыщения в тонком запирающем слое больше во столько раз, во сколько тепловая скорость v_0 носителей больше их скорости дрейфа v_d .

Обе эти особенности являются достоинствами тонкого запирающего слоя. Недостатком его является тот факт, что он не выдерживает высоких обратных напряжений из-за электростатической ионизации. Соотношение, характеризующее тонкий запирающий слой, соблюдается в германиевых детекторах; при концентрации примесей порядка 10^{19} толщина запирающего слоя падает до 10^{-7} . В то же время подвижность носителей в германии велика, при комнатной температуре $\mu \approx 10^3$ см²/в·сек, и соответственно велика длина свободного пробега; таким образом, в данном случае мы имеем дело с ситуацией, обратной селеновым выпрямителям *).

8.5. ТЕОРИЯ ФИЗИЧЕСКОГО ЗАПИРАЮЩЕГО СЛОЯ (ТЕОРИЯ ИСТОЩЕНИЯ ШОТТКИ)

Физический запирающий слой в равновесии мы рассмотрели в начале этой главы; как видно из (8.7), его толщина и заряд пропорциональны корню из приложенного к нему

*) В действительности в данном случае длина свободного пробега не намного больше, а одного порядка с толщиной запирающего слоя, и изложенную ниже теорию можно применять с некоторой натяжкой.

напряжения — в этом заключается его основное отличие от химического запорного слоя *). Согласно (8.7) при $l \gg d$

$$U = V_{\kappa} \pm V = \frac{2\pi n e l^2}{\kappa}, \quad (8.38)$$

где U — суммарная разность потенциалов на запорном слое.

Напряженность поля в физическом запорном слое также непостоянна, а растет линейно и достигает максимального значения на границе с металлом:

$$E(0) = \frac{4\pi n e l}{\kappa} = \left[\frac{8\pi n e (V_{\kappa} \pm V)}{\kappa} \right]^{1/2}. \quad (8.39)$$

Если концентрация примесей очень велика $l \ll \lambda$ (где λ — длина свободного пробега), то опять следует пользоваться диодной теорией Бете и будут применимы изложенные выше результаты (так как ток не зависит при этих условиях от толщины слоя). В противоположном случае, когда $l \gg \lambda$, выводы оказываются существенно иными. Будем исходить, как и при анализе работы химического запорного слоя, из диффузионного уравнения

$$I = euE(x) - eD \frac{dn}{dx} = eD \left[\frac{en(x)}{kT} \frac{dU}{dx} - \frac{dn}{dx} \right]. \quad (8.40)$$

Умножая и левую и правую часть (8.40) на $e^{-\frac{eU}{kT}}$, получаем

$$I e^{-\frac{eU}{kT}} = eD \left[\frac{en(x)}{kT} \frac{dU}{dx} e^{-\frac{eU}{kT}} - \frac{dn}{dx} e^{-\frac{eU}{kT}} \right] = eD \frac{d}{dx} [n(x) e^{-\frac{eU}{kT}}]. \quad (8.41)$$

Интегрируя и левую и правую часть уравнения (8.41), получаем

$$I \int_0^l e^{-\frac{eU}{kT}} dx = eD (A - n e^{-\frac{eU}{kT}}). \quad (8.42)$$

*) Второе отличие физического запорного слоя заключается в том, что в зависимости от приложенного напряжения меняется его заряд, это означает, что появляется дополнительная и динамическая, или, как ее называют, барьерная емкость $C_3 = \frac{dV}{dQ}$.

При отсутствии внешнего напряжения $U = V_{\text{к}}$ и $I = 0$, откуда $A = n_0 e^{-eV_{\text{к}}/kT}$ *) и, наконец, при $U = V_{\text{к}} \pm V$ получаем выражение для тока

$$I = \frac{eDne^{-\frac{eV_{\text{к}}}{kT}} (1 - e^{-\frac{eV}{kT}})}{\int_0^l e^{-\frac{eU}{kT}} dx}. \quad (8.43)$$

В выражении для тока (8.43) только знаменатель зависит от формы барьера (оно применимо также и к химическому запирающему слою), если в него подставить $U = [(V_{\text{к}} \pm V)/l]x$ и считать (как это мы делали раньше) $e(V_{\text{к}} \pm V) \gg kT$, то получим старое (8.21) выражение для вольтамперной характеристики. В рассматриваемом нами случае физического запирающего слоя

$$U = \frac{(2lx - x^2) 2\pi ne}{\kappa} \quad (8.44)$$

и интеграл в (8.43) берется приближенно при $e(V_{\text{к}} \pm V) \gg kT$:

$$\int_0^l e^{-\frac{eU}{kT}} dx \approx \frac{\kappa kT}{4\pi ne^2 l}. \quad (8.45) **$$

Подставляя (8.45) в (8.43), получаем

$$I = \frac{eDne^{-\frac{eV_{\text{к}}}{kT}} (1 - e^{-\frac{eV}{kT}})}{\frac{\kappa kT}{4\pi ne^2 l}} = \sigma \frac{4\pi enl}{\kappa} e^{-\frac{eV_{\text{к}}}{kT}} (1 - e^{-\frac{eV}{kT}}). \quad (8.46)$$

Подставив в (8.46) выражение для l из (8.38), получим

$$I = \sigma \left[\frac{(V_{\text{к}} \pm V)^{1/2} 8\pi ne}{\kappa} \right] e^{-\frac{eV_{\text{к}}}{kT}} (1 - e^{\pm \frac{eV}{kT}}). \quad (8.47)$$

Согласно (8.47) сопротивление запирающего слоя при $V \rightarrow \infty$ будет

$$\rho^{-1} = \frac{I}{V} = \frac{\sigma e}{kT} \left(\frac{8\pi neU}{\kappa} \right)^{1/2} e^{-\frac{eV_{\text{к}}}{kT}}. \quad (8.48)$$

*) n здесь так же, как и раньше, концентрация носителей в объеме полупроводника.

***) Интеграл (8.45) приводится к виду $\int_0^{s_1} e^{-s} ds$ путем подстановки $s = (l - x) (2\pi ne^2)^{1/2} / (\kappa kT)^{1/2}$.

С учетом (8.48) выражение (8.47) можно переписать в виде*)

$$I = \left(1 + \frac{V}{V_K}\right)^{1/2} \left(1 - e^{\pm \frac{eV}{kT}}\right) \frac{kT}{e\rho_0}. \quad (8.49)$$

Согласно (8.48) дифференциальное сопротивление при больших обратных напряжениях

$$\rho = \frac{\partial V}{\partial I} = \left(1 + \frac{V}{V_K}\right)^{1/2} 2\rho_0 \frac{eV}{kT}. \quad (8.50)$$

Как видно из (8.49), ρ растет с ростом напряжения, в то время как для химического запорного слоя сопротивление стремилось к постоянному значению; объясняется это различие тем, что с ростом напряжения растет толщина физического запорного слоя, что обуславливает дополнительный рост сопротивления.

8.6. ТЕОРИЯ *p-n* ПЕРЕХОДА УРАВНЕНИЕ НЕПРЕРЫВНОСТИ

Основным уравнением для решения многих задач, касающихся движения электронов в полупроводнике, является уравнение непрерывности, которое выражает закон сохранения материи по отношению к потоку частиц, число которых в некотором объеме может меняться за счет разности входящего и выходящего потоков, а также за счет каких-либо реакций, происходящих в этом объеме. В случае электронов и дырок такими реакциями может быть их генерация и рекомбинация.

Для электронов это уравнение будет иметь следующий вид:

$$\frac{\partial n}{\partial t} = \frac{n - n_0}{\tau} + \frac{1}{e} \operatorname{div} j_n + g'_n, \quad (8.51)$$

где n_0 — равновесная для данной температуры концентрация носителей;

$\operatorname{div} j_n$ — разность входящего и выходящего из данного объема тока;

*) В этой главе все расчеты проводятся для 1 см^2 контакта; поэтому мы не делаем различия между током и плотностью тока ($I = j$).

$(n - n_0)/\tau$ — разность скоростей тепловой генерации и рекомбинации (которая обращается в нуль при $n = n_0$);

g'_n — скорость генерации электронов в рассматриваемом объеме за счет сторонних факторов (фотонов, ионизирующих поток частиц, или других причин).

Точно такой же вид имеет уравнение непрерывности для дырок:

$$\frac{\partial p}{\partial t} = \frac{p - p_0}{\tau} + \frac{1}{e} \operatorname{div} j_p + g'_p; \quad (8.52)$$

j_n в (8.51) и j_p в (8.52) — это полные токи дырок и электронов, возникающие за счет дрейфа в электрическом поле и диффузии:

$$j_n = eu_n E + eD_n \nabla n \quad (8.53)$$

и

$$j_p = eu_p E - eD_p \nabla p. \quad (8.54)$$

p-n ПЕРЕХОД В РАВНОВЕСИИ

На рис. 1.21, б, в, г и д был представлен контакт электронного и дырочного полупроводника в равновесии (т. е. после того, как электроны перешли в дырочную область в таком количестве, что уровни химического потенциала сравнялись).

В соответствии с этим рисунком ход потенциала в *p-n* переходе определяется дифференциальными уравнениями

$$x < 0 \quad \frac{d^2\varphi}{dx^2} = \frac{4\pi e^2 N_d}{\kappa}, \quad (8.55)$$

$$x > 0 \quad \frac{d^2\varphi}{dx^2} = \frac{4\pi e^2 N_a}{\kappa}. \quad (8.56)$$

Решив уравнения (8.55) и (8.56) при граничных условиях (при $x = l_n$, $\varphi = 0$ и $\frac{d\varphi}{dx} = 0$ и при $x = -l_p$, $\varphi = \varphi_0$ и $\frac{d\varphi}{dx} = 0$), получим

$$0 < x < l_n, \quad \varphi_n = \frac{2\pi e^2 N_d}{\kappa} (l_n - x)^2, \quad (8.57)$$

$$-l_p < x < 0, \quad \varphi_p = \frac{2\pi e^2 N_a}{\kappa} (l_p + x)^2. \quad (8.58)$$

Полный заряд справа и слева от p - n перехода должен быть одинаков, поэтому

$$\frac{l_n}{l_p} = \frac{N_a}{N_d} = \frac{\rho_p}{n_n}. \quad (8.59)$$

Уравнения (8.57), и (8.58) (8.59) позволяют определить полную толщину слоя объемного заряда l ; действительно:

$$l = l_n + l_p, \quad (8.60)$$

$$\varphi_0 = \varphi_n + \varphi_p = W_2 - W_1 = eV_K. \quad (8.61)$$

Подставив (8.57) и (8.58) в (8.61) и выразив согласно (8.60) и (8.59) l_n и l_p через l :

$$l_n = \frac{\rho_p}{\rho_p + n_n} l \quad \text{и} \quad l_p = \frac{n_n}{n_n + \rho_p} l, \quad (8.62)$$

найдем

$$l = \left(\frac{\kappa V_K}{2\pi e^2} \frac{n_n + \rho_p}{n_n \rho_p} \right)^{1/2}. \quad (8.63)$$

ИНЖЕКЦИЯ И ЭКСТРАКЦИЯ НОСИТЕЛЕЙ

Рассмотрим теперь p - n переход в неравновесных условиях, т. е. когда к нему приложено внешнее напряжение V . Начнем с пропускного направления. В этом случае (см. гл. 1) число электронов в n -области, энергии которых достаточно, чтобы преодолеть потенциальный барьер (который уменьшился на величину eV , см. рис. 1.23) и перейти в p -область, будет

$$n_1 = \frac{2(2\pi m_n kT)^{3/2}}{h^3} e^{-\frac{\mu - eV}{kT}} = n_p e^{\frac{eV}{kT}}. \quad (8.64)$$

Предположим, что рекомбинацией носителей в слое объемного заряда можно пренебречь, и поэтому толщина его не играет существенной роли в работе p - n перехода; в соответствии с этим слой объемного заряда на рис. 1.22 сокращен до минимума и ход потенциала в нем изображен почти вертикальными линиями. При этих предположениях концентрацию носителей в слое объемного заряда можно считать постоянной и равной n_1 , т. е. на величину

$$\Delta n_p = n_1 - n_p = n_p \left(e^{\frac{eV}{kT}} - 1 \right) \quad (8.65)$$

больше равновесной концентрации электронов в p -области n_p .

Эти избыточные носители будут диффундировать в p -область и постепенно рекомбинировать с подходящими дырками, таким образом, концентрация избыточных электронов в p -области будет спадать по экспоненциальному закону:

$$\Delta n_p(x) = \Delta n_{p0} e^{-\frac{x}{L_n}}, \quad (8.66)$$

где L_n — так называемая диффузионная длина, выражающаяся через коэффициент диффузии и время жизни носителей, и x — расстояние от слоя объемного заряда. Это увеличение концентрации неосновных носителей носит название инжекции (впрыскивания) неосновных носителей через p - n переход.

При нарушении нейтральности любого проводника возникают чрезвычайно большие электростатические силы, которые притягивают заряды противоположного знака и восстанавливают нейтральность. Поэтому на величину концентрации неосновных носителей Δn_p возрастет и концентрация основных

$$\Delta p_p = \Delta n_p. \quad (8.67)$$

Однако $p_p \gg \Delta n_p$, и поэтому изменением p_p мы можем пренебречь.

Если ограничиться тем случаем, когда падение напряжения на p - n переходе мало по сравнению с контактной разностью потенциалов V_K между p - и n -областью *), то

$$n_p \ll n_1 \ll p_p. \quad (8.68)$$

Следовательно, убыль концентрации дырок за счет рекомбинации с избыточными электронами будет также ничтожной, поэтому p_p во всей p -области будет выражаться соотношением

$$p_p = \frac{2(2\pi m_p kT)^{3/2}}{h^3} e^{\frac{\mu}{kT}}, \quad (8.69)$$

где μ' — расстояние от уровня химического потенциала до валентной зоны.

*) С другой стороны, мы будем всегда считать, что $eV \gg kT \approx 0,04$ эв и, следовательно, $n_1 \gg n_p$.

С другой стороны, концентрация электронов в p -области согласно (8.68) будет во много раз больше равновесной:

$$n_1 \gg n_p = \frac{2(2\pi m_n kT)^{3/2}}{h^3} e^{\frac{\mu}{kT}}, \quad (8.70)$$

где μ — расстояние от того же уровня Ферми до дна свободной зоны.

Таким образом, условие равновесия $n_p n_n = n_i^2$ в данном случае будет нарушено, так как концентрация неосновных носителей стала много больше равновесной.

В тех случаях (как, например, в рассматриваемом здесь), когда концентрация носителей отстает от равновесной, оказывается чрезвычайно полезным ввести новое понятие — так называемый квазиуровень Ферми μ_K^n , через который неравновесная концентрация выражается так же, как равновесная — через нормальный уровень Ферми:

$$n_1 = \frac{2(2\pi m_n kT)^{3/2}}{h^3} e^{\frac{\mu_K^n}{kT}}. \quad (8.71)$$

Выражение (8.71) и является определением величины μ_K^n ; из (8.71) находим

$$\mu_K^n = kT \ln \frac{n_1}{N_c}, \quad (8.72)$$

где N_c — эффективная плотность состояний в интервале энергий kT зоны проводимости, равная

$$N_c = \frac{2(2\pi m_n kT)^{3/2}}{h^3}. \quad (8.73)$$

Согласно (8.71) и (8.65) в точке O (см. рис. 8.4,а)

$$\mu_K^n = \mu + eV; \quad (8.74)$$

на диффузионной длине согласно (8.66) $n_1(x)$ спадает экспоненциально, приближаясь к n_p . Следовательно, квазиуровень Ферми понижается примерно линейно, приближаясь к нормальному уровню Ферми.

Все сказанное об электронах в p -области можно без изменений отнести к дыркам в n -области, поэтому мы не будем повторять здесь этих рассуждений.

На рис. 8.4,а представлен ход квазиуровней Ферми для дырок и электронов при пропускном направлении приложенного напряжения. Во всех случаях, когда концентрация носителей превышает равновесную, квазиуро-

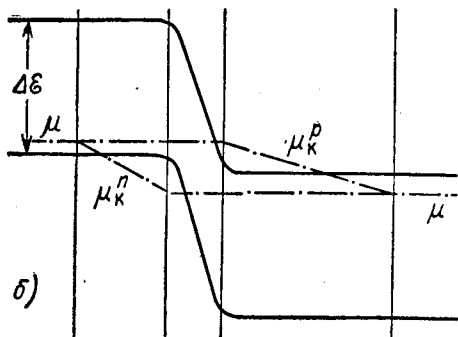
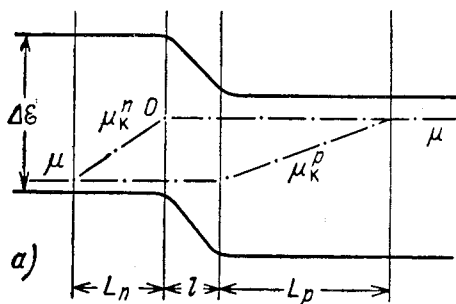


Рис. 8.4. Ход квазиуровней Ферми при напряжении, приложенном в пропускном (а) и запиорном (б) направлениях:

l — область объемного заряда, L_n и L_p — диффузионные длины электронов и дырок.

вень Ферми подходит ближе к соответствующей зоне, чем нормальный уровень Ферми; выражение (8.74) может быть также переписано в форме

$$\mu_K^n = \mu + \ln \frac{n_i}{n_p}. \quad (8.75)$$

Как мы уже упоминали, рассмотренный выше случай, когда концентрация неосновных носителей повышается за

счет их диффузии через p - n переход, называется инжекцией. Существует целый ряд других возможностей повышения концентрации (основных и неосновных) носителей: фотоионизация (см. гл. 9), ударная и электростатическая ионизация; почти во всех случаях для количественного анализа оказывается полезным ввести квазиуровни Ферми.

На рис. 8.4,б представлен случай, когда к p - n переходу приложено напряжение в запиорном направлении. В этом случае концентрация носителей в n -области, имеющих энергию, достаточную для перехода в p -область, падает:

$$n_1 = n_p e^{-\frac{eV}{kT}}. \quad (8.76)$$

Для того чтобы понять, как это отражается на концентрации электронов в p -области, восстановим картину для равновесия и пропускного направления тока.

В случае равновесия ($V = 0$) $n_1 = n_p$, т. е. потоки слева направо и справа налево одинаковы; убыль числа электронов в любой точке p -области за счет рекомбинации и диффузии в n -область компенсируется генерацией и встречной диффузией. При пропускном направлении диффузия электронов из n -области значительно превышает встречную диффузию, концентрация электронов в p -области растет и соответственно растет скорость рекомбинации. За счет увеличившейся скорости рекомбинации неравновесная концентрация падает на диффузионной длине по экспоненциальному закону, постепенно приближаясь к равновесию.

В случае запиорного направления приложенного напряжения и при $eV \gg kT$ (а для запиорного направления это неравенство практически всегда соблюдается) потоком электронов из n -области в p -область в первом приближении можно пренебречь и поток электронов из p -области в n -область ничем не компенсируется. Поэтому концентрация электронов на диффузионной длине в p -области падает по экспоненциальному закону. Так как в сечении, находящемся на расстоянии диффузионной длины L от p - n перехода, электроны из n -области благодаря рекомбинации не доходят в любом случае, то здесь концентрация носителей во всех случаях, а следовательно, и в рассматриваемом, должна остаться приблизительно равновесной; в соответствии с этим на диффузионной длине она будет убывать по закону

$$\Delta n_p(x) = \Delta n_{p0} e^{-\frac{x}{L_n}},$$

где

$$\Delta n_{p0} = n_p \left(e^{-\frac{eV}{kT}} - 1 \right). \quad (8.77)$$

Это явление (т. е. убывание концентрации неосновных носителей за счет диффузии через p - n переход) называется экстракцией.

ВЫПРЯМЛЕНИЕ ПРИ ПОСТОЯННОМ НАПРЯЖЕНИИ

Полученные выражения (8.65) и (8.77) позволяют нам воспроизвести картину выпрямления на p - n переходе.

Рассмотрим сначала пропускное направление тока. В этом случае электроны инжектируются в n -область и движутся в ней, постепенно рекомбинируя с дырками; для того чтобы вычислить электронный ток согласно (8.65), мы должны учесть их диффузию и дрейф в электрическом поле. То же самое можно сказать о дырках в p -области: они движутся, постепенно рекомбинируя с электронами, а оставшиеся инжектируются в n -область.

Однако и на дырки и на электроны действует одно и то же поле, но концентрация электронов в p -области во много раз меньше, чем концентрация дырок, поэтому омическим электронным током в p -области можно пренебречь.

С другой стороны, мы предположили, что $V \ll V_K$, и, следовательно, концентрация инжектированных электронов много меньше концентрации дырок:

$$n_1 \ll p_p. \quad (8.78)$$

Поэтому изменением концентрации дырок за счет рекомбинации с электронами можно пренебречь. Это значит, что градиент концентрации и диффузионный ток дырок ничтожно малы, т. е. дырочный в p -области ток — чисто омический. Таким образом, вырисовывается следующая картина прохождения тока через контакты. Справа вдали от перехода течет практически чисто дырочный ток (так как $n_p \ll p_p$). В области p - n перехода этот ток можно разбить на два:

$$I_p = I'_p + I''_p, \quad (8.79)$$

где I'_p — это та часть дырок, которая рекомбинирует с инжектированными в p -область электронами, и I''_p — другая часть, которая инжектируется в n -область и там рекомбинирует с электронами. Ток I'_p мы можем определить, вычис-

лив диффузионный поток электронов, инжектированных в p -область:

$$I'_p = eD_n \left. \frac{\partial \Delta n(x)}{\partial x} \right|_{x=0} = \frac{eD_n n}{L_n} (e^{\frac{eV}{kT}} - 1). \quad (8.80)$$

Выражение для числа дырок, инжектированных в p -область, можно написать по аналогии с (8.66):

$$\text{при } 0 < x < L_p \quad \Delta p(x) = p_n e^{-\frac{x}{L_p}} (e^{\frac{eV}{kT}} - 1). \quad (8.81)$$

Следовательно,

$$I''_p = eD_p \frac{\partial \Delta p(x)}{\partial x} = \frac{eD_p p_n}{L_p} (e^{\frac{eV}{kT}} - 1). \quad (8.82)$$

Таким образом, полный ток через p - n переход

$$I = I'_p + I''_p = e \left(\frac{D_n n_p}{L_n} + \frac{D_p p_n}{L_p} \right) (e^{\frac{eV}{kT}} - 1). \quad (8.83)$$

Рассмотрим теперь работу p - n перехода в запорном направлении. В случае пропускного направления внешнее напряжение было направлено таким образом, что тянуло носители через p - n переход и инжектировало их в область с противоположным знаком проводимости.

Теперь мы имеем обратную картину, рассмотрим ее на примере электронов и дырок в p -области, а потом перенесем сделанные выводы на n -область. По-прежнему в p -области вдали от p - n перехода течет практически чисто дырочный ток (так как концентрация дырок p_p много больше концентрации электронов n_p), только ток этот теперь течет в направлении от перехода. Точно так же в глубине n -области течет чисто электронный ток и тоже в направлении от перехода. Как уже упоминалось, при тепловом равновесии ($V = 0$) генерация компенсирует рекомбинацию и диффузия носителей слева направо и справа налево также компенсируют друг друга. При запорном направлении этот баланс нарушается: почти все неосновные носители (в данном случае электроны), генерированные в p -область на расстоянии диффузионной длины, уходят в n -область. Согласно (8.77) этот диффузионный ток равен

$$I_n = eD_n \frac{\partial \Delta n}{\partial x} = \frac{eD_n n_p}{L_n} (1 - e^{-\frac{eV}{kT}}). \quad (8.84)$$

Точно так же почти все дырки, созданные в n -области, переходят в p -область; по аналогии с (8.84)

$$I_p = eD_p \frac{\partial \Delta p}{\partial x} = e \frac{D_p p_n}{L_p} (1 - e^{-\frac{eV}{kT}}). \quad (8.85)$$

Таким образом, полный ток через p - n переход в этом случае будет

$$I = I_p + I_n = e \left(\frac{D_n n_p}{L_n} + \frac{D_p p_p}{L_p} \right) (1 - e^{-\frac{eV}{kT}}). \quad (8.86)$$

Формулы (8.83) и (8.86) можно объединить, если току и напряжению для пропускного направления приписать знак плюс, а для запиорного — минус. При неограниченном возрастании запиорного напряжения ток согласно (8.85) стремится к насыщению:

$$I_s = e \left(\frac{D_n n_p}{L_n} + \frac{D_p p_p}{L_p} \right). \quad (8.87)$$

Подставив в выражения для L_p и L_n

$$L_p = \sqrt{D_p \tau_p} \quad \text{и} \quad L_n = \sqrt{D_n \tau_n} \quad (8.88)$$

и избавившись от иррациональности в знаменателе, получим

$$I_s = e \left(\frac{n_p}{\tau_n} L_n + \frac{p_p}{\tau_p} L_p \right), \quad (8.89)$$

т. е. ток насыщения равен полному числу пар, создаваемых на диффузионных длинах в n - и p -области.

С учетом всего сказанного выше работу p - n перехода в запиорном направлении можно обрисовать следующим образом.

На диффузионной длине в p -области создается в одну секунду следующее число пар электронов:

$$I_{sn} = \frac{n_p}{\tau_n} L_n; \quad (8.90)^*$$

большая часть созданных электронов

$$I_{sn} (1 - e^{-\frac{eV}{kT}}) \quad (8.91)$$

* $n_p/\tau_n = p_p/\tau_p^p$, где τ_p^p — время жизни дырки в p -области, точно так же $p_n/\tau_p = n_n/\tau_n^n$, где τ_n^n — время жизни электрона в n -области.

переходит в n -область и там превращается в омический ток, меньшая —

$$I_{sp} e^{-\frac{eV}{kT}} \quad (8.92)$$

рекомбинирует с дырками в p -области.

Точно так же на диффузионной длине в n -области создается в секунду

$$I_{sp} = \frac{p_n}{\tau_p} L_p \quad (8.93)$$

пар, причем $I_{sp} (1 - e^{-\frac{eV}{kT}})$ дырок диффундирует в p -область и превращается в омический ток, а остальные рекомбинируют.

ВЫПРЯМЛЕНИЕ НА p - n ПЕРЕХОДЕ ПРИ ПЕРЕМЕННОМ НАПРЯЖЕНИИ

Предположим, что напряжение V в (8.84) на p - n переходе состоит из двух частей, отвечающих постоянному и переменному току:

$$V = V_0 + v_1 e^{i\omega t}, \quad (8.94)$$

где сигнал переменного тока v_1 мы будем считать настолько малым, что можно при разложении экспоненты в (8.94) в ряд воспользоваться линейным приближением (т. е. $ev_1 \ll kT$). Тогда согласно (8.64)

$$n_1 = n_p e^{\frac{eV_0}{kT}} \left(1 + \frac{ev_1}{kT} e^{i\omega t} \right). \quad (8.95)$$

Представим эту концентрацию в виде трех слагаемых:

$$n_1 = n_p + n_0 + n_{\sim} e^{i\omega t}, \quad (8.96)$$

где

$$n_0 = n_p \left(e^{\frac{eV_0}{kT}} - 1 \right) \quad (8.97)$$

и

$$n_{\sim} = \frac{en_p v_1}{kT} e^{\frac{eV_0}{kT}}. \quad (8.98)$$

До тех пор пока $n_1 \ll p_p$, мы можем считать время жизни τ_n и коэффициент диффузии электронов D_n

в p -области не зависящими от V и рассматривать их в уравнении непрерывности как постоянные величины *).

Напишем уравнение непрерывности для электронов в p -области. Подставляя в уравнение непрерывности (8.51) выражение для тока (8.53) и пренебрегая в соответствии с сказанным выше омическим членом, получим

$$\frac{\partial n}{\partial t} = \frac{n_p - n(x)}{\tau_n} - D_n \frac{\partial^2 n}{\partial x^2}. \quad (8.99)$$

Решение этого уравнения при учете граничных условий: при $x=0$ $n=n_0$ и при $x=L_n$ $n=n_p$, имеет вид

$$n = n_p + n_0 e^{-\frac{x}{L_n}} + n_{\sim} \exp \left[i\omega t - \frac{x(1+i\omega\tau_n)^{\frac{1}{2}}}{L_n} \right]. \quad (8.100)$$

Выражение (8.100) позволяет вычислить диффузионный ток электронов в n -области:

$$I_n = eD_n \frac{\partial n}{\partial x} = \frac{en_0 D_n}{L_n} + \frac{en_{\sim} D_n e^{i\omega t} (1+i\omega\tau_n)^{\frac{1}{2}}}{L_n}. \quad (8.101)$$

Подставляя n_0 из (8.97) в (8.101), получаем прежнее выражение (8.84) для компоненты I , отвечающей постоянному току:

$$I_{n0} = \frac{en_p D_n}{L_n} \left(e^{\frac{eV_0}{kT}} - 1 \right). \quad (8.102)$$

Подставляя в (8.101) выражение для n_{\sim} , получаем выражение для переменной компоненты

$$I_{n\sim} = \frac{en_p u_n}{L_n} e^{\frac{eV_0}{kT}} (1+i\omega\tau_n)^{\frac{1}{2}} v_1 e^{i\omega t} \quad (8.103)$$

или

$$I_{n\sim} = (G_{\text{дин}} + iS_{\text{дин}}) v_1 e^{i\omega t} = A_n v_1 e^{i\omega t}, \quad (8.104)$$

где A_n — полная электронная проводимость, состоящая из вещественной части $G_{\text{дин}}$ и мнимой $S_{\text{дин}}$.

Полученное выражение позволяет сделать два важных вывода:

1. Несмотря на то, что мы не учитывали обычную электростатическую емкость, в выражении (8.104) появилась

*) До сих пор мы молчаливо пользовались этим предположением.

емкостная составляющая тока (опережающая напряжение на 90°). Качественно ее происхождение можно объяснить следующим образом. Согласно (8.57) и (8.58) заряд на p - n переходе зависит от приложенного к нему напряжения (т. е. заряд будет меняться с изменением напряжения); по отношению к внешней цепи это изменение заряда будет проявляться как дополнительная (по сравнению с электростатической) емкость. Эта емкость называется диффузионной емкостью p - n перехода, которая наряду с электростатической ограничивает применение p - n переходов на высоких частотах, но, с другой стороны, открывает целый ряд важных применений, связанных с ее зависимостью от напряжения.

2. Если частота приложенного напряжения мала и

$$\omega\tau_n \ll 1,$$

то вещественная составляющая проводимости

$$G_{n0} = en_p e^{\frac{eV_0}{kT}} \frac{u}{L_p} = \frac{e(n_p + n_0)}{L} \quad (8.105)$$

и при $V \rightarrow 0$

$$G_{n0} = \frac{en_p u_n}{L_p} = \frac{\sigma_n^p}{L_n}, \quad (8.106)$$

где σ_n^p — электронная проводимость в p -области, т. е. сопротивление на 1 см^2 p - n перехода будет равно сопротивлению электрическому току, вызванному малой n_p концентрацией электронов в p -области.

Точно так же дырочная проводимость (мы проводим ее здесь по аналогии без вывода)

$$G_{0p} = \frac{ep_n u_p}{L_n} = \frac{\sigma_p^n}{L_n} \quad (8.107)$$

будет определяться дырочной электропроводностью в n -области.

Приведенные выше выражения (8.106) и (8.107) позволяют сделать важный вывод: дифференциальное (динамическое) сопротивление p - n перехода намного больше, чем сопротивление собственного полупроводника (так как n_p и $p_n \ll n_i$).

Нетрудно выразить активную электропроводность G_{n0} p - n перехода через ток насыщения I_{ns} . Разлагая выражение (8.102) в ряд, получаем

$$I = I_{ns} \frac{eV_0}{kT}. \quad (8.108)$$

Сопоставив (8.105) и (8.108), находим

$$G_{0n} = I_{ns} \frac{e}{kT}$$

и точно также

$$G_{0p} = I_{ps} \frac{e}{kT}. \quad (8.109)$$

Нетрудно также получить выражение для емкостной составляющей проводимости. Согласно (8.104)

$$A_n = G_n + iS_n = (1 + i\omega\tau_n)^{1/2} G_{n0} e^{\frac{eV_0}{kT}} \quad (8.110)$$

при низких частотах, когда $\omega \ll 1/\tau_n$; разлагая выражения в скобках в (8.110) в ряд, получаем

$$G_n + iS_n = G_{n0} e^{\frac{eV_0}{kT}} + i\omega \frac{\tau_n}{2} g_n e^{\frac{eV_0}{kT}}. \quad (8.111)$$

Таким образом, емкость на единицу площади

$$C_n = \frac{\tau_n}{2} G_{n0} e^{\frac{eV_0}{kT}}. \quad (8.112)$$

При $V_0 = 0$, подставляя в (8.112) выражения G_{n0} и τ_p и воспользовавшись соотношением Эйнштейна

$$u_n = \frac{eD_n}{kT}, \quad (8.113)$$

получим

$$C_n = \frac{e^2 n_p L_n}{2kT}. \quad (8.114)$$

Полученное выражение имеет простой физический смысл. При небольшом приложенном напряжении ΔV изменение электронного заряда перехода будет

$$\Delta Q = \frac{1}{2} e n_p \left(e^{\frac{e\Delta V}{kT}} - 1 \right) L_n \approx \frac{e^2 n_p L_n}{2kT} \Delta V. \quad (8.115)$$

Поделив ΔQ на ΔV , получим точное выражение (8.114) для диффузионной емкости p - n перехода. Аналогичные выражения могут быть получены для дырочной активной проводимости и емкости. При низких частотах G_n и C_n ведут себя, таким образом, как проводимость и емкость, включенные параллельно друг другу. При высоких частотах, однако, p - n переход ведет себя совершенно иначе.

Как следует из (8.103), при $\omega\tau_n \gg 1$ и $V_0 = 0$

$$G_n - S_n = \sqrt{\frac{\tau_n}{2}} G_{n0} \sqrt{\omega} = \frac{en_p\mu_n \sqrt{\tau_n}}{\sqrt{D_n\tau_n}} \sqrt{\omega} = \frac{en_p\mu_n}{\sqrt{D_n}} \sqrt{\omega}. \quad (8.116)$$

Как видно из (8.116), и активная и реактивная составляющие перестают зависеть от времени жизни и определяются только диффузией электронов. Качественно это понятно: при $\tau_n \gg 1/\omega$ электрон совершит очень много колебаний, прежде чем будет захвачен, и процессы рекомбинации перестают играть существенную роль. При еще больших частотах, когда частота колебаний становится больше числа столкновений электрона, надо учитывать его инерцию (т. е. эффективную массу), что приводит к индуктивному сдвигу фазы — этот вопрос рассмотрен в следующей главе.

ВЫПРЯМЛЕНИЕ НА ТОЛСТОМ p - n ПЕРЕХОДЕ

Развитые выше представления были основаны на предположении, что рекомбинацией в слое объемного заряда можно пренебречь; это действительно имеет место, когда толщина слоя много меньше диффузионной длины дырок и электронов. Такой p - n переход называется тонким.

Но тонкий p - n переход оказывается прозрачным для туннельного эффекта, поэтому, когда в запиорном направлении прикладывается напряжение, большее

$$V_k = \mu_n + \mu'_p, \quad (8.117)$$

где μ_n — расстояние от уровня химического потенциала до дна зоны проводимости в n -области и μ'_p — расстояние от уровня химического потенциала до верхнего края зоны в p -области, то становятся возможными туннельные переходы из валентной зоны p -проводника в свободную зону n -проводника; при этом запиорный ток начинает экспоненциально возрастать.

Кроме этого, напряжение поля в тонком запиорном слое при запиорном напряжении настолько возрастает, что становится возможной ударная ионизация, которая при размножении пар ведет к пробоему выпрямителя.

Для силовых выпрямителей одной из важнейших характеристик является высокое обратное напряжение, и эти

явления недопустимы. Поэтому в данном случае применяются толстые p - n переходы, в которых уже нельзя пренебречь рекомбинацией в слое объемного заряда.

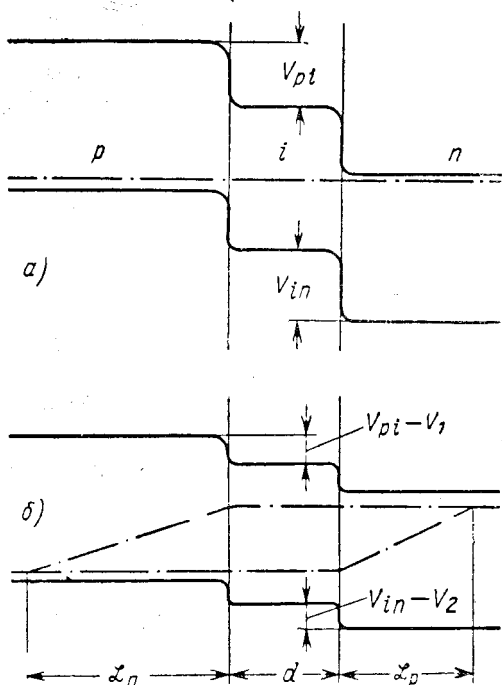


Рис. 8.5. Ступенчатый (p - i - n) переход: а — в состоянии равновесия; б — при напряжении, приложенном в пропускном направлении.

Мы рассмотрим теорию толстого перехода на простейшем примере ступенчатого p - i - n перехода, в котором p -и n -области разделены областью собственной проводимости (см. рис. 8.5, а и б). На рис. 8.5, а изображен такой переход в состоянии равновесия. На рис. 8.5, б представлена зонная диаграмма этого перехода, когда к нему приложено напряжение V в пропускном направлении, которое распределяется каким-то образом между p - i и i - n переходами *). При

*) Выше было показано, что сопротивление выпрямляющих переходов много больше, чем собственного полупроводника, поэтому падением напряжения в i -области мы в первом приближении можем пренебречь.

этом полный ток через p - n переход будет состоять из двух частей:

$$I = I_p + I_c, \quad (8.118)$$

где I_p — рекомбинационный ток, обусловленный рекомбинацией электронов и дырок в i -области; I_c — «сквозной» ток, обусловленный рекомбинацией электронов, инжектируемых в p -область, и дырок, инжектируемых в n -область.

Как будет видно из дальнейшего, выражение для сквозного тока ничем не отличается от соответствующего выражения (8.118) для тонкого p - n перехода.

Мы начнем с вычисления рекомбинационного тока I_p . При приложении напряжения V_1 к первому переходу число электронов на границе n - и i -области возрастет на величину

$$\Delta n = (n_1 - n_i) = n_i \left(e^{\frac{eV_1}{kT}} - 1 \right). \quad (8.119)$$

Аналогично можно вычислить увеличение концентрации дырок на границе p - и i -области:

$$\Delta p = (p_1 - n_i) = p_i \left(e^{\frac{eV_2}{kT}} - 1 \right). \quad (8.120)$$

Если ограничиться рассмотрением того случая, когда толщина (d) области i много меньше диффузионной длины, то эта же концентрация избыточных электронов (и дырок) приблизительно сохранится во всей i -области; следовательно, общее число электронов, рекомбинирующих в 1 сек в i -области, будет

$$I_{pn} \approx \frac{\Delta n}{\tau_n^i} d \approx \frac{n_i}{\tau_n^i} d \left(e^{\frac{eV_1}{kT}} - 1 \right). \quad (8.121)$$

Точно так же можно вычислить рекомбинационный ток дырок, инжектируемых из p -области в i -область:

$$I_{pp} = \frac{\Delta p}{\tau_p^i} d = \frac{p_i}{\tau_p^i} d \left(e^{\frac{eV_2}{kT}} - 1 \right). \quad (8.122)$$

Из условия стационарности $I_{pn} = I_{pp} = I_p$. Но $\Delta n = \Delta p$ — это следует из условия нейтральности; следовательно, из (8.121) получаем

$$\tau_n^i = \tau_p^i. \quad (8.123)$$

Мы поэтому в дальнейшем индексы p и n будем опускать. Учитывая что $n_i = p_i$ (что также следует из условия

нейтральности в состоянии равновесия), из (8.121) и (8.122) получаем

$$V_1 = V_2, \quad (8.124)$$

и так как $V_1 + V_2 = V$ равно приложенному напряжению, то

$$V_1 = V_2 = \frac{V}{2}.$$

Таким образом, мы получаем окончательное выражение для рекомбинационного тока:

$$I_p = \frac{en_i d}{\tau_i} (e^{\frac{eV}{2kT}} - 1). \quad (8.125)$$

Вычислим теперь «сквозной» ток. Если бы число электронов i -области сохранилось равновесным, но к i - p переходу было приложено напряжение V_2 , то за счет инжекции из i -области число электронов в p -области на границе с i -областью возросло бы до величины

$$n_2 = n_p e^{\frac{eV_2}{kT}} = n_i e^{-\frac{e(V_{pi} - V_2)}{kT}}, \quad (8.126)$$

где V_{pi} — контактная разность потенциалов между i - и p -областью. Но число электронов в i -области возрастает благодаря инжекции из n -области:

$$n_i = n_i^0 e^{\frac{eV_1}{kT}}. \quad (8.127)$$

Следовательно, соответственно должно возрасти число электронов, инжектируемых в p -область:

$$n_2 = n_p e^{\frac{e(V_1 + V_2)}{kT}} = n_p e^{\frac{eV}{kT}}. \quad (8.128)$$

По мере удаления от i - p перехода число избыточных электронов в p -области $\Delta n_2(x)$ убывает по экспоненциальному закону

$$\Delta n_2(x) = n_p (e^{\frac{eV}{kT}} - 1) e^{-\frac{x}{L_n}}. \quad (8.129)$$

Следовательно, диффузионный ток электронов в p -области

$$I_{cn} = eD_n \frac{\partial n}{\partial x} = \frac{eD_n n_p}{L_n} (e^{\frac{eV}{kT}} - 1). \quad (8.130)$$

Точно так же можно получить выражение для сквозного дырочного тока:

$$I_{cp} = \frac{eD_p p_n}{L_p} \left(e^{\frac{eV}{kT}} - 1 \right). \quad (8.131)$$

Таким образом, согласно (8.125), (8.130) и (8.131) полный ток через p - n переход будет

$$I = I_p + I_c = I_p + I_{cn} + I_{cp} = \frac{en_i d}{\tau_i} \left(e^{\frac{eV}{2kT}} - 1 \right) + \left(\frac{eD_n n_p}{L_n} + \frac{eD_p p_n}{L_p} \right) \left(e^{\frac{eV}{kT}} - 1 \right). \quad (8.132)$$

Нетрудно показать, что выражение (8.132) сохраняет силу и для запертого направления тока. Можно показать, что при большой толщине перехода, в мощных выпрямителях, основную роль играет рекомбинационный ток.

ТУННЕЛЬНЫЕ ДИОДЫ

Нам остается рассмотреть противоположный случай — очень тонкого запертого слоя, прозрачного для туннельного эффекта. На рис. 8.6, *а, б* и *в* изображена серия таких p - n переходов (с постепенно возрастающей концентрацией основных носителей).

В последнем случае (рис. 8.6, *в*) и дырки и электроны находятся в вырожденном состоянии, уровень химического потенциала проходит внутри соответствующей зоны. На рис. 8, *г, д, е* изображены вольтамперные характеристики тех же p - n переходов.

Остановимся сначала на рис. 8.6, *а, б, г, д*. Как видно из этих рисунков, при увеличении концентрации примесей и носителей *) уменьшается критическое напряжение V_0 , при котором запертый ток благодаря туннельному эффекту начинает резко возрастать. В случае *д* ток в запертом направлении начинает возрастать более резко и при меньших напряжениях, чем в пропускном направлении, т. е. по существу запертое направление становится пропускным, а пропускное — запертым.

*) Т. е. по мере приближения уровня химического потенциала ко дну зоны проводимости в n -области и ко дну валентной зоны в p -области.

Диоды, основанные на этом явлении, называются обращенными. По сравнению с обычным, обращенный диод обладает тремя преимуществами:

1) возможностью применения для детектирования малых сигналов, поскольку обратный ток начинает возрастать при очень малых напряжениях;

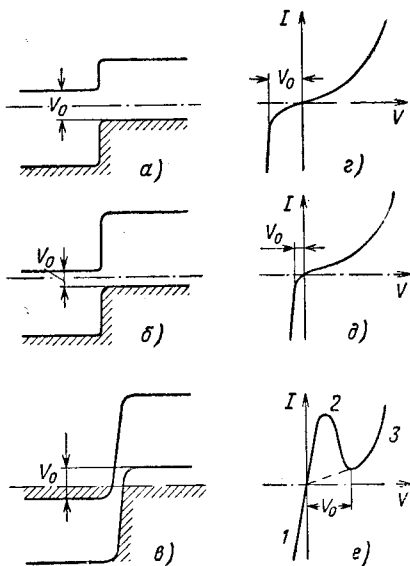


Рис. 8.6. Зонные диаграммы и вольтамперные характеристики тонких $p-n$ переходов (прозрачных для туннельного эффекта).

2) возможностью применения его в ключевых схемах, так как обратный ток возрастает очень резко;

3) возможностью работы на высоких частотах, так как для туннельного эффекта требуется ничтожное время, а частотные возможности обычных диодов ограничены инерционностью рекомбинации и диффузии носителей.

Еще более своеобразный вид имеет характеристика $p-n$ перехода, изображенная на рис. 8.6, *в*; здесь носители по обе стороны от перехода находятся в вырожденном состоянии. Такие диоды называются туннельными. На рис. 8.6, *в* напряжение, приложенное к переходу, равно нулю; несмотря на то, что переход и при этом прозрачен для туннельного эффекта, результирующий ток в этом случае равен нулю; это обусловлено следующими причинами.

Переходить и справа налево, и слева направо могут только электроны, находящиеся выше уровня Ферми слева,

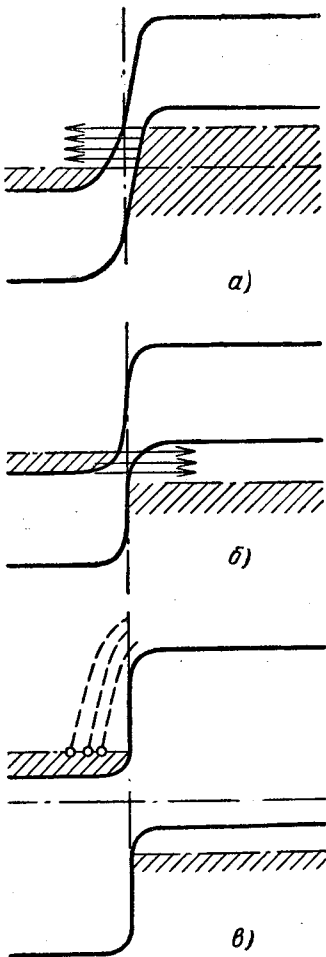
и дырки, находящиеся ниже уровня Ферми справа; число и тех и других (относительно) очень мало, поэтому каждый из этих токов очень мал. Кроме того, эти токи равны друг другу, поэтому результирующий ток равен нулю.

Картина резко меняется, когда к переходу приложено напряжение. На рис. 8.7, а представлена зонная схема перехода, когда к нему приложено напряжение в запиорном направлении. В этом случае в n -область могут переходить все электроны из p -области, энергия которых выше уровня Ферми в n -области, и ток очень резко возрастает. Участок характеристики туннельного диода 1 (на рис. 8.6, е) подобен характеристике обращенного с той лишь разницей, что возрастание тока начинается сразу же с напряжения, равного нулю.

Чем больше запиорное напряжение, тем больше число электронов, которые могут переходить благодаря туннельному эффекту из p -области в n -область, и тем больше обращенный ток. Совершенно поиному обстоит дело при противоположной полярности напряжения. На рис. 8.7, б и в изображен p - n переход при двух напряжениях, приложенных в пропускном направлении. Как видно из рис. 8.7, б,

Рис. 8.7. Зонная диаграмма туннельного диода при напряжении, различном по величине и направлению.

в этом случае в валентную зону p -области могут переходить



Все электроны из свободной зоны n -области, и туннельный ток в этом случае максимален (участок 2 на рис. 8.6, e).

Напротив на участке 3 (рис. 8.6, e) свободная зона поднялась настолько высоко, что напротив нее в p -области находится запрещенная зона и туннельный ток в этом случае отсутствует. Однако при этом потенциальный барьер уже значительно упал и начинает возрастать обычный диффузионный ток через p - n переход. Поэтому вольтамперная характеристика туннельного диода имеет вид, представленный на рис. 8.6, e .

Чрезвычайно большое значение имеет в данном случае наличие участка с отрицательным сопротивлением ($dV/dI < 0$), позволяющим очень широко использовать туннельные диоды для усиления, генерации и в схемах переключения. Благодаря своей малой инерционности туннельные диоды с успехом используются в СВЧ диапазоне. Меняя степень легирования по обе стороны от p - n перехода, можно в широких пределах менять вольтамперную характеристику и частотные параметры туннельных диодов.

ОПТИЧЕСКИЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

9.1. ПОГЛОЩЕНИЕ СВЕТА

Распространение электромагнитного излучения в проводящей среде описывается уравнением затухающей волны

$$u(x, t) = U_0 e^{i\omega(t - \frac{x}{v})}, \quad (9.1)$$

где u — мгновенное значение напряженности электрической или магнитной составляющей электромагнитного поля; ω — частота и v — комплексная скорость ($v = c/N$, где N — комплексный показатель преломления и c — скорость света в пустоте).

Показатель преломления выражается через диэлектрическую постоянную κ , электропроводность σ и магнитную проницаемость μ (на данной частоте ω):

$$N^2 = \mu\kappa - \frac{i\mu\sigma}{\omega}. \quad (9.2)$$

Полагая

$$N = n - ik, \quad (9.3)$$

т. е. выделив в N вещественную и мнимую часть, получаем

$$n^2 - k^2 = \mu\kappa \quad \text{и} \quad 2nk = \frac{\sigma\mu}{\omega}. \quad (9.4)$$

Согласно (9.1) и (9.4)

$$u(x, t) = U_0 e^{i\omega t} e^{-\frac{i\omega n x}{c}} e^{-\frac{\omega k x}{c}}. \quad (9.5)$$

Согласно (9.5) интенсивность света (\mathcal{J}) убывает по закону

$$\mathcal{J} = \mathcal{J}_0 e^{-Kx}, \quad (9.6)$$

где

$$K = \frac{2\omega k}{c} = \frac{4\pi k}{\lambda}. \quad (9.7)$$

Для того чтобы найти частотную зависимость коэффициента поглощения и преломления, мы должны изучить взаимодействие света с различными группами электронов.

СВЯЗАННЫЕ ЭЛЕКТРОНЫ

Напомним классическое уравнение движения электрона

$$m \frac{d^2x}{dt^2} + mg \frac{dx}{dt} + m\omega_0^2 x = eEe^{i\omega t}. \quad (9.8)$$

Решение этого уравнения

$$x = \frac{\frac{eE}{m}}{\omega_0^2 - \omega^2 + i\omega g} = \chi E, \quad (9.9)$$

где χ — атомная поляризуемость.

Поляризация единицы объема

$$P = eN_{\text{эл}}\chi, \quad (9.9a)$$

где $N_{\text{эл}}$ — число электронов в 1 см^3 , и диэлектрическая постоянная

$$\kappa = 1 + 4\pi\chi. \quad (9.10)$$

Полагая для связанных электронов $\sigma = 0$ и $\mu = 1$, из (9.2) находим $N^2 = \kappa$; следовательно, согласно (9.9) и (9.10)

$$(n - ik)^2 = \frac{4\pi N_{\text{эл}} e^2}{m(\omega_0^2 - \omega^2 + i\omega g)} + 1, \quad (9.11)$$

откуда

$$n^2 - k^2 - 1 = \frac{4\pi N_{\text{эл}} e^2}{m} \left[\frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \omega^2 g^2} \right] \quad (9.12)$$

и

$$2nk = \frac{4\pi N_{\text{эл}} e^2}{m\kappa_0} \frac{\omega g}{(\omega_0^2 - \omega^2)^2 + \omega^2 g^2}. \quad (9.13)$$

Выражения (9.12) и (9.13) показывают, что при $\omega_0 \gg \omega$ $n \approx 1$ и $k = 0$; это значит, что электроны, расположенные на глубоких уровнях (рентгеновских термов), практически не участвуют в преломлении и поглощении видимого и инфракрасного излучения.

Вблизи $\omega = \omega_0$ имеется максимум поглощения, а затем при дальнейшем увеличении частоты K уменьшается, оставаясь все же больше, чем в длинноволновой области спектра. Квантовомеханическая теория дисперсии дает соотношения, подобные (9.12) и (9.13). Согласно этой теории атом можно описать набором независимых осцилляторов, каждый из которых характеризуется своей частотой ω_i , коэффициентом затухания g_i и силой f_i . Уравнения (9.12) и (9.13) заменяются:

$$n^2 - k^2 - 1 = 4\pi \sum \frac{N_{\text{эл}} e^2 f_i / m (\omega_i^2 - \omega^2)}{(\omega_i^2 - \omega^2)^2 + \omega^2 g_i^2}, \quad (9.14)$$

$$2nk = 4\pi \sum \frac{N_{\text{эл}} e^2 f_i / m \omega g_i}{(\omega_i^2 - \omega^2)^2 + \omega^2 g_i^2}, \quad (9.15)$$

где суммирование проводится по всем осцилляторам.

ПОГЛОЩЕНИЕ СВЕТА СВОБОДНЫМИ НОСИТЕЛЯМИ

В сильнолегированных полупроводниках начинает играть существенную роль поглощение и преломление света за счет колебаний свободных носителей в электромагнитном поле. При этом следует учитывать, что совершенно свободные носители не могут поглотить энергию излучения — их колебания лишь вносят свой вклад в коэффициент преломления.

Однако зонный электрон, движущийся (или колеблющийся) в электромагнитном поле, непрерывно испытывает столкновения, и если частота их достаточно велика, то значительная часть электромагнитной энергии превращается в тепловую. Этим объясняется поглощение света свободными носителями и его увеличение по мере приближения к низким частотам. Эти рассуждения нетрудно оформить количественно. Количество электромагнитной энергии в единице объема

$$\mathcal{E} = \frac{E^2}{4\pi}. \quad (9.16)$$

Коэффициент поглощения определяется соотношением $\mathcal{J} = \mathcal{J}_0 e^{-Kx}$, откуда

$$\frac{\partial \mathcal{J}}{\partial x} = -K\mathcal{J} \quad \text{и} \quad K = -\frac{\partial \mathcal{J}}{\mathcal{J} \partial x}, \quad (9.17)$$

но $\partial \mathcal{J} / \partial x$ — это мощность, поглотившаяся в единице объема и равная

$$\frac{\partial \mathcal{J}}{\partial x} = P = jE = \sigma E^2, \quad (9.18)$$

$$\mathcal{J} = \mathcal{E}c = \frac{cE^2}{4\pi}. \quad (9.19)$$

Следовательно,

$$K = \frac{\sigma E^2}{\frac{cE^2}{4\pi}} = \frac{4\pi\sigma}{c}, \quad (9.20)$$

где $\sigma = eun$, n — концентрация носителей и u — их подвижность (на данной частоте).

Таким образом, нахождение коэффициента поглощения сводится к определению частотной зависимости подвижности. Для решения последней задачи напишем уравнение движения электрона в электрическом поле:

$$m\dot{\bar{v}} - \gamma\bar{v} = eE. \quad (9.21)$$

Здесь \bar{v} — скорость дрейфа в электрическом поле E .

Второй член слева в (9.21) описывает потери энергии электрона при столкновениях. При постоянном поле $\dot{\bar{v}} = 0$ и $\bar{v} = eE/\gamma$; но, с другой стороны, мы знаем, что в этом случае $\bar{v} = uE = (e/m)\tau E$; сопоставляя эти два выражения для \bar{v} , находим, что $\gamma = m/\tau$ или $\gamma = \nu m$, где ν — число столкновений в секунду. Переходя теперь к переменному полю и полагая $E = E_0 \sin \omega t$, найдем из (9.21)

$$\bar{v} = \frac{e}{m} \frac{1}{\nu^2 + \omega^2} (\nu \sin \omega t + \omega \cos \omega t) E_0 \quad (9.22)$$

и

$$j = en\bar{v} = \frac{e^2 n}{m} \frac{\nu}{\nu^2 + \omega^2} E + \frac{e^2 n}{m} \frac{1}{\nu^2 + \omega^2} \frac{\partial E}{\partial t}. \quad (9.23)$$

Второй член в (9.23) не создает потерь (поглощения), так как он сдвинут на $\pi/2$ по фазе по отношению к напряжению (т. е. для него $j\bar{E} = 0$). Таким образом,

$$K = \frac{4\pi e^2 n \nu}{cm(\nu^2 + \omega^2)} = \frac{4\pi\sigma_0}{c} \frac{1}{1 + \left(\frac{\omega}{\nu}\right)^2}, \quad (9.24)$$

где σ_0 — электропроводность на постоянном токе.

В том случае, когда мы имеем дело с несколькими сортами носителей, выражение (9.24) должно быть заменено суммой по всем сортам носителей.

Дифференциальное уравнение (9.22) было решено в предположении, что τ , время релаксации, не зависит от скорости v (или, что то же самое, энергии носителей). Если отбросить это ограничение и положить, как мы это делали раньше, $l \sim \varepsilon^r$, то в выражении для коэффициента поглощения появляется дополнительный множитель

$$A_r = \frac{\Gamma(r+2) \Gamma(2-r)}{\left[\Gamma\left(\frac{5}{2}\right) \right]^2}, \quad (9.25)$$

где Γ — гамма-функция от соответствующих аргументов; например, при рассеянии на ионах примеси $A_r = 3,4$.

МЕЖЗОННЫЕ ПЕРЕХОДЫ

Согласно формуле (2.53) вероятность перехода электрона из состояния ψ_1 в состояние ψ_2 под действием возмущения пропорциональна матричному элементу возмущения

$$M_{1,2} = \int \psi_1 V \psi_2 dv. \quad (9.26)$$

В рассматриваемом нами случае ψ_1 и ψ_2 — модулированные плоские волны, описывающие состояния электрона до и после перехода, V — энергия возмущения, т. е. световой волны:

$$\psi_1 = u_{k_1} \exp\left(2\pi i k_1 r - \frac{\mathcal{E}_1}{h} t\right), \quad (9.27)$$

$$\psi_2 = u_{k_2} \exp\left(2\pi i k_2 r - \frac{\mathcal{E}_2}{h} t\right) \quad (9.28)$$

и V — энергия электрона в переменном электромагнитном поле, равная

$$V = V_0 \exp[i(qr - \omega t)]. \quad (9.29)$$

Подставив выражения (9.27), (9.28) и (9.29) в (9.26), нетрудно убедиться, что в нем появятся множители

$$\exp[i(k_1 - k_2 - q)r] \text{ и } \exp\left[\frac{i(\mathcal{E}_1 - \mathcal{E}_2 - h\omega)t}{h}\right],$$

которые очень быстро периодически меняются во времени и пространстве за исключением случая, когда

$$k_1 - k_2 = q \quad (9.30)$$

$$\mathcal{E}_1 - \mathcal{E}_2 = h\omega. \quad (9.31)$$

Умножив и левую и правую части (9.30) на h , перепишем его в виде

$$p_1 - p_2 = p_\Phi. \quad (9.32)$$

Таким образом, из теории переходов квантовой механики непосредственно вытекают законы сохранения энергии и импульсов (9.32). Импульс световой волны (фотона)

$$p_\Phi = \frac{h\omega}{c} \quad (9.33)$$

ничтожно мал по сравнению с импульсом электрона, поэтому из (9.30) непосредственно вытекает очень важное следствие:

$$k_1 \approx k_2. \quad (9.34)$$

Таким образом, под действием электромагнитного излучения возможны только «вертикальные» переходы, т. е. переходы, представленные на рис. 9.1, *a* в пространстве импульсов вертикальными стрелками.

Соотношение (9.34) является дополнением принципа Франка — Кондона, согласно которому под действием света возможны только вертикальные переходы. Это вытекает из того, что время взаимодействия с фотоном очень мало и поэтому электрон не успевает изменить за это время своего положения.

Соотношения (9.31) и (9.32) для свободного электрона не могут быть удовлетворены одновременно; это значит, что внутрizonные переходы могут происходить лишь при участии третьего тела. При наличии третьего тела (фонона, примесного атома или другого дефекта) в выражение (9.32) должна быть включена соответствующая энергия возмущения и в (9.31) и (9.32) появятся дополнительные члены, описывающие изменение импульса и энергии третьего тела (фонона или примесного атома). Таким образом, мы получаем два уравнения с тремя неизвестными, которые имеют бесчисленное число решений. Такие переходы (с участием фонона и фотона), при которых импульс электрона меняется, изображаются на энергетической диаграмме наклон-

ными стрелками и поэтому называются непрямыми *) (см. рис. 9.1, б, в, г).

С точки зрения сказанного выше, на первый взгляд, кажется, что не прямые переходы более вероятны, чем

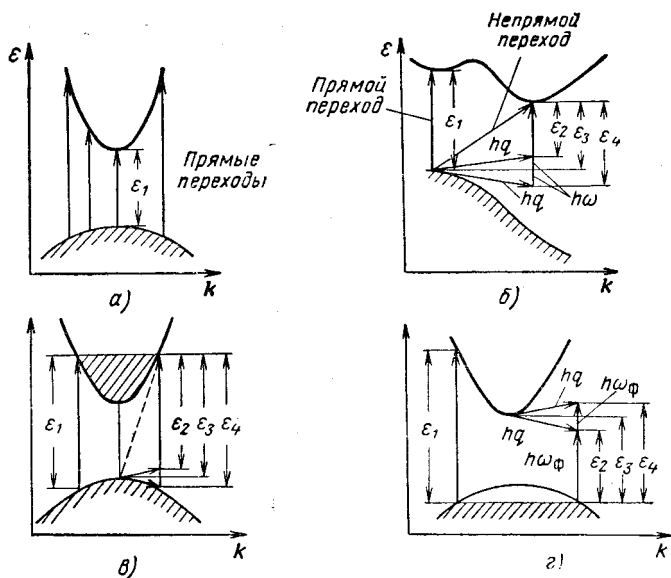


Рис. 9.1. Схема различного расположения зон в k -пространстве, иллюстрирующая прямые и не прямые переходы:

а — прямые межзонные переходы; (ϵ_1 — минимальная энергия фотона, определяющая красную границу полосы собственного поглощения и фотоэффекта); б — не прямые межзонные переходы с участием фононов при смещенных экстремумах; в — не прямые межзонные переходы при вырожденном электронном газе; г — не прямые межзонные переходы при вырожденном дырочном газе.

прямые. В действительности это неверно. Можно показать, что вероятность их значительно меньше, поскольку при вычислении вероятности перехода появляются еще множители типа $\int \Phi^2 dg$, связанные, например, с интегрированием по какой-то части «фононного конфигурационного пространства»; эти множители всегда меньше единицы.

*) В гл. 4 мы показали, что соударения электрона с фононом почти упруги, т. е. что в этом случае наоборот, изменение энергии электрона очень мало. При тройном соударении электрон — фотон — фонон одновременно меняются и энергия и импульс электрона.

Качественно это объясняется тем, что вероятность такого процесса так же мала, как одновременное столкновение трех частиц. Поэтому непрямые переходы начинают играть существенную роль в тех случаях, когда прямые переходы по какой-либо причине невозможны. Это иллюстрируется рис. 9.1, б, на котором представлен энергетический спектр полупроводника, в котором экстремумы валентной и свободной зоны смещены относительно друг друга. В случае, представленном на этом рисунке, прямые межзонные переходы возможны лишь при энергиях фотонов $\epsilon > \epsilon_1$, в то время как переходы с одновременным поглощением фона возможны начиная с энергии фотона ϵ_2 , переходы с испусканием фона — начиная с энергии ϵ_3 и, наконец, забросы под действием одного только теплового движения — начиная с энергии ϵ_4 . Как видно из рисунка, все четыре энергии активации различны:

$$\epsilon_1 > \epsilon_3 > \epsilon_4 > \epsilon_2.$$

Второй случай, когда непрямые переходы начинают играть существенную роль, проиллюстрирован рис. 9.1, в; на этом рисунке представлена зонная диаграмма полупроводника, в которой электронный газ в зоне проводимости находится в вырожденном состоянии. Здесь сохранены те же обозначения, что и на предыдущем рисунке, и имеет место то же соотношение между различными энергиями активации. На рис. 9.1, г представлена аналогичная ситуация для вырожденного дырочного полупроводника.

К сказанному выше следует добавить, что все три упомянутых выше типа поглощения (за счет прямых переходов, с захватом и испусканием фононов) обладают разной температурной зависимостью, так как вероятность прямых переходов непосредственно не связана с температурой, вероятность переходов с поглощением фононов пропорциональна числу последних, а переходы с излучением фона могут происходить спонтанно и вынужденно, вероятность вынужденных переходов также пропорциональна наличию фононов.

ЭКСИТОННОЕ ПОГЛОЩЕНИЕ

Представление об экситонах *) (т. е. о возбужденных, но не проводящих состояниях кристалла) было введено в 1931 г.

*) Название «экситон» происходит от английского слова *excitation* — возбуждение.

Я. И. Френкелем для того, чтобы объяснить тот факт, что поглощение света, связанное с переходами электронов, не всегда сопровождается появлением фотопроводимости.

Рассмотрим экситонное состояние на примере цепочки из N одинаковых атомов, один из которых возбужден. Так как все атомы одинаковы, то энергия этого состояния не зависит от того, какой из атомов возбужден, и поэтому такое состояние было бы N -кратно вырождено, если бы не существовало возмущения — взаимодействия атомов. В результате этого взаимодействия такое локализованное состояние становится неустойчивым и образуется квазинепрерывная зона экситонных состояний, каждое из которых описывается плоской модулированной волной, движущейся с определенной скоростью по кристаллу.

При этом существенно подчеркнуть, что движение экситона — это *движение возбужденного состояния*, а не перенос заряда, т. е. не движение электрона или дырки в отдельности.

Качественно механизм этого движения можно представить следующим образом: электрон, находящийся в возбужденном состоянии на i -м атоме, переходит в нормальное состояние; освободившаяся при этом энергия передается валентному электрону соседнего атома и т. д.

Каждому возбужденному состоянию атома соответствует зона экситонных состояний. Таким образом, в кристалле должна быть целая серия экситонных зон, часть из которых может перекрываться. Пределом этой серии является зона проводимости, которая образуется в результате коллективизации уровней, соответствующих ионизации атома.

Мотт предложил рассматривать экситон как совместное движение электрона и дырки, вращающихся синхронно по квантованной орбите вокруг общего центра тяжести и одновременно движущихся поступательно по кристаллу. При этом радиус орбиты определяется главным квантовым числом, т. е. номером экситонного состояния (или экситонной зоны), а скорость поступательного движения — положением уровня в зоне.

На первый взгляд кажется, что спектр поглощения, обусловленный возбуждением экситонных состояний, должен состоять из серии постепенно расширяющихся полос, ширина каждой из которых равна ширине соответствующей экситонной зоны. В действительности это неверно: дело в том, что, как мы уже неоднократно упоминали, импульс

фотона $h\nu/c$ очень мал. Поэтому под действием света могут образовываться только неподвижные экситоны, соответствующие нижним состояниям каждой зоны. Поэтому спектр экситонного поглощения состоит из серии дискретных линий, расширяющихся с ростом температуры за счет тепловых колебаний атомов.

9. 2. ФОТОПРОВОДИМОСТЬ СПЕКТРАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Спектральное распределение fotocувствительности зависит от того, что является источником фотоэлектронов — примесные центры или валентная зона (рис. 9.2).

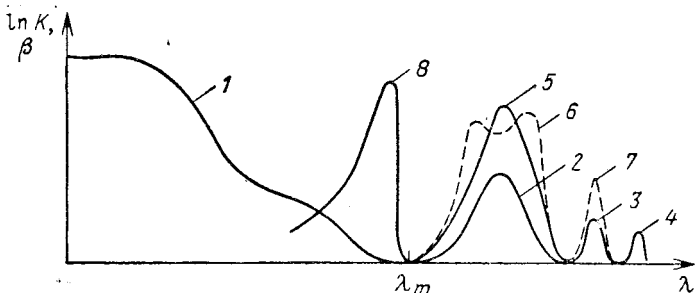


Рис. 9.2. Схема спектрального распределения поглощения и fotocувствительности:

1 — полоса собственного поглощения; 2 — примесное поглощение, соответствующее ионизации примесных атомов; 3 и 4 — кривые, соответствующие возбуждению примесных атомов; 5, 6 и 7 — кривые спектрального распределения fotocувствительности при примесном фотоэффекте; 8 — спектральное распределение fotocувствительности при собственном фотоэффекте.

В случае примесного фотоэффекта спектральное распределение fotocувствительности (кривая 5) обычно совпадает с широкой полосой поглощения, соответствующей переходу с примесных уровней в зону проводимости (кривая 2). В некоторых случаях, впрочем, в центре полосы поглощения наблюдается провал fotocувствительности (кривая 6), который объясняется тем, что в этой области при большой концентрации примесей поглощение очень велико и весь свет поглощается в поверхностном слое полупроводника. При этом скорость рекомбинации носителей оказывается очень велика (за счет поверхностной рекомбинации и большой концентрации фотоносителей) и время жизни носителей очень мало. Кроме того, эффективная

подвижность носителей в поверхностном слое значительно ниже, так как движению носителей препятствуют многочисленные дефекты приповерхностного слоя.

При достаточно высоких температурах может быть достаточно эффективной и двухступенчатая ионизация — электрон переводится светом в возбужденное состояние на примесном центре, а затем тепловым движением перебрасывается в зону проводимости. При этом появляется дополнительная область fotocувствительности (кривая 7) в полосе поглощения (кривая 3), соответствующей возбуждению примесных центров. По этой же причине при достаточно высоких температурах может наблюдаться фотоэффект в области экситонного поглощения.

При «собственной» фотопроводимости красная граница фотоэффекта λ_m (кривая 8) совпадает с красной границей собственного поглощения (кривая 1). Однако по мере увеличения коэффициента поглощения fotocувствительность сначала достигает максимума, а затем более или менее быстро спадает. Этот факт объясняется причинами, о которых мы упоминали выше: малыми временами жизни и эффективной подвижностью в поверхностном слое.

СРАВНЕНИЕ ОПТИЧЕСКОЙ И ТЕРМИЧЕСКОЙ ЭНЕРГИИ АКТИВАЦИИ

Примесная фотопроводимость. Мы проиллюстрируем этот вопрос на примере двухатомной молекулы, а затем

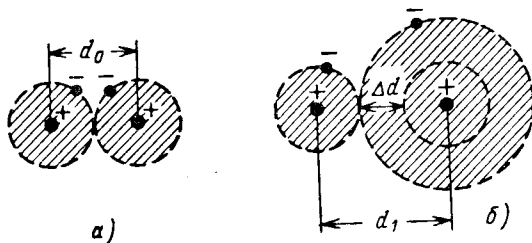


Рис. 9.3. Схема двухатомной молекулы в нормальном (а) и возбужденном (б) состояниях.

перенесем полученные выводы на кристалл. На рис. 9.3, а представлена двухатомная молекула в нормальном состоянии и на рис. 9.3, б — в возбужденном; в последнем случае в одном из атомов, образующих молекулу, валентный элект-

трон находится на возбужденной орбите и при этом радиус этого атома соответственно возрастает. На рис. 9.4 представлена зависимость энергии молекулы от расстояния между центрами атомов для нормального (кривая *a*) и возбужденного (кривая *b*) состояний. Как видно из рисунка, во втором случае равновесное расстояние между атомами больше на величину $d_1 - d_0$ и минимальная энергия молекулы

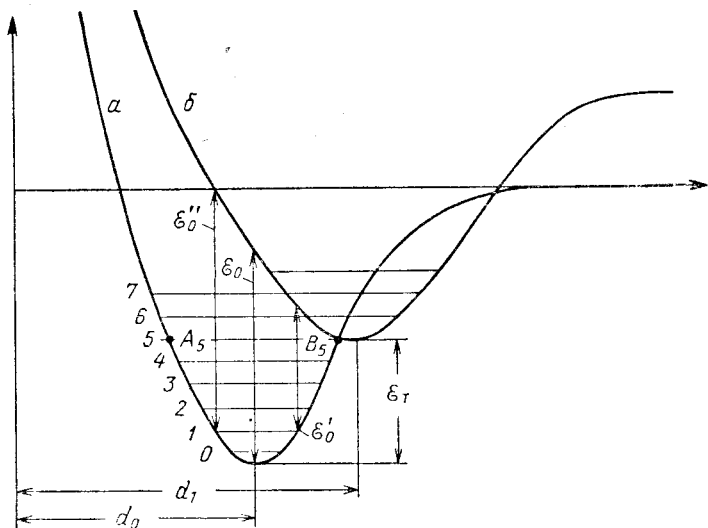


Рис. 9.4. Энергетический спектр двухатомной молекулы в нормальном (*a*) и возбужденном (*b*) состояниях.

больше на величину ϵ_T , необходимую для перевода электрона в возбужденное состояние. Эта величина и будет соответствовать энергии термического возбуждения атома. При этом сам механизм возбуждения можно представить следующим образом.

На рис. 9.4 горизонтальными тонкими линиями представлены колебательные уровни молекулы в нормальном и возбужденном состояниях. Если энергия теплового движения будет достаточно велика, то молекула может оказаться в состоянии 5, соответствующем колебаниям между точками A_5 и B_5 . При этом в тот момент, когда молекула находится в состоянии, соответствующем точке B_5 , электрон первого атома может без затраты энергии перейти в возбуж-

денное состояние, и, таким образом, энергия тепловых колебаний перейдет в энергию термического возбуждения электрона

Однако переход, аналогичный представленному выше, не может произойти под действием света. Мы уже упоминали, что такие переходы происходят почти «мгновенно», т. е. за время, в течение которого атомы практически не успевают сместиться, следовательно, такие переходы должны изображаться на рис. 9.4 вертикальными линиями. Таким образом, энергия оптической активации при абсолютном нуле будет соответствовать отрезку \mathcal{E}_0 на рис. 9.4. При температуре, отличной от абсолютного нуля, молекулы будут совершать тепловые колебания, например на уровне 1, и поэтому энергия оптической активации будет колебаться в пределах от \mathcal{E}'_0 до \mathcal{E}''_0 ; этим объясняется температурное расширение линий поглощения. Из представленной картины видно, что оптическая энергия активации отличается от термической.

Представленную выше картину можно с небольшими изменениями перенести на примесный атом в кристалле. Основное отличие в данном случае заключается в том, что у примесного атома даже только ближайших соседей будет не один, а много — число их будет определяться координацией решетки. Однако несомненно то, что и в этом случае равновесное расстояние до ближайших соседей будет также определяться радиусом (а следовательно, состоянием) примесного атома, и в возбужденном состоянии оно будет больше, чем в нормальном. Поэтому если мы под r будем понимать среднее расстояние до ближайших соседей (так называемый «координационный радиус»), то проведенный выше качественный анализ можно перенести на примесный атом в кристалле.

Собственная фотопроводимость. Эффекты, описанные выше, могут явиться причиной для различия оптической и термической энергии активации и при собственной фотопроводимости.

Как мы уже много раз говорили, электронные состояния атомов в известной мере определяются периодическим потенциалом решетки; справедливо и обратное утверждение: геометрическая и кристаллохимическая структура решетки определяются электронными волновыми функциями. Поэтому, если электрон переходит под действием света из валентной зоны в свободную, то атомы должны

в соответствии с этим как-то перестроиться, что будет сопровождаться возвратом энергии.

Однако последний процесс требует времени. Поэтому на переход электрона будет затрачена световая энергия, соответствующая неизменной конфигурации атомов, а затем избыток ее «возвратится» в виде тепловых колебаний атомов. В случае термической ионизации эти процессы могут происходить одновременно или в обратном порядке, т. е. тепловое движение сначала создает благоприятную конфигурацию атомов, а затем совершается электронный переход. Таким образом, и в случае фотовыравнивания электрона из основной зоны величина оптической энергии активации может быть больше термической. Это различие наиболее ярко выражено в ионных кристаллах, где изменение электронных состояний вызывает значительные смещения ионов (ионную поляризацию).

Вторая причина различия оптической и термической энергий активации может лежать в зонной структуре.

Если экстремумы зон не совпадают, то для прямых переходов оптическая энергия активации будет больше термической. Для непрямых переходов эти энергии могут быть приблизительно одинаковы (см. рис. 9.1, б).

Наконец, третья причина различия $\Delta\mathcal{E}_T$ и $\Delta\mathcal{E}_0$ может заключаться в наличии вырождения; этот случай проиллюстрирован на рис. 9.1, в и г.

РЕКОМБИНАЦИЯ ФОТОНОСИТЕЛЕЙ

Мы уже упоминали, что историю каждого фотоносителя можно разбить на четыре основные периода:

1) рождение (генерацию)— этот процесс мы рассмотрели выше;

2) время, в течение которого фотоноситель в результате столкновений теряет избыточную энергию и приходит в тепловое равновесие с решеткой; на этом вопросе мы также останавливались и установили, что это время превышает примерно в 10^2 — 10^3 раз время «упругой» релаксации, но обычно на много порядков меньше, чем полное время пребывания в зоне;

3) самый длительный период жизни фотоносителя, в течение которого он, смешавшись с темновыми носителями и ничем не отличаясь от них (кроме своего происхождения), участвует во всех явлениях переноса наравне

с ними. Этот период может прерываться более или менее длинными периодами прилипания, что будет уменьшать эффективную подвижность и коэффициент диффузии.

Следует, однако, подчеркнуть, что в течение всего этого периода общая концентрация носителей превышает равновесную и рекомбинация преобладает над тепловой генерацией. Поэтому с того момента, как световая (или созданная другим путем) генерация прекратится, концентрация носителей начинает спадать по тому или иному закону;

4) последний и едва ли не самый главный момент в жизни фотоносителя — его «смерть», т. е. рекомбинация. На этом вопросе мы останавливались уже в гл. 1, однако ввиду его важности вернемся к нему еще раз.

Рекомбинация ограничивает и тем самым определяет время жизни фотоносителей τ . Поэтому скорость рекомбинации определяет стационарную концентрацию носителей. Число создаваемых светом в 1 сек фотоносителей (g) пропорционально интенсивности света (J). Если число генерируемых в единицу времени носителей g , а число рекомбинирующих $\Delta n/\tau$, где Δn — общее число фотоносителей, накопившихся к данному моменту времени, то условие стационарности будет иметь вид

$$\frac{d\Delta n}{dt} = g - \frac{\Delta n}{\tau} = 0, \quad (9.35)$$

откуда

$$\Delta n_{ст} = g\tau. \quad (9.36)$$

В ряде случаев время жизни фотоносителей обратно пропорционально их концентрации $\tau^{-1} = \gamma\Delta n$, при этом условие стационарности будет иметь вид

$$\frac{d\Delta n}{dt} = g - \gamma\Delta n^2 = 0, \quad (9.37)$$

откуда $\Delta n_{ст} = \sqrt{\frac{g}{\gamma}}$, т. е. в данном случае стационарная фотопроводимость пропорциональна корню из интенсивности света. В первом случае (9.35) рекомбинация называется линейной, во втором (9.37) — квадратичной.

Законы нарастания и спадания фотопроводимости также определяются характером рекомбинации. В случае линейной рекомбинации закон нарастания определяется дифференциальным уравнением

$$\frac{d\Delta n}{dt} = g - \frac{\Delta n}{\tau}, \quad (9.38)$$

откуда

$$\Delta n = g\tau (1 - e^{-\frac{t}{\tau}}); \quad (9.39)$$

после прекращения генерации

$$\frac{d\Delta n}{dt} = -\frac{\Delta n}{\tau} \quad (9.40)$$

и

$$\Delta n = g\tau e^{-\frac{t}{\tau}}.$$

В случае квадратичной рекомбинации

$$\frac{d\Delta n}{dt} = g - \gamma\Delta n^2 \quad (9.41)$$

и

$$n = \sqrt{\frac{g}{\gamma}} \operatorname{th} t \sqrt{g\gamma}, \quad (9.42)$$

и после выключения света

$$\frac{d\Delta n}{dt} = -\gamma\Delta n^2, \quad (9.43)$$

$$n = \sqrt{\frac{g'}{\gamma}} \frac{1}{t \sqrt{g'\gamma} + 1}.$$

Как видно из выражений (9.42) и (9.43), в случае квадратичной рекомбинации концентрация фотоносителей нарастает и спадает по разным законам и ее нельзя характеризовать постоянным временем релаксации. Как видно из (9.42) и (9.43), спад происходит значительно медленнее, чем нарастание*). Однако мы и в этом случае можем ввести мгновенное время жизни, определив его как величину, обратную средней вероятности рекомбинации $\gamma\Delta n$ для отдельного электрона:

$$\tau = \frac{1}{\gamma\Delta n}, \quad (9.44)$$

где γ может зависеть от времени, концентрации фотоносителей и т. д. Это выражение сохраняет силу для любого закона рекомбинации.

*) Из сказанного выше также следует, что характером рекомбинации определяется инерционность фотоэффекта, так, в частности, квадратичный закон неблагоприятен для точного воспроизведения характера спада коротких фотосигналов.

В действительности линейный и квадратичный законы являются лишь сравнительно редкими частными случаями. Как уже упоминалось в гл. I, вероятность рекомбинации и время жизни носителей зависят от мгновенной заселенности всех примесных уровней и поэтому с течением времени должны меняться по сложному закону. В заключение рассмотрим работу Шокли и Рида, заложившую фундамент статистики рекомбинации через примесные центры.

СТАТИСТИКА РЕКОМБИНАЦИИ ШОКЛИ — РИДА*)

В модели, рассмотренной авторами, в запрещенной зоне имеется один тип уровней захвата, которые могут захватить электрон из валентной зоны и затем эмиттировать его в зону проводимости или наоборот (см. рис. 9.5, а).

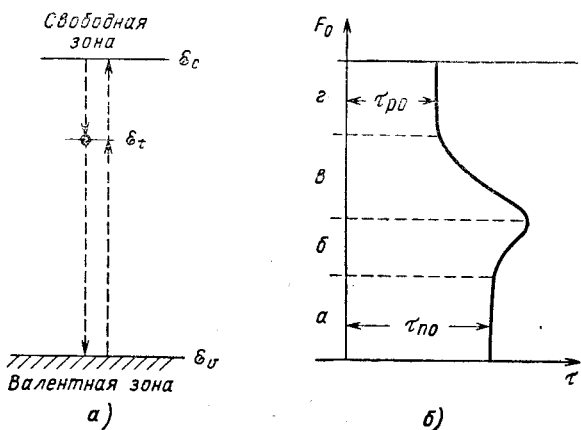


Рис. 9.5. Схема рекомбинации по Шокли — Риду (а) и зависимость времени жизни от положения уровня Ферми (б).

При этом предполагается, что у захваченного электрона имеется только одно нормальное состояние (или время пребывания в возбужденных состояниях пренебрежимо мало). Если число таких центров захвата (ловушек) в единице объема N_t , то скорость захвата электронов, т. е. число актов захвата в 1 сек, будет иметь

*) Работу Шокли и Рида мы излагаем здесь в сокращенном и упрощенном виде; для того чтобы читателю было легче дополнить свои сведения чтением оригинальной статьи, здесь сохранены обозначения авторов [28].

$$f_{pt} N_t c_n(\mathcal{E}) f(\mathcal{E}) N(\mathcal{E}) d\mathcal{E}, \quad (9.45)$$

где f_{pt} — доля свободных ловушек (занятых дырками);
 $c_n(\mathcal{E})$ — вероятность захвата электрона с энергией \mathcal{E} такой ловушкой;

$f(\mathcal{E})$ — функция распределения Ферми;

$N(\mathcal{E})$ — плотность электронных состояний в зоне проводимости.

Вероятность обратного процесса будет

$$f_t N_t e_n f_p(\mathcal{E}) N(\mathcal{E}) d\mathcal{E}, \quad (9.46)$$

где f_t — доля ловушек, занятых электроном;

e_n — вероятность обратного перехода, соответствующая c_n *;

$f_p(\mathcal{E})$ — функция распределения дырок

$$f_p(\mathcal{E}) = 1 - f(\mathcal{E}). \quad (9.47)$$

Результирующая скорость захвата электронов в зоне проводимости может быть найдена путем вычитания (9.46) из (9.45):

$$dU_{cn} = \left[f_{pt} f(\mathcal{E}) - \frac{e_n}{c_n} f_t f_p(\mathcal{E}) \right] N_t c_n(\mathcal{E}) N(\mathcal{E}) d\mathcal{E}, \quad (9.48)$$

и интегрирования полученного выражения по всем энергиям в зоне проводимости.

В состоянии теплового равновесия, исходя из принципа детального равновесия, мы можем сделать вывод, что выражение, стоящее в (9.48) в скобках, должно быть равно нулю, откуда

$$e_n = c_n e^{-\frac{\mathcal{E}_t - \mathcal{E}}{kT}}. \quad (9.49)$$

В случае нарушения теплового равновесия функции распределения электронов, дырок и степень заполнения ловушек определяются квазиуровнями Ферми F_n , F_p и F_t :

$$f_n = \frac{1}{e^{\frac{\mathcal{E} - F_n}{kT}} + 1}, \quad f_p = 1 - f_n, \quad (9.50)$$

$$f_t = \frac{1}{e^{\frac{\mathcal{E} - F_t}{kT}} + 1} \quad \text{и} \quad f_{pt} = 1 - f_t. \quad (9.51)$$

Подставляя (9.50) и (9.51) в (9.48) и проведя интегрирование, получаем окончательно

$$u_{cn} = \left[1 - e^{\frac{F_t - F_n}{kT}} \right] f_{pt} N_t \int_{\mathcal{E}}^{\infty} f(\mathcal{E}) N(\mathcal{E}) c_n(\mathcal{E}) d\mathcal{E}, \quad (9.52)$$

*) Соотношения c_n и e_n будут ниже найдены из принципа детального равновесия.

где интегрирование проводится по всем уровням зоны проводимости. Полученное выражение весьма наглядно. Оно показывает, что скорость захвата пропорциональна: 1) числу свободных ловушек $f_{pl} N_t$, 2) числу электронов в зоне проводимости (подынтегральное выражение), умноженному на $c_n(\mathcal{E})$, их вероятности захвата и 3) степени отступления от равновесного состояния — выражение в скобках.

Если система находится в состоянии равновесия, квазиуровни Ферми для зоны и для ловушек совпадают и превращаются в общий уровень Ферми

$$F_t = F_n = F. \quad (9.53)$$

Выражение в скобках уравнения (9.52) обращается в нуль. Если $F_n > F_t$, т. е. концентрация носителей в зоне превышает необходимую для поддержания равновесия с ловушками и рекомбинация превалирует над генерацией, U_{cn} положительно. В случае $F_t > F_n$ число забросов в зону превышает число захватов — концентрация свободных носителей растет, U_{cn} отрицательно.

Выражение, аналогичное U_{cn} , сможет быть выведено для U_{cp} — интегральной скорости захвата дырок.

Выражение (9.52) сильно упрощается в случае невырожденного полупроводника; как нетрудно убедиться, оно при этом приобретает вид

$$u_{cn} = \left[1 - e^{-\frac{F_t - F_n}{kT}} \right] f_{pl} n C_n, \quad (9.54)$$

где n — полная концентрация электронов в зоне проводимости, равная

$$n = N_c e^{\frac{F_n - \mathcal{E}_c}{kT}}; \quad (9.55)$$

\mathcal{E}_c — уровень, соответствующий дну зоны; N_c — эффективная плотность состояний в зоне в интервале kT , равная

$$N_c = \int_{\mathcal{E}_c}^{\infty} e^{-\frac{\mathcal{E}_c - \mathcal{E}}{kT}} N(\mathcal{E}) d\mathcal{E} = \frac{2(2\pi mkT)^{\frac{3}{2}}}{h^3}, \quad (9.56)$$

и C_n — средняя интегральная вероятность захвата для одного электрона всеми ловушками (если бы все они были свободны),

$$C_n = N_t \langle c_n \rangle,$$

где

$$\langle c_n \rangle = \int_{\mathcal{E}_c}^{\infty} e^{-\frac{\mathcal{E}_c - \mathcal{E}}{kT}} c_n(\mathcal{E}) \frac{N(\mathcal{E}) d\mathcal{E}}{N_c}. \quad (9.57)$$

Как мы уже упоминали, первый член в квадратных скобках (9.52) (т. е. единица) соответствует захвату носителей, а второй $e^{(F_n - F_t)/kT}$ — их тепловому забросу.

В случае отсутствия вырождения вероятность заброса не должна зависеть от F_n , так как практически все уровни в зоне свободны. Раскрывая скобки в (9.54), убедимся, что это действительно так ($-F_n$ в экспоненциальном члене и n входят с обратными знаками и взаимно уничтожаются). Введя обозначение

$$f_{pt} n e^{\frac{F_t - F_n}{kT}} = f_t N_c \exp\left(\frac{\mathcal{E}_t - \mathcal{E}_c}{kT}\right) = f_t n_1, \quad (9.58)$$

где

$$n_1 = N_c \exp\left(\frac{\mathcal{E}_t - \mathcal{E}_c}{kT}\right), \quad (9.59)$$

получим окончательно

$$U_{cn} = C_n f_{pt} n - C_n f_t n_1. \quad (9.60)$$

Аналогичные выражения для скорости рекомбинации дырок приведем без вывода:

$$U_{cp} = C_p f_{pt} p - C_p f_t p_1, \quad (9.61)$$

где

$$p_1 = N_p \exp\left(\frac{\mathcal{E}_p - \mathcal{E}_t}{kT}\right). \quad (9.62)$$

Рассмотрим рекомбинацию в стационарных условиях, когда пары электрон — дырка генерируются светом или каким-либо другим ионизирующим агентом с постоянной скоростью u и когда прошло достаточно времени для установления стационарного состояния; при этом скорости рекомбинации электронов и дырок должны быть одинаковы и равны U :

$$U_{cn} = U_{cp} = U. \quad (9.63)$$

Приравняв (9.61) и (9.60) и заменив f_{pt} через f_t согласно

$$f_{pt} = 1 - f_t, \quad (9.64)$$

получим

$$C_n (1 - f_t) n - C_n f_t n_1 = C_p f_t p - C_p (1 - f_t) p_1. \quad (9.65)$$

Решив это уравнение относительно f_t , получим

$$f_t = \frac{C_n n + C_p p_1}{C_n (n + n_1) + C_p (p + p_1)} \quad (9.66)$$

и

$$f_{pt} = \frac{C_n n_1 + C_p p}{C_n (n + n_1) + C_p (p + p_1)}. \quad (9.67)$$

Подставив полученные выражения в (9.60) или (9.61), получим

$$U = U_{cn} = U_{cp} = \frac{C_n C_p (pn - p_1 n_1)}{C_n (n + n_1) + C_p (p + p_1)}. \quad (9.68)$$

В выражении (9.68) произведение $p_1 n_1$, как нетрудно убедиться, не зависит от уровня ловушек и имеет вид

$$p_1 n_1 = N_c N_p e^{\frac{\mathcal{E}_v - \mathcal{E}_c}{kT}} = N_c N_p e^{-\frac{\mathcal{E}_g}{kT}} = n_i^2. \quad (9.69)$$

Обозначим \mathcal{E}_i — положение уровня химического потенциала в собственном полупроводнике:

$$\mathcal{E}_i = \frac{1}{2} (\mathcal{E}_c - \mathcal{E}_v) + \frac{1}{2} kT \ln \frac{N_D}{N_C}. \quad (9.70)$$

В дальнейшем будем предполагать, что $\mathcal{E}_t > \mathcal{E}_i$ и, следовательно,

$$n_1 > n_i > p_1. \quad (9.71)$$

Рассмотрим на основе полученного выражения случай малых отступлений от состояния равновесия, полагая при этом, что концентрация основных носителей велика по сравнению с концентрацией ловушек.

Таким образом, мы предположим, что

$$n = n_0 + \delta n \text{ и } p = p_0 + \delta p, \quad (9.72)$$

где n_0 и p_0 — равновесные концентрации носителей;
 δp и δn — малые величины.

Согласно (9.35) скорость рекомбинации связана с временем жизни соотношением

$$U = \frac{dn}{dt} = \frac{\delta n}{\tau}; \quad (9.73)$$

воспользовавшись (9.73) и (9.72), мы можем получить выражение для времени жизни

$$\begin{aligned} \tau &= \frac{n_0 + n_1 + \delta n}{(n_0 + p_0 + \delta n)C_p} + \frac{p_0 + p_1 + \delta p}{(n_0 + p_0 + \delta n)C_n} = \\ &= \tau_{p0} \frac{n_0 + n_1 + \delta n}{(n_0 + p_0 + \delta n)} + \tau_{n0} \frac{p_0 + p_1 + \delta n}{(n_0 + p_0 + \delta n)}, \end{aligned} \quad (9.74)$$

где

$$\tau_{p0} = \frac{1}{C_p} \text{ и } \tau_{n0} = \frac{1}{C_n}. \quad (9.75)$$

Величина τ_{p0} — время жизни дырок в образце ярко выраженного n -типа. В этом случае ловушки практически все заполнены электронами и скорость рекомбинации дырок

$$U_{p0} = C_{p0} \delta p = \frac{\delta p}{\tau_{p0}}. \quad (9.76)$$

Точно так же τ_{n0} — время жизни электронов в образцах ярко выраженного p -типа. В рассматриваемом нами случае малых отступлений от состояния равновесия выражение (9.74) принимает вид

$$\tau = \tau_{p0} \frac{n_0 + n_1}{n_0 + p_0} + \tau_{n0} \frac{p_0 + p_1}{n_0 + p_0}. \quad (9.77)$$

На рис. 9.5, б представлена зависимость времени жизни τ от положения уровня Ферми, построенная согласно (9.77). Как видно из рисунка, здесь следует рассматривать четыре основные области:

1) область *a*, где $\tau = \tau_{n0}$. Когда уровень Ферми проходит в этой области, все примесные уровни свободны и число дырок в валентной зоне велико. Любой электрон, появившийся в зоне проводимости, тотчас же сваливается в ловушку и вслед за этим на пустое место в валентной зоне. В этом случае весь процесс ограничивается скоростью захвата электронов, поэтому время жизни пары

$$\tau = \frac{1}{C_n} = \tau_{n0};$$

2) область *b*; здесь ловушки еще почти все пусты, но число пустых мест в валентной зоне (т. е. число дырок в валентной зоне) начинает ограничивать рекомбинацию. Электрон может, не дождаввшись прихода дырки, вновь уйти в зону проводимости, поэтому время жизни растет по мере подъема уровня Ферми;

3) на участке *в* число электронов на примесных уровнях по-прежнему достаточно, но число свободных электронов в зоне проводимости уже мало и начинает ограничивать вероятность рекомбинации;

4) аналогичные рассуждения можно провести для верхней половины зоны; здесь мы уже имеем дело с ярко выраженными образцами *n*-типа. Все ловушки забиты, и весь процесс лимитируется скоростью захвата дырок.

9.3. ФОТОВОЛЬТАИЧЕСКИЕ ЭФФЕКТЫ

До сих пор мы рассматривали только изменение проводимости под действием света. Для того чтобы освещение полупроводника вызвало появление в цепи электродвижущей силы (фото-э. д. с.), должно выполняться одно из следующих двух условий:

а) либо освещение должно быть неоднородным, т. е. в различных элементах объема полупроводника должно поглощаться различное количество фотонов и соответственно появляться различное количество фотоносителей. Наличие градиента концентрации фотоносителей будет вызывать их диффузию и за счет этого возникнет фото-э. д. с. (если диффузия дырок и электронов не будет абсолютно одинакова);

б) либо освещаемый полупроводник должен быть неоднородным, в этом случае распределение концентрации фотоносителей также может быть непостоянным и вследствие этого может возникнуть фото-э. д. с.

Рассмотрим коротко три фотовольтаических эффекта: эффект Дембера, фотомагнитный эффект Кикоина — Носкова и вентильный фотоэффект.

Эффект Дембера. Если на поверхность полупроводника падает свет, энергия квантов которого достаточна для генерации фотоносителей, и коэффициент поглощения света при этом достаточно велик, то свет будет поглощаться в основном в приповерхностном слое полупроводника и в этом слое будут создаваться фотоносители. Это как раз и есть случай неоднородной освещенности полупроводника, упомянутый выше. Электроны или дырки (или и те и другие) будут при этом диффундировать в область с меньшей освещенностью и за счет этого возникнет э. д. с. пропорциональная разности коэффициентов диффузии (или, что то же самое, подвижностей) носителей противоположного знака. Эта разность потенциалов скомпенсирует разность коэффициентов диффузии, и в стационарном состоянии электроны и дырки будут уже в равном количестве диффундировать в затененную область и там рекомбинировать.

Приведем выражение для э. д. с. Дембера без вывода:

$$V_D = \frac{kT}{e} \frac{u_p - u_n}{u_p + u_n} \ln \frac{\sigma_{\Pi}}{\sigma_T}, \quad (9.78)$$

здесь σ_{Π} — удельная электропроводность полупроводника в непосредственной близости от поверхности, т. е. сумма темновой и фотопроводимости:

$$\sigma_{\Pi} = \sigma_{cp} + \sigma_T, \quad (9.79)$$

и σ_T — проводимость в глубине полупроводника в той области, куда уже не проникает свет, т. е. темновая проводимость.

Из (9.78) видно, что фото-э. д. с. Дембера тем больше, чем больше разность подвижностей. Формула (9.78) может быть применена также к примесной фотопроводимости (т. е. к тому случаю, когда электроны вырываются с доноров или забрасываются на акцепторы). При этом либо u_n , либо u_p должна быть приравнена к нулю. Эффект Дембера в известной мере аналогичен явлению термо-э. д. с.

Фотомагнитный эффект Кикоина — Носкова [34]. Как мы уже упоминали, э. д. с. Дембера возникает лишь за счет разности подвижностей дырок и электронов, при $u_p = u_n$ согласно (9.78) $V_D = 0$. Но если теперь мы к этому же образцу приложим магнитное поле, направленное перпендикулярно к потоку лучистой энергии и к одной из граней, то и на электроны и на дырки, диффундирующие от освещенной поверхности, будет действовать магнитная состав-

ляющая силы Лоренца

$$F_{\text{л}} = \frac{e [vH]}{c} \quad (9.80)$$

и отклонять их (электроны и дырки) в противоположные стороны (так как направление скорости у них в данном случае одинаково и заряды противоположны).

Таким образом, возникает поперечная э. д. с., перпендикулярная направлению магнитного поля и потока лучистой энергии. Величина э. д. с. Кикоина — Носкова в стационарном состоянии определяется тем условием, что ток, созданный этой э. д. с., должен скомпенсировать ток, создаваемый магнитным полем.

Если эффект Дембера аналогичен явлению термо-э. д. с., то эффект Кикоина — Носкова аналогичен (разумеется, так же, как и в первом случае, с целым рядом оговорок) поперечному эффекту Нернста — Эттингсгаузена.

Если соединить накоротко грани образца, на которых скапливаются носители противоположного знака, то фотомагнитная э. д. с. создает ток

$$I_{\text{фм}} = \frac{\eta NeH}{l} [u_n L_n + u_p L_p], \quad (9.81)$$

где N — число квантов света, падающих на 1 см^2 полупроводника в 1 сек ; η — квантовый выход фотоэффекта:

$$\eta = \frac{n_{\text{ф}}}{N}; \quad (9.82)$$

L_p и L_n — диффузионная длина дырок и электронов; l — длина образца в направлении тока.

Фотомагнитная э. д. с. (холостого хода) может быть найдена путем умножения тока короткого замыкания $I_{\text{фм}}$ на сопротивление образца с учетом его уменьшения за счет появления фотоносителей.

Вентильный фотоэффект. Вентильным фотоэффектом называют фото-э. д. с., возникающую при освещении вентильного (т. е. выпрямляющего) контакта. Выпрямляющими свойствами обладают запиорный слой на границе полупроводника и металла и p - n переход. Полупроводниковые приборы, основанные на вентильном фотоэффекте и предназначенные для превращения световой энергии в электрическую или световых сигналов в электрические, называют фотоэлементами. За первым видом фотоэлементов, т. е.

основанных на выпрямляющих свойствах контакта полупроводника и металла, укрепилось название вентильных, которые обычно называют фотоэлементами с $p-n$ переходом. Мы здесь рассмотрим коротко лишь фотоэлемент с $p-n$ переходом.

Напряжение холостого хода. Представим себе $n-p$ переход, освещаемый светом, энергия квантов которого достаточна для создания пар электрон — дырка. Предположим, что расстояние перехода d от поверхности фотоэлемента много меньше диффузионной длины электронов и что весь свет поглощается активно в p -области. Поверхностной рекомбинацией мы вначале пренебрежем *).

Тогда все электроны, генерируемые светом, будут достигать $n-p$ перехода и под действием контактного поля переходить в n -область, дырки же, напротив, будут задерживаться контактным полем и оставаться в p -области.

Таким образом, под действием света через $p-n$ переход пойдет фототок

$$I_{\Phi} = eg, \quad (9.83)$$

где g — число электронов (или, что то же самое, пар), создаваемых светом в 1 сек.

Этот ток, идущий в запертом направлении, будет создавать на $n-p$ переходе разность потенциалов V в пропускном направлении (точнее, уменьшать контактную разность потенциалов) и вследствие этого через переход потечет ток I_y , называемый током утечки, в обратном направлении:

$$I_y = I_s \left(e^{\frac{eV}{kT}} - 1 \right). \quad (9.84)$$

Стационарное состояние установится при таком V , при котором ток утечки будет равен фототоку.

Таким образом, условие стационарности будет иметь вид

$$I_{\Phi} - I_s \left(e^{\frac{eV_{xx}}{kT}} - 1 \right) = 0, \quad (9.85)$$

откуда напряжение холостого хода

$$V_{xx} = \frac{kT}{e} \ln \left(\frac{I_{\Phi}}{I_s} + 1 \right). \quad (9.86)$$

*) При $d \ll L_n$ мы можем также пренебречь и объемной рекомбинацией.

Вольтамперная характеристика фотоэлемента. Если фотоэлемент замкнут на сопротивление r , то стационарное состояние установится при меньшем напряжении V ; при этом ток во внешней цепи будет равен разности фототока I_{Φ} и тока утечки I_y :

$$I = I_{\Phi} - I_s \left(e^{\frac{eV}{kT}} - 1 \right), \quad (9.87)$$

с другой стороны,

$$I = \frac{V}{r}. \quad (9.88)$$

Уравнения (9.87) и (9.88) совместно дают ток и напряжение фотоэлемента; из них может быть легко исключено одно из неизвестных (ток или напряжение), но в обоих случаях мы получаем трансцендентное уравнение, которое может быть решено только приближенно. Если исключить из (9.87) и (9.88) ток, то это уравнение принимает следующий вид:

$$\frac{V}{r} = I_{\Phi} - I_s \left(e^{\frac{eV}{kT}} - 1 \right). \quad (9.89)$$

Если выразить I_{Φ} согласно (9.85) через напряжение холостого хода и подставить в (9.87), то оно принимает более простой вид:

$$I = I_s \left(e^{\frac{eV_{xx}}{kT}} - e^{\frac{eV}{kT}} \right), \quad (9.90)$$

это, однако, не меняет существенно положения дел: система уравнений (9.90) и (9.88) по-прежнему остается трансцендентной.

Работа фотоэлемента с p-n переходом в фотодиодном режиме. Если фотоэлемент используется для регистрации световых сигналов, то в ряде случаев оказывается полезным включить в его цепь дополнительный источник э. д. с. U в заперном направлении. С этой э. д. с. уравнение (9.89) принимает вид

$$\frac{V-U}{r} = I_{\Phi} - I_s \left(e^{\frac{eV}{kT}} - 1 \right); \quad (9.91)$$

если напряжение U включено в заперном направлении, то, беря его достаточно большим, можно сделать темновой ток фотодиода равным току насыщения. На фототок величина U

почти не влияет и даже несколько его увеличивает, так как увеличивает тянущее поле и уменьшает таким образом потери на рекомбинацию. В этом состоит большое преимущество фотодиодов перед фотоспротивлением.

Второе преимущество фотодиода заключается в том, что он обладает очень малой инерционностью и поэтому удобен для регистрации коротких сигналов.

Учет рекомбинации. Если учесть и объемную и поверхностную рекомбинацию, то выражение для фототока принимает вид

$$I_{\Phi} = eg(1 - \beta), \quad (9.92)$$

где β — коэффициент, учитывающий оба вида рекомбинации и равный

$$\beta = \beta_{об} + \beta_s = \frac{\frac{d}{L_s + \frac{1}{2} \left(\frac{d}{L_n} \right)^2}}{1 + \frac{d}{L_s}}, \quad (9.93)$$

где L_n — диффузионная длина в объеме: $L_n = \sqrt{D_n \tau_n}$ и L_s — диффузионная длина вблизи поверхности: $L_s = D_n/s$ (где s — скорость поверхностной рекомбинации). Выражение для тока насыщения I_s при учете поверхностной рекомбинации приобретает также более сложный вид [31].

9.4. ЦИКЛОТРОННЫЙ РЕЗОНАНС

Явление циклотронного резонанса в настоящее время широко используется как один из наиболее тонких методов исследования зонной структуры полупроводников. Есть два подхода к изучению этого явления — классический и квантовомеханический.

Классический подход. Как было показано в гл. 7, свободный (или зонный) электрон в магнитном поле движется по винтовой линии. Если составляющая скорости электрона вдоль магнитного поля равна нулю, то эта винтовая линия вырождается в окружность с радиусом

$$r = \frac{m\omega v}{eH}, \quad (9.94)$$

При этом круговая частота обращения электрона по ней (так называемая циклотронная частота) не зависит от его скорости:

$$\omega_0 = \frac{eH}{mc} . \quad (9.95)$$

Представим теперь, что одновременно с постоянным магнитным полем на электрон действует слабое электромагнитное излучение, поляризованное по кругу, направление распространения которого совпадает с направлением постоянного магнитного поля и частота которого $\omega = \omega_0$, причем направление вращения плоскости поляризации совпадает с направлением движения электрона по окружности. Тогда электрическая составляющая излучения будет ускорять электрон, скорость его при этом будет увеличиваться, но так как согласно (9.95) частота обращения не зависит от скорости *), то этот процесс будет продолжаться до тех пор, пока электрон в результате какого-либо столкновения не отдаст накопленную энергию, и затем все начнется сначала.

Если частота ω заметно отличается от ω_0 , то действие излучения на электрон будет ничтожным, так как в среднем за большой промежуток времени оно будет столько же времени ускорять электрон, сколько и замедлять (так как разность фаз все время будет меняться).

Таким образом, соблюдение условия $\omega = \omega_0$ выразится в резонансном поглощении электромагнитной энергии; на этом и основано явление циклотронного резонанса, позволяющее согласно формуле (9.94) определить эффективную массу электрона. Если в исследуемом материале масса электрона анизотропна, то, изменяя направление магнитного поля и одновременно с ним направление электромагнитного излучения (или просто вращая кристалл), можно изучить форму эллипсоида эффективных масс.

Квантовомеханический подход. Как мы уже упоминали, любое движение электрона в ограниченном объеме квантуется; это, следовательно, должно относиться и к движению электрона по окружности в магнитном поле. Условия квантования в данном случае можно получить несколькими способами:

*) До тех пор пока скорость не достигнет релятивистских значений и масса электрона не начнет зависеть от скорости.

1. Согласно гипотезе Бора — де Бройля на орбите должно укладываться целое число волн:

$$2\pi r = n\lambda = \frac{n\hbar}{mv} \quad (9.96) *$$

или

$$mv = \frac{n\hbar}{r}, \quad (9.97)$$

с другой стороны,

$$\frac{v}{r} = \omega. \quad (9.98)$$

Умножая (9.97) на (9.98), найдем, что кинетическая энергия электрона

$$T = \frac{1}{2} mv^2 = \frac{1}{2} n\hbar\omega. \quad (9.99)$$

Формула (9.84) дает точное значение энергии для первого уровня, соответствующего при $n = 1$

$$\varepsilon_1 = \frac{1}{2} \hbar\omega. \quad (9.100)$$

Для всех остальных уровней строгая теория дает выражение

$$\varepsilon = \left(n + \frac{1}{2} \right) \hbar\omega_0. \quad (9.101)$$

2. Соотношение (9.99) можно также получить, если учесть, что движение по кругу с частотой ω_0 эквивалентно сложению двух колебаний с той же частотой, но, как мы видели выше, энергия колебаний может принимать значения, удовлетворяющие формуле (9.101).

Так или иначе энергия электрона в плоскости, перпендикулярной магнитному полю, может принимать лишь квантованные значения, удовлетворяющие соотношению (9.99).

Следовательно, для того чтобы падающее электромагнитное излучение резонансно поглощалось, его кванты должны удовлетворять условию

$$\hbar\omega = \hbar\omega_0 \quad (9.102)$$

или

$$\omega = \omega_0. \quad (9.103)$$

*) Соотношение (9.96) совершенно эквивалентно условию квантования момента количества движения.

К сказанному выше следует сделать ряд дополнений. Падающее электромагнитное излучение не должно быть обязательно поляризовано по кругу — вместо вращающегося поля можно применять и линейно поляризованное; при правильном соотношении фаз электрон все время будет ускоряться.

Условие правильного соотношения фаз не является обязательным: если оно не соблюдается, электрон будет тормозиться, потеряет свою скорость и начнет снова ускоряться уже в правильной фазе. Таким образом, здесь так же, как в циклотроне, имеет место самосинхронизация.

Значительно более жесткие ограничения на условия опыта накладывает ширина резонансных линий. Оценка резонансной частоты ω_0 по формуле (9.95) показывает, что при эффективной массе, равной массе свободного электрона, и магнитном поле $H \approx 10^3 \div 10^4$ эрст резонансная частота ω_0 лежит в области сантиметрового диапазона, $\omega_0 \approx 10^{10}$, т. е. период обращения электрона на орбите $T \approx 10^{-10}$. Для того чтобы резонанс был заметным, электрон должен совершить без столкновений несколько оборотов, иначе соотношение фаз при столкновениях будет нарушаться слишком часто.

Таким образом, время свободного пробега должно удовлетворять условию

$$\tau \geq T \approx 10^{-10}. \quad (9.103a)$$

На языке квантовой механики это означает, что ширина уровней

$$\Delta \mathcal{E} = \frac{\hbar}{\tau} \quad (9.104)$$

должна быть меньше расстояния между уровнями $\hbar\omega$:

$$\hbar\omega = \frac{\hbar}{T} > \frac{\hbar}{\tau} \quad (9.105)$$

или

$$\tau > T, \quad (9.106)$$

т. е. опять получим для времени свободного пробега то же условие (9.103a).

В настоящее время этому требованию удовлетворяет лишь весьма ограниченное число очень чистых полупроводников (притом при низких температурах), на которых до сих пор и проводились опыты по циклотронному резонансу.

Так как плавные изменения частоты излучения в сантиметровом диапазоне представляют весьма сложную задачу, то практически в условиях опыта меняется не ω , а ω_0 , т. е. плавно меняется постоянное магнитное поле при фиксированной частоте падающего электромагнитного излучения. При этом, когда магнитное поле достигает величины, соответствующей условию $\omega = \omega_0$, наблюдается пик поглощения. Если в материале имеется несколько сортов носителей, то наблюдается несколько пиков.

9.5. СТИМУЛИРОВАННОЕ ИЗЛУЧЕНИЕ

Представим себе какую-либо систему частиц (например, газ, жидкость, твердое тело), в которой электроны могут находиться в состояниях с двумя энергиями, 1 и 2, при этом возможны два типа переходов.

Переход из нижнего состояния в верхнее требует энергии $\Delta\mathcal{E}$; поэтому такие переходы могут быть осуществлены, если электрон в результате взаимодействия (которое мы называем столкновением) с какой-либо другой частицей (фононом, фотоном или другим электроном) получит необходимую энергию.

Переход из одного состояния в другое в результате взаимодействия с другой частицей называется вынужденным или индуцированным. Таким образом, переходы «снизу — вверх» могут быть только вынужденными.

Переход из верхнего состояния в нижнее не требует затраты энергии, поэтому он может происходить двумя путями: 1) спонтанно (самопроизвольно) и при этом освободившаяся энергия выделится в виде кванта света и 2) вынужденно (индуцированно), т. е. в результате столкновения с какой-либо другой частицей; при этом освободившаяся энергия может также выделиться в виде кванта света, фонона или перейти в кинетическую энергию одной из частиц, участвовавших в столкновении.

Для того чтобы индуцированный переход осуществился под действием кванта света, этот квант должен удовлетворять резонансному условию

$$h\nu = \Delta\mathcal{E}.$$

В результате этого рождается второй фотон с той же частотой, с той же фазой, с тем же направлением движения и поляризации, т. е. неотличимый от первого.

В результате такого индуцированного (или как его также называют стимулированного) перехода энергия электронной системы превращается в энергию излучения, и можно сказать, что имеет место отрицательное поглощение. На этом явлении в принципе могут быть основаны и усиление и генерация электромагнитного излучения. Но для этого необходимо соблюдение двух условий:

— чтобы увеличение энергии электромагнитных колебаний за счет индуцированных переходов превышало бы уменьшение этой энергии за счет всех видов поглощения, т. е. чтобы отрицательное поглощение превалировало над положительным;

— чтобы энергия электронной системы, расходуемая на индуцированное излучение, все время компенсировалась, т. е. каким-то образом все время осуществлялась «подкачка» энергии в электронную систему.

Рассмотрим, как может осуществляться каждое из этих условий.

СОСТОЯНИЯ С ОТРИЦАТЕЛЬНОЙ ТЕМПЕРАТУРОЙ

В тепловом равновесии числа переходов сверху вниз и снизу вверх одинаковы, отношение числа частиц в первом состоянии к числу частиц во втором определяется статистической Ферми:

$$n_1 = \frac{g_1}{e^{\frac{\varepsilon_1 - \mu}{kT}} + 1}, \quad (9.107)$$

$$n_2 = \frac{g_2}{e^{\frac{\varepsilon_2 - \mu}{kT}} + 1}, \quad (9.108)$$

где g_1 и g_2 — «кратность» состояний 1 и 2, т. е. число электронных мест в каждом из этих состояний; согласно (9.107) и (9.108)*)

$$\frac{n_1}{n_2} = \frac{e^{\frac{\varepsilon_2 - \mu}{kT}} + 1}{e^{\frac{\varepsilon_1 - \mu}{kT}} + 1} \frac{g_1}{g_2}. \quad (9.109)$$

*) В дальнейшем мы для простоты будем полагать, что $g_1 = g_2$, и что никаких других состояний у электронов нет. При этих условиях уровень химического потенциала будет проходить посередине между уровнями 1 и 2.

Если электронный газ не вырожден, то статистика Ферми переходит в статистику Больцмана и выражения (9.107) и (9.108) соответственно упрощаются.

Представим теперь, что на нашу систему падает излучение с частотой, удовлетворяющей условию

$$h\nu = \Delta\mathcal{E}. \quad (9.110)$$

Как уже упоминалось, под действием такого излучения возможны два типа переходов:

переход из состояния 1 в состояние 2, связанный с поглощением кванта $h\nu$. Обозначим вероятность такого перехода $B_{1,2}$;

переход из состояния 2 в состояние 1, связанный с испусканием второго кванта (отрицательное поглощение). Обозначим вероятность этого перехода $B_{2,1}$.

Кроме этого, возможны спонтанные переходы из состояния 2 в состояние 1 — $A_{2,1}$.

Эйнштейн вывел соотношения между величинами $B_{2,1}$, $B_{1,2}$ и $A_{2,1}$:

$$B_{2,1} = B_{1,2} \quad (9.111) *$$

и

$$A_{2,1} = \frac{8\pi h\nu^3}{c^3} B_{2,1}. \quad (9.112)$$

Общее число ($N_{1,2}$) переходов первого рода ($1 \rightarrow 2$) в 1 сек будет равно произведению числа частиц в первом состоянии n_1 на вероятность такого перехода $B_{1,2}$:

$$N_{1,2}^u = n_1 B_{1,2}; \quad (9.113)$$

соответственно число обратных переходов

$$N_{2,1}^u = n_2 B_{2,1}. \quad (9.114)$$

Так как $B_{2,1} = B_{1,2}$ и в состоянии теплового равновесия $n_1 > n_2$, то $N_{1,2}^u > N_{2,1}^u$, т. е. в состоянии теплового равновесия система поглощала бы электромагнитное излучение, если бы не было спонтанных переходов $N_{2,1}^c$ **).

*) Если $g_1 \neq g_2$, то $g_2 B_{2,1} = g_1 B_{1,2}$.

***) С учетом спонтанных переходов ($N_{2,1}^c$) суммарное $N_{2,1}^u + N_{2,1}^c = N_{1,2}^u$, и благодаря этому соблюдается закон Кирхгофа: твердое тело в термодинамическом равновесии с излучением столько же излучает, сколько поглощает.

Следовательно, для того чтобы наша система могла усиливать или тем более генерировать электромагнитное излучение, надо с помощью какого-то дополнительного источника энергии создать условия, при которых заселенность верхнего уровня будет больше, чем нижнего:

$$n_2 > n_1.$$

Такие состояния получили название состояний с инверсной заселенностью или с отрицательной температурой *) (так как формально $n_2/n_1 > 1$ можно получить, положив в (9.108) $T < 0$).

Если n_2 не только больше чем n_1 , но равно (или близко) к g_2 и $f_2 = n_2/g_2 \approx 1$, а n_1 близко к нулю и $f_1 = n_1/g_1 \approx 0$, то это особенно благоприятно для усиления и генерации колебаний.

Действительно, при этом переходы с поглощением излучения (из первого во второе состояние) невозможны по двум причинам:

1) так как «некому» переходить, все состояния 1 свободны и

2) так как некуда переходить, все состояния 2 заняты; напротив, для возникновения индуцированного излучения при этом создаются самые благоприятные условия.

Напомним, что согласно статистике Ферми при абсолютном нуле температуры все состояния ниже уровня Ферми заняты, а состояния выше уровня Ферми свободны.

В случае, рассмотренном только что, $f_2 = 1$, $f_1 = 0$, мы имеем обратную картину: такое положение называют состоянием отрицательного нуля, его можно формально получить, если в распределении Ферми считать температуру отрицательной величиной, стремящейся к нулю.

В рассматриваемом нами простейшем случае, когда $g_1 = g_2 = g$ и $n_1 + n_2 = g$, условие $n_2 > n_1$ означает, что верхние уровни должны быть заполнены больше чем наполовину, т. е. электроны должны быть там в состоянии по крайней мере частичного вырождения, а низкие уровни должны быть более чем наполовину пусты, т. е. здесь дырки должны быть в состоянии частичного вырождения.

*) Более правильное название — состояний с инверсной (обратной) заселенностью.

Как мы уже упоминали, для того чтобы усиливать электромагнитное излучение, необходимо настолько нарушить тепловое равновесие, чтобы создать состояние с отрицательной температурой, т. е. чтобы заселенность верхних уровней была больше, чем нижних. В принципе в твердом теле можно представить два способа получения состояния с отрицательной температурой.

а) нарушение внутризонного равновесия, т. е. при помощи какого-то источника энергии электроны в полупроводнике или полуметалле переводятся из нижней половины зоны в верхнюю. Но время восстановления внутризонного равновесия очень мало, $\sim 10^{-13}$ сек, это значит, что такой путь создания инверсной заселенности очень невыгоден, так как требует исключительно мощного источника энергии, а в настоящее время просто неосуществим;

б) нарушение межзонного равновесия: при помощи какого-то источника энергии из валентной зоны в свободную перебрасывается количество электронов, необходимое для создания состояния с отрицательной температурой.

Время релаксации для межзонных переходов колеблется от 10^{-3} до 10^{-9} сек, поэтому эта возможность оказывается вполне реальной. Так как установление внутризонного равновесия происходит значительно быстрее (10^{-13} сек), то мы можем считать, что все переброшенные вверх электроны подавляющую часть «своей жизни» находятся в тепловом равновесии, т. е. сосредоточены вблизи нижнего края зоны проводимости; точно так же все дырки сосредоточены вблизи верхнего края валентной зоны.

Следовательно, усиливаться может излучение, энергия кванта которого приблизительно равна ширине запрещенной зоны *).

Для того чтобы осуществлялось необходимое условие для усиления и генерации ($N_{2,1} > N_{1,2}$), число электронов вблизи нижнего края свободной зоны n_2 должно быть больше, чем число электронов вблизи верхнего края заполненной зоны n_1 . Это значит, что нижние состояния в свободной зоне должны быть более чем наполовину заполнены, а верх-

*) Если энергия кванта будет меньше, ее будет недостаточно, чтобы вызвать переход; если она будет намного больше $\Delta\mathcal{E}$, то соответствующие состояния наверху будут пусты, а внизу заполнены, т. е. не будет условий для отрицательного поглощения.

ние состояния в зоне проводимости — более чем наполовину пусты. Иными словами и электроны и дырки должны находиться в вырожденном (по крайней мере частично) состоянии. На рис. 9.1 представлены зависимости энергии от волнового числа для двух вариантов зонной структуры:

а) в первом случае (рис. 9.1, а) минимум зоны проводимости расположен точно над максимумом валентной зоны; в этом случае переход электрона из одной зоны в другую изображается вертикальной стрелкой, такие переходы называются прямыми;

б) во втором случае (рис. 9.1, б) экстремумы зон разделены некоторым расстоянием в k -пространстве и переход изображается наклонной стрелкой; такой переход, связанный с изменением импульса электрона на величину $\Delta p = p_2 - p_1$, не может осуществляться с участием одного только фотона, так как импульс фотона ничтожно мал*). Такой переход, который может осуществляться только с участием фотона и фонона, называется непрямым.

В этом случае вероятности индуцированного излучения и поглощения могут быть неодинаковы, так как для того чтобы перейти в возбужденное состояние, электрон должен поглотить не только фотон, но и фонон; если температура кристалла низка и необходимых фононов не будет в наличии, то такой переход окажется запрещенным законом сохранения импульса.

Таким образом, второй вариант зонной структуры с этой точки зрения более благоприятен для усиления излучения: здесь нет необходимости в том, чтобы и электроны в верхней зоне и дырки в нижней находились в вырожденном состоянии.

Свет может поглощаться не только при межзонных переходах, но и при внутризонных, т. е. свободными носителями.

Для того чтобы кристалл усиливал излучение, суммарная вероятность всех процессов, связанных с поглощением кванта, должна быть меньше, чем вероятность перехода с индуцированным излучением.

Вероятности прямых переходов ($B_{2,1} = B_{1,2}$) относительно велики и значительно превышают вероятность поглощения свободными носителями (так как последнее

*) Импульс фотона $p_\phi = h\nu/c$ очень мал, так как c — скорость света — очень велика.

требует испускания фотона, т. е. эквивалентно тройному столкновению).

При непрямых переходах оба процесса поглощения свободными носителями, и «выгодный» $2 \rightarrow 1$ и «паразитный», связаны с испусканием фотона и могут иметь сравнимые вероятности; в этом случае поэтому необходимо выбрать такой полупроводник, в котором вероятность межзонных переходов больше, чем внутризонных.

Для получения состояний с инверсной заселенностью могут быть также использованы переходы между примесными уровнями или между одной из зон и примесным уровнем.

СТИМУЛИРОВАННОЕ ИЗЛУЧЕНИЕ В ПОЛУПРОВОДНИКАХ И ПОЛУПРОВОДНИКОВЫЕ ЛАЗЕРЫ

Выше нами были сформулированы условия, необходимые для усиления излучения: 1) рабочие уровни в кристалле должны быть инверсно заселены и 2) полная интенсивность поглощения падающего излучения должна быть меньше интенсивности индуцированного излучения.

Условия для генерации являются более жесткими, так как в данном случае нет падающего излучения, а генератор должен поддерживать и стабилизировать случайно возникшие в нем колебания.

По сути дела любой генератор с самовозбуждением является усилителем шумов, в данном случае этим шумом является спонтанное излучение.

Для создания условий генерации необходима обратная связь *). В газовых лазерах она осуществляется системой зеркал. В случае твердого лазера зеркалами являются его торцы, которые обрабатываются таким образом, что один из них полностью отражает падающую на него энергию, а другой делается полупрозрачным, таким образом, каждый фотон, попавший в систему, несколько раз отражается от зеркал, а затем либо поглощается, либо выходит из кри-

*) Если лавина индуцированного излучения беспрепятственно пройдет через кристалл, то кристалл «погаснет» почти мгновенно и не сможет служить генератором; для того чтобы стимулированное излучение поддерживалось, коэффициент размножения должен быть настолько больше единицы, чтобы компенсировать все потери. Для этого каждый фотон должен несколько раз пройти кристалл, прежде чем выйти.

стала через полупрозрачный торец. Для того чтобы осуществилось самовозбуждение, необходимо, чтобы каждый фотон, прежде чем поглотиться или выйти из системы, породил второй фотон; при этих условиях фотоны будут саморазмножаться и число их будет расти до тех пор, пока энергия излучения, выходящего из кристалла, не станет равной энергии подкачки. В заключение этого раздела рассмотрим, какими способами может осуществляться подкачка и создаваться и поддерживаться инверсная заселенность.

МЕТОДЫ ПОЛУЧЕНИЯ ИНВЕРСНОЙ ЗАСЕЛЕННОСТИ

Оптическое возбуждение является одним из основных методов создания состояния с инверсной заселенностью. Недостатком этого метода является низкий коэффициент преобразования, обусловленный широким спектром обычных источников света. При этом большая часть энергии коротковолновых фотонов теряется (переходит в тепло) при переходе носителей на низшие уровни зоны, а длинноволновые фотоны не участвуют в возбуждении.

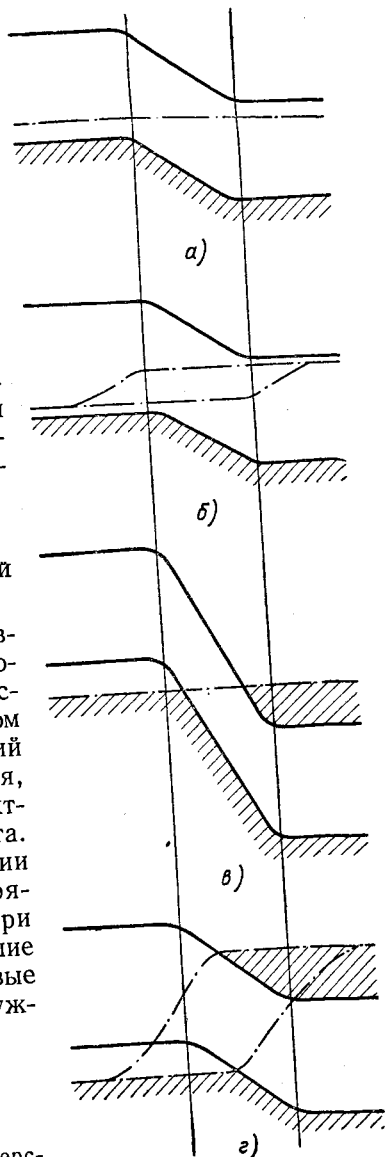


Рис. 9.6. Схема создания инверсной заселенности при помощи $p-n$ перехода.

Таблица интегралов Ферми $F_m(\mu) = \int_0^{\infty} \frac{x^m}{e^{\mu-x} + 1} dx$; $F_m(\mu) = a \cdot 10^b$

| μ | $F_{-\frac{1}{2}}$ | | F_0 | | $F_{\frac{1}{2}}$ | | F_1 | | $F_{\frac{3}{2}}$ | | F_2 | | $F_{\frac{5}{2}}$ | | F_3 | |
|-------|--------------------|----|--------|----|-------------------|----|--------|---|-------------------|----|--------|---|-------------------|----|--------|----|
| | a | b | a | b | a | b | a | b | a | b | a | b | a | b | a | b |
| -4 | 0,3204 | -1 | 0,1815 | -1 | 0,16128 | -1 | 0,0182 | 0 | 2,42685 | -2 | 0,0366 | 0 | 6,07731 | -2 | 1,0977 | -1 |
| -3 | 0,8526 | -1 | 0,4859 | -1 | 0,43366 | -1 | 0,0492 | 0 | 6,56115 | -2 | 0,0990 | 0 | 1,64742 | -1 | 2,9780 | -1 |
| -2 | 2,1918 | -1 | 1,2693 | -1 | 1,14588 | -1 | 0,1310 | 0 | 1,75800 | -1 | 0,2662 | 0 | 4,44554 | -1 | 8,0532 | -1 |
| -1 | 5,2114 | -1 | 3,1326 | -1 | 2,90501 | -1 | 0,3387 | 0 | 4,60848 | -1 | 0,7050 | 0 | 1,18597 | 0 | 2,1598 | 0 |
| 0 | 1,0722 | -1 | 6,9315 | -1 | 6,78094 | -1 | 0,8225 | 0 | 1,15280 | 0 | 1,8030 | 0 | 3,08259 | 0 | 5,6822 | 0 |
| 1 | 1,8204 | 0 | 1,3137 | 0 | 1,39638 | 0 | 1,8062 | 0 | 2,66168 | 0 | 4,3120 | 0 | 7,62653 | 0 | 1,4395 | 0 |
| 2 | 2,5954 | 0 | 2,1270 | 0 | 2,50246 | 0 | 3,5135 | 0 | 5,53725 | 0 | 9,4450 | 0 | 1,75294 | 0 | 3,4304 | 0 |
| 3 | 3,2852 | 0 | 3,0486 | 0 | 3,97699 | 0 | 6,0957 | 0 | 1,03537 | 1 | 1,8870 | 1 | 3,69321 | 1 | 7,5722 | 1 |
| 4 | 3,8743 | 0 | 4,0182 | 0 | 5,77073 | 0 | 9,6267 | 0 | 1,76277 | 1 | 3,4592 | 1 | 7,13480 | 1 | 1,5421 | 1 |
| 5 | 4,3832 | 0 | 5,0067 | 0 | 7,83797 | 0 | 1,4138 | 1 | 2,78024 | 1 | 5,8120 | 1 | 1,27489 | 2 | 2,9094 | 2 |
| 6 | 4,8338 | 0 | 6,0025 | 0 | 1,01443 | 1 | 1,9642 | 1 | 4,12610 | 1 | 9,1744 | 1 | 2,13098 | 2 | 5,1300 | 2 |
| 7 | 5,2416 | 0 | 7,0009 | 0 | 1,26646 | 1 | 2,6144 | 1 | 5,83422 | 1 | 1,3736 | 2 | 3,36814 | 2 | 8,5342 | 2 |
| 8 | 5,6170 | 0 | 8,0003 | 0 | 1,53805 | 1 | 3,3645 | 1 | 7,93526 | 1 | 1,9699 | 2 | 5,08084 | 2 | 1,3512 | 3 |
| 9 | 5,9674 | 0 | 9,0001 | 0 | 1,82776 | 1 | 4,2145 | 1 | 1,04574 | 2 | 2,7261 | 2 | 7,37087 | 2 | 2,0513 | 3 |
| 10 | 6,2972 | 0 | 1,0000 | 1 | 2,13445 | 1 | 5,1645 | 1 | 1,34270 | 2 | 3,6623 | 2 | 1,03468 | 3 | 3,0048 | 3 |
| 11 | 6,6096 | 0 | 1,1000 | 1 | 2,45718 | 1 | 6,2145 | 1 | 1,68688 | 2 | 4,7986 | 2 | 1,41237 | 3 | 4,2687 | 3 |
| 12 | 6,9076 | 0 | 1,2000 | 1 | 2,79518 | 1 | 7,3645 | 1 | 2,08062 | 2 | 6,1548 | 2 | 1,88225 | 3 | 5,9060 | 3 |
| 13 | 7,1930 | 0 | 1,3000 | 1 | 3,14775 | 1 | 8,6145 | 1 | 2,52616 | 2 | 7,7510 | 2 | 2,45700 | 3 | 7,9856 | 3 |
| 14 | 7,4672 | 0 | 1,4000 | 1 | 3,51430 | 1 | 9,9645 | 1 | 3,02564 | 2 | 9,6072 | 2 | 3,14983 | 3 | 1,0592 | 4 |
| 15 | 7,7314 | 0 | 1,5000 | 1 | 3,89430 | 1 | 1,1415 | 2 | 3,58112 | 2 | 1,1743 | 3 | 3,97448 | 3 | 1,3778 | 4 |
| 16 | 7,9868 | 0 | 1,6000 | 1 | 4,28730 | 1 | 1,2965 | 2 | 4,19458 | 2 | 1,4180 | 3 | 4,94522 | 3 | 1,7659 | 4 |
| 17 | 8,2342 | 0 | 1,7000 | 1 | 4,69286 | 1 | 1,4615 | 2 | 4,86794 | 2 | 1,6936 | 3 | 6,07677 | 3 | 2,2318 | 4 |
| 18 | 8,4744 | 0 | 1,8000 | 1 | 5,11061 | 1 | 1,6365 | 2 | 5,60305 | 2 | 2,0032 | 3 | 7,38433 | 3 | 2,7854 | 4 |
| 19 | 8,7076 | 0 | 1,9000 | 1 | 5,54019 | 1 | 1,8215 | 2 | 6,40171 | 2 | 2,3488 | 3 | 8,88359 | 3 | 3,4373 | 4 |
| 20 | 8,9350 | 0 | 2,0000 | 1 | 5,98128 | 1 | 2,0165 | 2 | 7,26568 | 2 | 2,7325 | 3 | 1,05906 | 4 | 4,1985 | 4 |

| μ | $F_{\frac{7}{2}}$ | | F_4 | | $F_{\frac{9}{2}}$ | | F_6 | | $F_{\frac{11}{2}}$ | | F_8 | | F_7 | |
|-------|-------------------|----|--------|---|-------------------|----|--------|---|--------------------|---|--------|---|--------|---|
| | a | b | a | b | a | b | a | b | a | b | a | b | a | b |
| -4 | 2,12877 | -1 | 0,4390 | 0 | 9,58334 | -1 | 0,2198 | 1 | 5,27190 | 0 | 1,3195 | 1 | 0,0923 | 3 |
| -3 | 5,77852 | -1 | 0,1194 | 1 | 2,60317 | 0 | 0,5948 | 1 | 1,43253 | 1 | 0,3582 | 2 | 0,2509 | 3 |
| -2 | 1,56497 | 0 | 0,3234 | 1 | 7,06296 | 0 | 1,6210 | 1 | 3,89035 | 1 | 0,9725 | 2 | 0,6818 | 3 |
| -1 | 4,21325 | 0 | 0,8721 | 1 | 1,91050 | 1 | 0,4383 | 2 | 1,05487 | 2 | 0,2641 | 3 | 0,1852 | 4 |
| 0 | 4,21837 | 1 | 2,3340 | 1 | 5,12904 | 1 | 0,1183 | 3 | 2,84903 | 2 | 0,7148 | 3 | 0,5020 | 4 |
| 1 | 2,88313 | 1 | 6,0981 | 1 | 1,35419 | 2 | 0,3151 | 3 | 7,62405 | 2 | 1,9223 | 3 | 1,3597 | 4 |
| 2 | 7,07645 | 1 | 1,2324 | 2 | 3,46603 | 2 | 0,8174 | 3 | 2,00234 | 3 | 5,0913 | 3 | 3,6422 | 4 |
| 3 | 1,62566 | 2 | 3,6369 | 2 | 8,46252 | 2 | 2,0404 | 3 | 5,09479 | 3 | 1,3148 | 4 | 9,6007 | 4 |
| 4 | 3,46758 | 2 | 8,1464 | 2 | 1,94721 | 3 | 4,8738 | 3 | 1,23991 | 4 | 3,2619 | 4 | 2,4514 | 5 |
| 5 | 6,87309 | 2 | 1,6749 | 3 | 4,20034 | 3 | 1,0821 | 4 | 2,86055 | 4 | 7,7515 | 4 | 6,1153 | 5 |
| 6 | 1,27353 | 3 | 3,2492 | 3 | 8,50005 | 3 | 2,2763 | 4 | 6,23266 | 4 | 1,7435 | 5 | 1,4515 | 6 |
| 7 | 2,22342 | 3 | 5,9365 | 3 | 1,62060 | 4 | 4,5161 | 4 | 1,28314 | 5 | 3,7142 | 5 | 3,2837 | 6 |
| 8 | 3,68668 | 3 | 1,0286 | 4 | 2,92792 | 4 | 8,4889 | 4 | 2,50387 | 5 | 7,5073 | 5 | 7,0728 | 6 |
| 9 | 5,84734 | 3 | 1,7016 | 4 | 5,04304 | 4 | 1,5198 | 5 | 4,65139 | 5 | 1,4445 | 6 | 1,4522 | 7 |
| 10 | 8,92629 | 3 | 2,7034 | 4 | 8,32807 | 4 | 2,6052 | 5 | 8,26504 | 5 | 6,6570 | 6 | 2,8499 | 7 |
| 11 | 1,31835 | 4 | 4,1468 | 4 | 1,32532 | 5 | 4,2967 | 5 | 1,41122 | 6 | 4,6915 | 6 | 5,3628 | 7 |
| 12 | 1,89204 | 4 | 6,1682 | 4 | 2,04150 | 5 | 6,8481 | 5 | 2,32527 | 6 | 7,9845 | 6 | 9,7103 | 7 |
| 13 | 2,64816 | 4 | 8,9575 | 4 | 3,05550 | 5 | 1,0588 | 6 | 3,71137 | 6 | 1,3146 | 7 | 1,6975 | 8 |
| 14 | 3,62572 | 4 | 1,2626 | 5 | 4,45804 | 5 | 1,5934 | 6 | 5,75745 | 6 | 2,1011 | 7 | 2,8744 | 8 |
| 15 | 4,86842 | 4 | 1,7476 | 5 | 6,35840 | 5 | 2,3406 | 6 | 8,70635 | 6 | 3,2692 | 7 | 4,7278 | 8 |
| 16 | 6,42490 | 4 | 2,3739 | 5 | 8,88666 | 5 | 3,3645 | 6 | 1,28667 | 7 | 4,9651 | 7 | 7,5740 | 8 |
| 17 | 8,34884 | 4 | 3,1707 | 5 | 1,21959 | 6 | 4,7429 | 6 | 1,86247 | 7 | 7,3774 | 7 | 1,1846 | 9 |
| 18 | 1,06992 | 5 | 4,1710 | 5 | 1,64645 | 6 | 6,5692 | 6 | 2,64579 | 7 | 1,0746 | 8 | 1,8125 | 9 |
| 19 | 1,35402 | 5 | 5,4121 | 5 | 2,18987 | 6 | 8,9541 | 6 | 3,69492 | 7 | 1,5372 | 8 | 2,7183 | 9 |
| 20 | 1,69419 | 5 | 6,9355 | 5 | 2,87348 | 6 | 1,2028 | 7 | 5,08033 | 7 | 2,1629 | 8 | 4,0026 | 9 |

Использование излучения другого лазера для возбуждения полупроводникового лазера могло бы устранить этот недостаток. Однако такой прием, разумеется, может быть использован лишь в исследовательской работе (для отбора полупроводниковых материалов и др.).

Метод возбуждения электронным пучком обладает тем преимуществом, что позволяет подбирать энергию пучка, строго соответствующую энергии возбуждения, что позволяет получать интенсивную заселенность в полупроводниках с различной запрещенной зоной. Недостатком его является то, что к. п. д. при этом не может превышать 40%, так как, как показали расчеты, энергия, идущая на образование электронно-дырочной пары, приблизительно в 3 раза превышает ширину запрещенной зоны.

Наиболее перспективным является метод инъекции неравновесных носителей тока через $p-n$ переход вырожденных полупроводников, так как позволяет получить очень высокий коэффициент преобразования.

На рис. 9.6, *а* представлен обычный $p-n$ переход в равновесном состоянии (обычный в том смысле, что основные носители по обе стороны $p-n$ перехода не вырождены). На рис. 9.6, *б* представлен тот же переход при приложении небольшого напряжения в пропускном направлении. Как видно из рисунка, этого напряжения недостаточно, чтобы создать инверсную заселенность.

На таком переходе инверсная заселенность может быть создана только при напряжении, полностью снимающем потенциальный барьер, однако создать такие условия весьма трудно, так как ток через $p-n$ переход при этом будет очень велик.

На рис. 9.6, *в* и *г* представлены аналогичные схемы для вырожденного $p-n$ перехода. Как видно из рисунка, в этом случае для создания инверсной заселенности достаточно сравнительно небольшого напряжения.

Нетрудно убедиться, что для такой благоприятной ситуации достаточно, чтобы носители находились в вырожденном состоянии хотя бы по одну сторону от $p-n$ перехода.

На этом мы заканчиваем весьма краткое описание принципа действия полупроводниковых лазеров. Теория лазеров довольно подробно изложена в книге Г. Е. Пикуса [31].

1. Иоффе А. Ф. Физика полупроводников. Изд-во АН СССР, 1957.
2. «Полупроводники в науке и технике». Сборник статей под ред. академика А. Ф. Иоффе. Изд-во АН СССР, 1956—1957.
3. Соминский М. С. Полупроводники. Физматгиз, 1961.
4. Федотов Я. А. Основы физики полупроводниковых приборов. Изд-во «Советское радио», 1963.
5. Де Бройль Л. Революция в физике. Изд-во «Наука», 1965.
6. Компанеев А. С. Теоретическая физика. Гостехиздат, 1955.
7. Ландау Л. Д. и Лившиц Е. М. Квантовая механика. Часть I. Нерелятивистская теория. Физматгиз, 1963; Статистическая физика. Изд-во «Наука», 1964.
8. Жданов Г. С. Физика твердого тела. Изд-во МГУ, 1961.
9. Ансельм А. И. Введение в теорию полупроводников. Физматгиз, 1962.
10. Киттель Ч. Введение в физику твердого тела. Физматгиз, 1963.
11. Зейтц Ф. Современная теория твердого тела. Гостехиздат, 1949.
12. Займан Дж. Электроны и фононы. Изд-во иностранной литературы, 1962.
13. Шокли В. Теория электронных полупроводников. Изд-во иностранной литературы, 1953.
14. Смит Р. Полупроводники. Изд-во иностранной литературы, 1962.
15. Болтакс Б. И. Диффузия в полупроводниках. Физматгиз, 1961.
16. Цидильковский И. М. Термомагнитные явления в полупроводниках. Физматгиз, 1960.
17. Рывкин С. М. Фотоэлектрические явления в полупроводниках. Физматгиз, 1963.
18. Бьюб Р. Фотопроводимость твердых тел. Изд-во иностранной литературы, 1962.
19. Мосс Т. С. Оптические свойства полупроводников. Изд-во иностранной литературы, 1964.
20. Тауц Я. Фото- и термоэлектрические явления в полупроводниках. Изд-во иностранной литературы, 1962.

21. Драбл Дж. и Голдсмит Г. Теплопроводность полупроводников. Изд-во иностранной литературы, 1963.
22. Мотт и Герни. Электронные процессы в ионных кристаллах. Изд-во иностранной литературы, 1940.
23. Паулинг Л. Природа химической связи. Госхимиздат, 1947.
24. Физика твердого тела. АН СССР, Ин-т научной информации, 1965.
25. Лендьеел Б. Лазеры. Изд-во «Мир», 1964.
26. Жузе В. П., Гусенкова Е. И. Библиография по термоэлектричеству. Изд-во АН СССР, 1963.
27. Поверхностные свойства полупроводников. Под ред. А. Н. Фрумкина, А. В. Ржанова и Р. Х. Бурштейн. Изд-во АН СССР, 1962.
28. Полупроводниковые электронные приборы. Сб. переводов под ред. А. В. Ржанова. Изд-во иностранной литературы, 1953.
29. Богомолов В. Н. Устройства с датчиками Холла и датчиками магнетосопротивления. Госэнергоиздат, 1961.
30. Callaway Y. Phys. Rev., 1959, v. 133, p. 1046.
31. Пикве Г. Е. Основы теории полупроводниковых приборов. Изд-во «Наука», 1966.
32. Вавилов В. С. Действие излучения на полупроводники. Физматгиз, 1963.
33. Роуз А. Основы теории фотопроводимости. Изд-во «Мир», 1966.
34. Равич Ю. И. Фотомагнитный эффект в полупроводниках и его применение. Изд-во «Советское радио», 1967.
35. Займан Дж. Принципы теории твердого тела. Изд-во «Мир», 1966.
36. Киттель. Элементарная физика твердого тела. Изд-во «Наука», 1965.
37. Губанов А. И. Теория выпрямляющего действия полупроводников. Гостехиздат, 1956.
38. Томсон, Уэрт. Физика твердого тела. Изд-во «Мир», 1966.
39. Адирович Э. И. Некоторые вопросы теории люминесценции кристаллов. Гостехиздат, 1956.
40. Петровский И. И. Электронная теория полупроводников. г. Минск. Изд-во «Высшая школа», 1964.
41. Ван дер Зил. Флюктуационные явления в полупроводниках. Изд-во иностранной литературы, 1961.
42. Епифанов Г. И. Физика твердого тела. Изд-во «Высшая школа», 1964.
43. Spenke E. Elektronische Halbleiter. Berlin. Springer-Verlag, 1956.
44. Блекмор Дж. Статистика электронов в полупроводниках. Изд-во «Мир», 1964.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Адиабатическое приближение
190—194, 196

Альфа-подход (в теории термоэлектричества) 302

Ангармоничность (колебаний)
48, 54, 55, 186, 319

Антизапорный слой 64, 65

Атомные кристаллы 170

Базисный вектор (решетки)
152, 155, 157, 160, 198, 210

Бозе — Эйнштейна распределение 322, 324, 328

Больцмана распределение 184,

Бравэ решетки 156, 157

Бриллюэна зоны 204, 205,
211, 213, 216

Валентная зона 13, 14, 16,
18, 36, 109, 203, 434, 435

Вероятность нахождения частицы 122, 134

— перехода 130, 131, 138, 141

«Вертикальные» переходы —

см. переходы

Видемана и Франца закон 45,
256, 311, 315, 316

Волна электрона 121, 122, 128,
158, 249

Волновая функция 122, 123,
130, 190, 192, 196—200, 274,
275, 285

Волновое число 47, 160, 259,
285

Волновой вектор 47, 53, 117,
122, 159—161, 176, 179,
181, 202, 320, 322

Волны упругие 46, 176, 317,
320

Вольтамперная характеристика
контакта двух металлов 61

— — — полупроводника и металла
в диодной теории 66—68,
373—376

— — — — — в диффузионной
теории 66, 68, 69, 368—373,
376—378

— — *p-i-n* перехода 396

— — *p-n* перехода 72—75, 385—
396

— — фотоэлемента 425

Время жизни 72, 106, 392,
409, 416, 420

Время релаксации 41, 218—
220, 256, 258—261, 267—
278, 355, 404, 415

— —, энергетическая зависимость
256, 260, 261, 271, 343,
357

— — фононов 300, 301, 317,
323—325

— —, дисперсионные зависимости
325—326

Время свободного пробега 37,
100, 247, 249, 257, 259—
261, 279, 429

Выпрямление на контакте
двух металлов 60, 61

— — — — — полупроводников
61

— — — — — полупроводника и
металла 66—70

Вырождение (см. также «масса
эффективная», «невыврожденное»,
«электронный газ», «дырочный газ»,
«состояния», «полупроводники»
31, 78, 235, 282, 339, 413

- , кратность 224
 - , критерий 242
 - ориентационное 235
 - полное 237, 241, 243, 342
 - , связь с эффективной массой 240
 - сильное 44, 242, 243
 - , снятие 31, 137, 138, 201, 235
 - , степень 30, 238, 261, 312, 316
 - частичное 31, 35, 239, 242, 433
- Гальвано-магнитные явления 83, 85—98, 99, 244, 331, 334, 341, 351
- —, коэффициенты 270
 - — поперечные, продольные 92, 99, 351
- Гамильтона оператор (гамильтониан) 131—134, 139, 142, 190, 207
- Генерация носителей 101, 104, 378, 379, 418
- Групповая скорость 117, 179, 180
- — распространения тепловых колебаний 49
 - — электронной волны 202, 236
- Дебаевский радиус экранирования 288, 363
- Дебая температура 41, 51, 52, 54, 55, 89, 96, 97, 187, 255, 276, 277, 283, 327
- функция 188
- Дембера эффект 422
- Дефекты решетки 34, 36, 37, 39, 48, 163—174, 319, 324
- Диффузии коэффициент 69, 72, 364, 422
- — для ионов 172, 173
- Диффузионная длина 72, 365, 381—383, 392, 394, 423
- Диффузионная теория — см. выпрямление
- Диффузия биполярная 82, 349, 357
- ионов 171
 - носителей 68, 72, 74, 78, 363, 397, 421, 422
- Диффузия электронно-дырочных пар 45
- Длина свободного пробега 37, 39, 40, 42, 44, 48, 49, 54, 66, 68, 81, 94, 96, 97, 247, 249, 257, 261, 275—276, 279, 317, 324, 337, 343, 368, 373, 374
- Емкость p - n перехода 280, 301, 364, 366, 376, 389—392
- Жидкость фермиевская 239
- электронная 30, 238
- Запорный слой 63, 64, 66, 72, 290, 363, 368—378, 423
- Заселенность инверсная 433, 434, 436, 437, 440
- Захват носителей 104, 105
- Зеркального изображения силы 371, 372
- Зона (энергетическая) 11—15, 56, 100, 209—216
- запрещенная 16, 18, 20, 32, 46, 108, 166, 202, 235, 286, 289, 306, 399, 416, 434, 435, 440
- Импульс фонона 53, 54, 118, 254, 258
- фотона 108, 109, 405, 408—409
 - электрона (см. также квазиимпульс) 27, 121, 190, 254, 259, 405—406
- Импульса закон сохранения 54, 104, 108, 110, 111, 251, 253, 405, 435
- Инжекция 74, 75, 362, 380—385, 394—395, 440
- Интеграл перекрытия 207, 208
- столкновений 267, 269, 323
- Ионизация двухступенчатая 410
- термоэлектронная 285—287, 373
 - ударная 285, 289—291, 362, 365, 384, 392
 - электростатическая 285, 362, 375, 384
- Квазиимпульс фонона 253
- электрона 27, 214, 241, 253, 257
- Квазистатический процесс 220

- Квазиуровень Ферми 382—384, 417
 Квазичастицы 219, 253
 Квантовый выход 106, 112
 Кванты энергии волны 114, 116
 — — излучения 120
 — — осциллятора 51, 53
 Кикоина — Носкова фотомагнитный эффект 422—423
 Кинетическое уравнение (Больцмана) 260—270, 340, 341
 — — для фононов 318, 325, 328
 Ковалентные кристаллы 89, 144, 147, 176, 196, 197
 Когерентного излучения генераторы 142
 Колебания атомов решетки, акустические 89, 181—183, 188, 317, 318
 — — — оптические 41, 89, 112, 181—183, 189, 255, 276, 283, 314, 316, 317
 — — — поляризация 176
 — — — поперечные, продольные 176, 317, 318
 — — — тепловые 14, 15, 29, 36, 37, 39, 40, 146, 148, 166, 174—184, 237, 238, 274—278, 409
 Контактная разность потенциалов 58, 59, 63, 78, 367, 368, 381, 395
 Контактные явления 55, 63, 362—399
 Концентрация носителей тока 19, 22, 24, 30, 35, 44, 68, 77, 88, 100, 226, 229—234, 241, 278, 304—306, 418
 — — , влияние поля 284—291
 — — — неравновесная 102—103, 380, 382, 384
 — электронов в металлах и полуметаллах 19, 30, 31, 43, 242
 Край полосы поглощения 108
 Кристаллическая решетка 36, 147—163, 171, 208, 211, 281
 Лавинный эффект 363, 365
 Лазер 104, 165, 436, 437, 440
 Ландау уровни 340
 Ловушки (см. также прилипание) 416—417, 420
 Лоренца сила 84, 85, 264, 331, 333, 423
 — число 44, 257, 316
 Магнитная сила 84, 85, 86, 90, 91, 95, 97, 332, 334, 335
 Магнитное поле сильное 94, 95, 338, 339, 341, 351—355
 — — —, классический и квантовый критерии 340, 351
 — — слабое 93, 94, 95, 337, 338, 341, 347
 Маджи — Риги — Ледюка эффект 100
 Максвелла — Больцмана распределение 35, 227, 239, 328, 347
 — закон распределения для скоростей 15
 Масса эффективная 24, 27, 35, 42, 120, 212, 213, 214, 277, 331, 392, 427
 Мелкие уровни 120
 Металлическая связь 144, 147
 Металлы 13, 16, 18, 58—62, 78, 176, 237, 242, 298
 —, схема валентной зоны 16
 — щелочно-земельные 17, 238
 — щелочные 16, 34, 199, 237
 Миллеровские индексы 163
 Модуль сжимаемости 49
 Невырожденное состояние электронного газа 30, 31, 35, 44, 77, 81, 227, 277
 Неопределенностей соотношения Гейзенберга 124, 162
 Неопределенности принцип 123—128, 175, 249
 Непрерывности уравнение 378—379, 389
 Нернста — Эттинггаузена эффекты 98—100, 356—358, 423

- Обратная решетка 158, 203, 205, 210
 — —, вектор 161, 200, 203, 321, 322
 — —, — базисный 162
- Обратное пространство 159, 161, 162, 198, 210
 — —, основная (или приведенная) область 160, 161
 — —, элементарная ячейка 160
- Обращенный диод 364, 397
- Объемный заряд 62, 64
- Одноэлектронное приближение 194—198, 209
- Паули принцип 13, 56, 126, 175, 196, 199
- Пельтье явление 63, 75, 76, 79, 292, 293, 300, 303
 — коэффициент 76, 81, 82, 303
 — теплота 80
- Переноса явления 244—292, 413
 — —, феноменологический анализ 270—271
- Переход *p-i-n* 393—396
 — электронно-дырочный 70, 71, 280, 362, 363, 366, 378—399, 423, 437, 440
 — — на границе полупроводник — металл 65, 66
 — — толстый, тонкий 72, 392—396
- Переходы (носителей) — см. также «рекомбинация»
 — безызлучательные 104, 194
 — «вертикальные» 256, 405
 — межзонные (прямые) 107, 108, 404—407, 435
 — «непрямые» 406, 407, 435, 436
 — при столкновениях горизонтальные, вертикальные 313—314
- Пи-подход (в теории термоэлектричества) 294, 302
- Писаренко формула 82
- Плотность состояний (электронных) 28, 32, 33, 42, 227—229, 308, 339, 382, 417, 418
- Поглощение (излучения, света) 107, 400—409
 — отрицательное 431, 432
 — примесное 107, 108, 110
 — свободными носителями 106, 112, 402—404, 435
 — собственное (фундаментальное) 107, 109, 409
 — —, красная граница 410
 —, спектр 107, 408
 —, спектральное распределение 409
 — экситонное 111, 407—410
- Подвижность ионов 173, 174
- Подвижность носителей тока 19, 36, 245, 256, 257, 260, 403, 410
 — — —, температурная зависимость 19, 256, 261
 — — —, экспериментальное определение 88, 97, 356
 — —, зависимость от электрического поля 279—284
- Полупроводники 13, 18
 — вырожденные (см. также вырождение) 35, 89, 227, 357, 407
 — ковалентные 41, 89
 — невырожденные 40, 41, 92, 237, 418
 — примесные 20, 31
 — — дырочные 23, 24, 36, 232
 — — электронные 21, 31, 35, 232
 — с атомной решеткой 96, 98
 — с ионной решеткой 41, 89, 96, 97
 — собственные 20, 32
- Полярон 192—194, 249
- Потенциал ионизации (ионизационный) 11, 55, 56, 145—147, 166
 — решетки периодический—195—198, 200—206, 210, 214, 412
 — химический (см. также уровень Ферми) 222, 223
 — — приведенный 226, 241, 304
 — —, уровень 33, 35, 57, 58, 71, 78, 80, 166, 227, 241, 294, 309, 381, 396, 431

- электростатический 157, 222
- Поток тепла 46, 98, 245, 318, 321, 326, 355
- энергии 245, 270, 329
- Преломления показатель (коэффициент) 112, 400, 401
- Приближение почти свободных электронов 198—206
- сильно связанных электронов 206—209
- Прилипание (см. также ловушки) 414
- , уровень 105
- Прищельное расстояние 248, 258
- Проводимость ионная 171—175
- Процессы нормальные («N» при фонон-фононных взаимодействиях) 321, 323, 326
- переброса (umclapp) 321—324, 326
- Работа выхода 55—58, 62, 64, 70, 78, 373
- Распределение фононов — см. Бозе
- — неравновесное 322, 323, 325, 329
- электронов по волновым векторам 266
- — — —, неравновесная добавка 269—270, 296
- — — —, — функция 266, 270, 315
- — — — скоростям 262, 263, 344
- — — —, изменение под действием градиента температуры 313, 314
- — — —, — — — электрического поля 262, 263, 312, 313, 344
- — — —, неравновесное 263—265, 312, 313, 345
- Рассеяние фононов 319
- — на фононах 319, 321, 324
- электронов 39
- — межзонное 310
- — на дефектах решетки 39, 245, 247, 314
- — на тепловых колебаниях решетки 40, 54, 81, 89, 261, 263, 274—278, 281, 282, 316
- — примесное 40, 81, 89, 97, 248, 251, 258, 261, 263, 271—274, 277, 281, 282, 394
- Рекомбинация 18, 45, 72, 74, 101, 104, 363, 378, 392, 397, 413—421, 426
- безызлучательная 104
- , вероятность 102, 106
- излучательная 104
- квадратичная 414
- линейная 414
- поверхностная 72, 409, 424, 426,
- прямая (межзонная) 101, 104, 105
- ударная 106
- через центры 101, 104, 105, 416
- экситона 112
- Риги — Ледюка эффект 98, 355—356
- Симметрия кристалла 198
- трансляционная 111, 157, 158, 192, 193, 198, 209—211
- Скорость дрейфа электронов 37, 38, 69
- теплового движения электронов 37, 39, 40, 78
- — — — средняя 43, 68, 254, 279
- Скрещенные электрическое и магнитное поля 334, 335
- Слой объемного заряда 63, 71, 383
- — — на контакте металл — полупроводник 64
- — —, толщина 380
- Столкновение 248—251
- Столкновения, вероятность 219, 246
- межэлектронные, неупругие, упругие, почти упругие 251, 252, 254—258, 268, 269, 283, 284, 314—316
- Столкновения фононов с дефектами, с фононами 54, 219, 320

- — — —, 2-3-и 4-фононные 320
- , частота 219, 246
- электронов с фононами 54, 251, 253, 254, 258
- Тепловое возбуждение электрона 15, 20
- — —, частота следования 15, 18
- Теплоемкость решетки 43, 49, 50, 52, 184—189, 317
- фотонная 329
- электронного газа 44, 311
- Теплопроводность 244, 256, 270, 271
- решеточная 14, 43, 46—49, 219, 317—329, 351
- фотонная 311, 329—330
- электронная 43—44, 46, 257, 264, 311—316, 351, 356
- —, отношение к электропроводности 44, 45
- Термомагнитные коэффициенты 257, 270
- Термомагнитные явления 83, 85, 98—100, 244, 331, 334, 355—357
- Термоэлектрические явления 75, 76, 78, 82, 100, 244, 264, 292—311, 315, 351, 422—423
- Томсона коэффициент 77
- соотношения 76—77
- теплота 77
- явление 75, 77, 79
- Туннельный диод 165, 280, 286, 364, 396—399
- эффект 11, 59, 62, 66, 193, 285, 363, 368, 373, 392, 396
- Увлечение электронов фононами 299—304
- Уровень (энергетический) см. энергетический
- Уровень Ферми (см. также потенциал химический) 33, 34, 59, 71, 80, 81, 105, 433
- — в невырожденном полупроводнике 82, 229—232, 234
- — в невырожденном полупроводнике, температурная зависимость 35, 36, 232, 233
- —, температурная зависимость 35, 36
- Фазовая скорость 177, 179, 180, 253
- Фазовое пространство 265—267, 328
- Ферми — Дирака статистика 199, 431
- интегралы 241, 243, 299
- поверхность 314, 339, 342
- распределение 32—34, 224—226, 328, 417, 431
- Фонона энергия 53, 104, 118, 253—255
- Фононы 49, 54, 104, 118, 186, 187, 255, 299—303, 317, 318, 324, 405—407
- , групповая скорость 317
- Фотовольтаические эффекты 421—426
- Фотодиод 425, 426
- Фотоионизация 384
- Фотопроводитель 100, 413, 421
- Фотопроводимость 100—103, 108, 409—421
- примесная, собственная 410—412
- стационарная 103
- Фоточувствительность 107, 109, 110, 409, 410
- Фото-э. д. с. 421
- Фотозффект 100, 109
- вентиляльный 423—426
- , квантовый выход 423
- , красная граница 410
- примесный 409
- «собственный» 409
- Фундаментальное поглощение, см. поглощение
- Химическая связь 143—147
- Холла постоянная 87—91, 347, 349—352
- э. д. с. 90, 341, 349
- эффект 86, 92, 341—352, 356
- Холловское поле 87, 90, 91, 346, 351, 354
- Центр рекомбинации примесный 101, 105, 110

Цепочка двухатомная 180—182

Циклотронный резонанс 334, 426—430

Шокли — Риды статистика рекомбинации 416—421

Шредингера уравнение 128—134, 190, 274

— — для нестационарных процессов 131—138

— — для электронов в адиабатическом приближении 192

— — — в одноэлектронном приближении 195

— — невозмущенное 135

Эйнштейна соотношение 69, 174, 370

Экситоны 106, 111, 196, 407—409

Экстракция 75

Электрическое поле сильное 278—291, 362

Электронный газ, вырожденное состояние 30—32, 34, 237—239, 256, 262, 406, 407, 435

— —, невырожденное состояние 30, 31, 35, 44, 77, 81, 227, 238, 239, 256, 262, 299

— — невырожденный, статистика (для полупроводников) 226—231

— —, плотность 36

Электронный газ, сохранение теплового равновесия 257 с кристаллом

Электрон-фононное взаимодействие 255

Электрохимический потенциал 79, 224, 293

Элементарная ячейка — см. ячейка

Эллипсоид энергии 216

Энергетический уровень захвата 416

— — электрона в кристалле 236

Энергетический спектр электронов в кристалле 10—13, 20, 209, 236, 244, 249

Энергия активации 11, 19, 27, 31, 172, 339, 341, 407

— — оптическая 410, 412, 413

— — примесных электронов 22, 35, 120, 166, 233, 287, 368, 373

— — термическая 410, 412, 413

— тепловых колебаний атомов кинетическая средняя 49, 52

— — — потенциальная средняя 50

— электронов в направленном потоке средняя 40, 80, 81, 293

— — тепловая (кинетическая) средняя 40, 242—243

Эттингсгаузена эффект 91, 92, 349—351

Ячейка (решетки кристаллической) элементарная 111, 152—155, 157, 160, 183

| | |
|---------------------------------------------------------------------------------------------|-----|
| Предисловие | 3 |
| Глава первая ОСНОВНЫЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ | 6 |
| 1. 1. Некоторые сведения о строении атома | 6 |
| 1. 2. Энергия и движение электрона в твердом теле | 10 |
| 1. 3. Электропроводность полупроводников | 36 |
| 1. 4. Теплопроводность полупроводников | 43 |
| 1. 5. Контактные явления | 55 |
| 1. 6. Термоэлектрические явления | 75 |
| 1. 7. Гальваномагнитные и термомагнитные явления | 83 |
| 1. 8. Фотопроводимость | 100 |
| Глава вторая СТРОЕНИЕ КРИСТАЛЛОВ | 113 |
| 2. 1. Некоторые вопросы квантовой теории | 113 |
| 2. 2. Геометрия кристаллической решетки | 147 |
| 2. 3. Дефекты в кристаллах | 163 |
| 2. 4. Тепловые колебания кристаллов | 174 |
| 2. 5. Теплоемкость | 184 |
| Глава третья ОСНОВНЫЕ ПРЕДСТАВЛЕНИЯ ЭЛЕКТРОННОЙ ТЕОРИИ КРИСТАЛЛОВ | 190 |
| 3. 1. Адиабатическое приближение | 190 |
| 3. 2. Одноэлектронное приближение | 194 |
| 3. 3. Приближение почти свободных электронов | 198 |
| 3. 4. Приближение сильно связанных электронов | 207 |
| 3. 5. Основные особенности структуры энергетических зон полупроводников | 209 |
| Глава четвертая СТАТИСТИКА ЭЛЕКТРОНОВ В ПОЛУПРОВОДНИКАХ | 217 |
| 4. 1. Некоторые понятия статистики и термодинамики | 217 |
| 4. 2. Распределение Ферми | 224 |
| 4. 3. Статистика невырожденного электронного газа в полупроводниках | 226 |
| 4. 4. Энергия электронов в зоне проводимости, вырождение | 235 |
| Глава пятая НЕКОТОРЫЕ ВОПРОСЫ ТЕОРИИ ЯВЛЕНИЙ ПЕРЕНОСА | 244 |
| 5. 1. Элементарный расчет электропроводности и подвижности | 245 |
| 5. 2. Кинетическое уравнение (учет энергетической зависимости времени релаксации) | 260 |

| | |
|------------------------------------------------------------|-----|
| 5. 3. Феноменологический анализ явлений переноса | 270 |
| 5. 4. Вычисление времени релаксации | 271 |
| 5. 5. Явления в сильных электрических полях | 278 |

Глава шестая
ТЕРМОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ И ТЕПЛОПРОВОДНОСТЬ 292

| | |
|-----------------------------------------------------------------------------------|-----|
| 6. 1. Термоэлектродвижущая сила | 294 |
| 6. 2. Вывод коэффициента термо- э.д.с. из кинетического уравнения | 296 |
| 6. 3. Увлечение электронов фононами | 299 |
| 6. 4. Зависимость термо- э.д.с. от температуры и концентрации носителей | 304 |
| 6. 5. Электронная теплопроводность | 311 |
| 6. 6. Теплопроводность кристаллической решетки | 317 |
| 6. 7. Фотонная теплопроводность | 329 |

Глава седьмая
ГАЛЬВАНОМАГНИТНЫЕ И ТЕРМОМАГНИТНЫЕ ЯВЛЕНИЯ 331

| | |
|-------------------------------------------------------------------------|-----|
| 7. 1. Общие сведения | 331 |
| 7. 2. Эффект Холла и изменение сопротивления в магнитном поле | 341 |
| 7. 3. Эффект Эттингсгаузена | 350 |
| 7. 4. Гальваномагнитные явления в сильных магнитных полях | 351 |
| 7. 5. Термомагнитные явления | 355 |

Глава восьмая
КОНТАКТНЫЕ ЯВЛЕНИЯ 362

| | |
|------------------------------------------------------------------------------------------------------------------------|-----|
| 8. 1. Особенности контактных явлений | 362 |
| 8. 2. Контакт полупроводника и металла | 366 |
| 8. 3. Диффузионная теория выпрямления Мотта (химический барьерный слой на границе металла с полупроводником) | 368 |
| 8. 4. Диодная теория Бете | 373 |
| 8. 5. Теория физического запорного слоя (теория истощения Шоттки) | 375 |
| 8. 6. Теория $p-n$ перехода | 378 |

Глава девятая
ОПТИЧЕСКИЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ 400

| | |
|-------------------------------------------|-----|
| 9. 1. Поглощение света | 400 |
| 9. 2. Фотопроводимость | 409 |
| 9. 3. Фотовольтаические эффекты | 421 |
| 9. 4. Циклотронный резонанс | 426 |
| 9. 5. Стимулированное излучение | 430 |

| | |
|--------------------------------|-----|
| Литература | 441 |
| Предметный указатель | 443 |