

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

**О.В. ТРЕТЯК  
В.З. ЛОЗОВСЬКИЙ**

## **ОСНОВИ ФІЗИКИ НАПІВПРОВІДНИКІВ**

**Підручник**

**Том 1**

*Затверджено  
Міністерством освіти і науки України  
як підручник  
для студентів вищих навчальних закладів*

УДК 537.311.322(075.8)

ББК 22.379.я73

T66

**Рецензенти:**

**чл.-кор. НАН України Ф.Ф.Сизов,  
д-р фіз.-мат наук, проф. В.А.Скришевський**

*Рекомендовано до друку вченою радою радіофізичного факультету  
(протокол № від 2006 року)*

**Третяк О.В., Лозовський В.З.**

**T66 Основи фізики напівпровідників: Підручник: У 2 т. – К.:  
Видавничо-поліграфічний центр "Київський університет",  
2007. – Т. 1. – 338 с.**

**ISBN 966-594-871-7 – Загальний**

**ISBN 966-594-872-5 – Т. 1**

Викладено основні ідеї фізики твердого тіла, на базі якої вивчається зонна структура напівпровідників; розглянуто статистику електронів та дірок, кінетичні явища та явища перенесення. Подано ефекти, що обумовлені наявністю домішок та дефектів у напівпровідниках. Розглянуто механізми розсіювання носіїв заряду, процеси дифузії та дрейфу. Обговорюються явища нестійкості у напівпровідниках.

Розраховано на студентів фізичних, радіофізичних і радіотехнічних спеціальностей вищих закладів освіти.

УДК 537.311.322(075.8)

ББК 22.379.я73

**Гриф надано Міністерством освіти і науки України  
(лист № від )**

ISBN 966-594-871-7 – Загальний  
ISBN 966-594-872-5 – Т. 1

© О.В.Третяк, В.З.Лозовський, 2007  
© Київський національний університет імені Тараса Шевченка,  
ВПЦ "Київський університет", 2007

## ЗМІСТ

<b>ВІД АВТОРІВ</b> .....	7
<b>■ Розділ 1. ЕЛЕМЕНТИ КРИСТАЛОГРАФІЇ</b> .....	9
1.1. Кристалічні ґратки.....	9
1.1.1. Ґратки Браве.....	9
1.1.2. Класифікація ґраток Браве та кристалічних структур.....	12
1.2. Пряма та обернена ґратки кристала.....	16
1.3. Електрони у періодичному потенціалі. Загальні властивості.....	22
1.4. Теорема Блоха.....	25
1.5. Циклічні граничні умови Борна–Кармана. Приведення до зони Бриллюена.....	27
1.6. Задачі.....	31
<i>Список літератури</i> .....	34
<b>■ Розділ 2. ПРИНЦИПИ КВАНТОВОЇ ТЕОРІЇ БАГАТЬОХ ЧАСТИНОК     У ФІЗИЦІ ТВЕРДОГО ТІЛА</b> .....	35
2.1. Адіабатичне наближення.....	35
2.2. Одноелектронне наближення.....	40
2.3. Квазіімпульс.....	44
2.4. Ізоенергетичні поверхні.....	47
2.5. Задачі.....	51
<i>Список літератури</i> .....	51
<b>■ Розділ 3. ЕЛЕКТРОН У ПЕРІОДИЧНОМУ ПОТЕНЦІАЛІ</b> .....	53
3.1. Рівняння Шрединґера для електрона у періодичному потенціалі.....	53
3.2. Електрон у полі потенціалу типу гребінця Дірака.....	54
3.3. Наближення майже вільного електрона.....	57
3.4. Утворення енергетичних зон.....	64
3.5. Електрон у полі потенціалу Кроніґа–Пенні.....	65
3.6. Утворення енергетичних зон у теорії квазізв'язаного електрона.....	69
3.7. Типи міжатомних зв'язків у кристалах.....	74
3.8. Метали, діелектрики, напівпровідники. Елементарна класифікація.....	80
3.9. Задачі.....	82
<i>Список літератури</i> .....	83
<b>■ Розділ 4. МЕТОДИ ОБЧИСЛЕННЯ ЗОННОЇ СТРУКТУРИ     КРИСТАЛІЧНИХ ТВЕРДИХ ТІЛ</b> .....	85
4.1. Метод плоских хвиль.....	88
4.2. Метод ортогоналізованих плоских хвиль.....	89
4.3. Метод приєднаних плоских хвиль.....	89
4.4. Метод псевдопотенціалу.....	90
4.5. k-p-метод.....	92
4.6. Метод ефективної маси.....	94
4.6.1. Розкладання енергії електрона поблизу екстремальних точок.....	94
4.6.2. Динамічний аспект проблеми ефективної маси.....	96
4.7. Спін-орбітальна взаємодія.....	98
4.8. Деякі приклади розрахунків зонної структури напівпровідників.....	102
4.9. Задачі.....	106
<i>Список літератури</i> .....	106

■ Розділ 5. ЕЛЕКТРОН У МАГНІТНОМУ ПОЛІ.....	107
5.1. Магнітний момент і спін електрона.....	107
5.2. Електронні статистичні системи.....	108
5.3. Рівні Ландау.....	111
5.4. Задачі.....	115
<i>Список літератури</i> .....	115
■ Розділ 6. КОЛИВАННЯ АТОМІВ КРИСТАЛІЧНОЇ ҐРАТКИ.....	117
6.1. Одновимірний лінійний ланцюжок атомів.....	117
6.2. Довгодіючі сили та метод оберненої ґратки.....	119
6.3. Коливання двоатомного лінійного ланцюжка.....	122
6.4. Коливання атомів кристала. Фонони.....	124
6.5. Статистика фононів.....	126
6.6. Електрон-фононна взаємодія.....	128
6.7. Задачі.....	132
<i>Список літератури</i> .....	133
■ Розділ 7. ДОМІШКИ ТА ДЕФЕКТИ В НАПІВПРОВІДНИКАХ.....	135
7.1. Типи дефектів у напівпровідниках.....	137
7.2. Енергетичні стани домішкових напівпровідників.....	139
7.3. Елементарна теорія дефектів.....	143
7.4. Задачі.....	146
<i>Список літератури</i> .....	147
■ Розділ 8. СТАТИСТИКА ЕЛЕКТРОНІВ У НАПІВПРОВІДНИКАХ В УМОВАХ ТЕРМОДИНАМІЧНОЇ РІВНОВАГИ.....	149
8.1. Статистика електронів і дірок.....	149
8.2. Електрони та дірки.....	152
8.3. Щільність станів.....	157
8.3.1. Сферичні ізоенергетичні поверхні.....	158
8.3.2. Еліпсоїдальні ізоенергетичні поверхні.....	159
8.3.3. Щільність станів для неквадратичного закону дисперсії.....	160
8.3.4. Щільність станів електронів напівпровідника в магнітному полі.....	162
8.3.5. Щільність станів електронів, що локалізовані на атомах домішок.....	165
8.4. Концентрація носіїв заряду. Вироджені та неvirоджені напівпровідники.....	166
8.5. Залежність положення рівня Фермі від температури.....	170
8.5.1. Рівняння електронейтральності.....	170
8.5.2. Концентрація електронів і дірок на простих домішкових центрах.....	170
8.5.3. Рівняння електронейтральності для власного напівпровідника.....	173
8.5.4. Напівпровідник, що має домішку одного типу.....	175
8.5.5. Напівпровідник, що містить акцепторну та донорну домішки.....	183
8.5.6. Вироджений напівпровідник.....	185
8.6. Задачі.....	187
<i>Список літератури</i> .....	190

<b>■ Розділ 9. ЯВИЩА ПЕРЕНЕСЕННЯ В НАПІВПРОВІДНИКАХ</b> .....	191
9.1. Кінетичне рівняння Больцмана.....	191
9.2. Час релаксації.....	194
9.3. Електропровідність напівпровідників.....	199
9.4. Залежність рухливості напівпровідника від температури.....	207
9.4.1. Розсіювання на коливаннях ґратки.....	207
9.4.2. Розсіювання на іонізованих домішках.....	207
9.4.3. Розсіювання на нейтральних атомах домішок.....	208
9.4.4. Розсіювання на дислокаціях.....	209
9.4.5. Залежність рухливості та електропровідності носіїв заряду в реальних напівпровідниках від температури.....	209
9.5. Гальваномагнітні явища в напівпровідниках.....	211
9.5.1. Ефект Холла.....	212
9.5.2. Ефект Холла в напівпровідниках із двома типами носіїв.....	215
9.5.3. Магніторезистивний ефект.....	222
9.5.4. Ефекти Еттінгсгаузена та Ернста.....	228
9.6. Термомагнітні ефекти.....	229
9.7. Термoeлектричні явища.....	233
9.8. Задачі.....	235
Список літератури.....	236
<b>■ Розділ 10. ГЕНЕРАЦІЯ ТА РЕКОМБІНАЦІЯ НОСІЇВ ЗАРЯДУ У НАПІВПРОВІДНИКАХ</b> .....	237
10.1. Розподіл нерівноважних носіїв заряду за енергіями.....	238
10.2. Механізми рекомбінації носіїв заряду.....	239
10.3. Кінетика генерації і рекомбінації носіїв заряду у напівпровідниках.....	243
10.3.1. Міжзонна випромінювальна рекомбінація.....	250
10.3.2. Міжзонна ударна рекомбінація.....	254
10.3.3. Рекомбінація через одностарядні локальні центри (за Шоклі-Рідом).....	257
10.3.4. Спін-залежна рекомбінація в напівпровідниках.....	265
10.4. Задачі.....	272
Список літератури.....	276
<b>■ Розділ 11. ДИФУЗІЯ ТА ДРЕЙФ НОСІЇВ ЗАРЯДУ</b> .....	277
11.1. Дифузійні та дрейфові струми.....	277
11.2. Нерівноважні напівпровідники. Квазірівні Фермі.....	279
11.2.1. Співвідношення Ейнштейна.....	281
11.2.2. Експериментальне визначення коефіцієнта дифузії за Хайнсом-Шоклі.....	283
11.3. Дифузія та дрейф нерівноважних носіїв заряду в монополярному напівпровіднику.....	286
11.4. Час релаксації Максвелла.....	288
11.5. Дифузія та дрейф неосновних надлишкових носіїв заряду в домішковому напівпровіднику.....	289
11.6. Дифузія неосновних носіїв заряду в електричному полі.....	292

11.7. Дрейф неосновних надлишкових носіїв заряду	295
11.8. Дифузія та дрейф носіїв заряду в напівпровідниках із провідністю, близькою до власної	297
11.9. Задачі	300
<i>Список літератури</i>	302
<b>■ Розділ 12. НЕСТІЙКОСТІ У НАПІВПРОВІДНИКАХ</b>	<b>303</b>
12.1. Електропровідність напівпровідників у сильних електричних полях	303
12.2. Електронна температура	308
12.3. Диференціальна провідність. <i>N</i> - та <i>S</i> -подібні вольт-амперні характеристики	311
12.4. Критерії виникнення нестійкості	315
12.5. Ударна іонізація	317
12.6. Ефект Ганна	318
12.7. Умови виникнення від'ємної диференціальної провідності	322
12.8. Задачі	324
<i>Список літератури</i>	326
<b>ДОДАТКИ</b>	<b>327</b>
Додаток А. Теорема Блоха для вироджених систем	327
Додаток В. Оператор координати в <i>k</i> -представленні	329
Додаток С. Функція розподілу за багатозарядними домішковими центрами	330
Додаток D. Час релаксації носіїв заряду за розсіювання на іонах домішок	332
Додаток Е. Енергія, яку втрачає електрон за розсіювання на оптичному фононі	336
Додаток F. Визначення відношення коефіцієнта дифузії до рухливості	337

# ВІД АВТОРІВ

Видання, яке Ви тримаєте в руках, шановний читачу, є підручником з фізики напівпровідників, що пропонується студентам, які обрали за спеціальністю фізику напівпровідників, електроніку (у широкому розумінні) та радіофізику. Воно може бути корисним для всіх, хто вивчає фізику твердого тіла або збирається студіювати сучасні нанотехнології.

Загальновідомо, що фізики, які зокрема вивчають її властивості напівпровідників, поділяються на дві великі групи. Одні, так звані теоретики, полюбляють пити каву, вести безперервні теревені, а між цими двома заняттями писати формули, знущатись над комп'ютером і жваво обговорювати результати, отримані їхніми колегами з іншого табору. Їхні ж колеги, яких називають експериментаторами, також полюбляють пити каву та теревенити – у вільний від роботи час. А такого часу у них не так і багато. Спробуйте знайти вільний час між походом у "кріогенку" за гелієм, ремонтом осцилографа та втовкмачуванням братам-теоретикам про неможливість забезпечення однакових умов у кожній квантовій точці, що геть не означає некоректності постановки експерименту, а свідчить про непридатність теоретичної моделі, яка має дуже мало спільного з реальною ситуацією. Розповідають, що одному із представників теоретиків експериментатор надав для пояснення результату останніх вимірювань. Прямо в коридорі теоретик почав пояснювати, що результати очевидні та є наслідком попередніх його робіт. Коли ж експериментатор розгледів, що колега тримає графік догори ногами й перевернув аркуша, то здивовано почув, що в такому випадку результат пояснюється ще простіше на основі нещодавно розробленої теоретиком моделі. Але попри це теоретики та експериментатори складають одну сім'ю дослідників і не можуть жити один без одного. І незважаючи на постійні "підколювання" ставляться один до одного з повагою та братерською любов'ю.

Цей підручник також написали експериментатор і теоретик, які, звісно, розглядають одні й ті самі явища із різних позицій. Автори дійшли консенсусу про існування кількох принципів питань, що потребують строгого викладання й формулювання у вигляді теорем. З іншого боку фізика і, зокрема фізика напівпровідників, є наукою експериментальною – це означає, що майже завжди явище відкривається в експериментальній роботі, і тільки згодом – знаходить теоретичне пояснення. Тому у підручнику вміщено детальні описи вирішальних експериментів, що сформували сучасні уявлення про фізику напівпровідників.

Підручник складається із двох томів. До першого – увійшли, так би мовити, класичні розділи фізики твердого тіла та фізики напівпровідників, за винятком оптичних властивостей і фізики поверхні. Таким чином, оскільки у першому томі в основному розглядатимуться кристалічні напівпровідники, почнемо подорож країною напівпровідників із розгляду деяких питань кристалографії. Далі детально розглянемо та обговоримо:

- основні ідеї фізики твердого тіла, зокрема з'ясуємо, чому при вивченні складних багаточастинкових систем, якими є напівпровідники, можна розглядати одноелектронну задачу та отримувати правильні результати;
- методи обчислення закону дисперсії електронів у напівпровіднику та викладемо основні ідеї за статистичного опису поведінки носіїв заряду у напівпровідниках;

- основні ідеї, які використовуються при вивченні дії магнітного поля на електронну систему та отримуємо ефект, відомий як "квантування Ландау";
- проблему неідеальності кристалічної ґратки за рахунок коливань атомів і введемо поняття фонону;
- ефекти, обумовлені наявністю легуючих домішок у напівпровідниках та явища електронного транспорту;
- наостанок першої частини – процеси генерації і рекомбінації носіїв заряду, дифузії та дрейф нерівноважних носіїв у напівпровідниках, явища в напівпровідниках у суттєво нерівноважних умовах і проблеми нестійкості.

У кінці кожного розділу подано задачі, велика кількість яких супроводжується детальними розв'язками. Тим самим продемонстровано підходи, які використовуються при розв'язанні невеличких проблем як необхідних кроків будь-якої наукової роботи. Крім того, подано чимало задач для самоконтролю за самостійного вивчення матеріалу. Для читачів, що воліють отримати детальнішу інформацію про явища, у кінці кожного розділу наведено список літератури.

Таким чином автори намагались поєднати традиційні підходи при вивченні фізики напівпровідників із новими сучасними ідеями, що правлять напівпровідниковою наукою та сучасними технологіями, вважаючи, що таке поєднання дасть більше користі, ніж шкоди та сподіваючись, що студенти відчують усю принадність цієї науки, яка без перебільшень стала основою сучасної цивілізації. Наука фізики напівпровідників стала суттю життя авторів, які мають надію, що молоде прийдешнє покоління знайде в ній власні шляхи та зробить свої відкриття.

Таким чином, на думку авторів, допитливий студент зможе знайти багато корисного й цікавого у цьому підручнику. Принаймні, автори сподівались на це, намагаючись докласти максимум зусиль для того, щоб вона вийшла цікавою.

Отже, уперед, шановний читачу. Бажаємо Вам успіхів у вивченні властивостей напівпровідників.



# Розділ 1

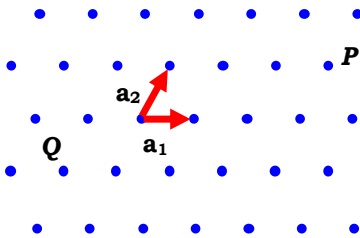
## ЕЛЕМЕНТИ КРИСТАЛОГРАФІЇ

### ■ 1.1. Кристалічні ґратки

#### 1.1.1. Ґратки Браве

Напівпровідники, що є основою сучасної мікроелектроніки, належать до великого класу твердих тіл, в яких атоми розташовуються впорядковано, або, іншими словами, утворюють кристалічну ґратку. Це означає, що завдяки розташуванню атомів у твердому тілі у вузлах періодичної ґратки електронні властивості великою мірою визначаються періодичним полем ґратки, яке діє на електронну підсистему. Таким чином, для адекватного опису електронних властивостей напівпровідників необхідно розглянути деякі загальні риси періодичних структур. Класифікацією можливих типів кристалічних структур і визначенням кристалічної структури реальних твердих тіл займається кристалографія, основною ідеєю якої є така:

**Ґратка Браве – нескінченна періодична структура, утворена дискретними точками, що має абсолютно однаковий просторовий порядок та орієнтацію незалежно від того, яку її точку обрано за вихідну.**



**Рис. 1.1.** Двовимірна ґратка Браве, точки якої можна отримати як лінійні комбінації основних векторів  $\mathbf{a}_1$  та  $\mathbf{a}_2$ . Наприклад,  $P = 3\mathbf{a}_1 + \mathbf{a}_2$ ,  $Q = -\mathbf{a}_1 - \mathbf{a}_2$

Приклад двовимірної ґратки Браве подано на рис. 1.1.

Тривимірна ґратка Браве утворена всіма точками із радіус-векторами

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \quad (1.1)$$

де  $\mathbf{a}_i$ ,  $i = 1, 2, 3$  – будь-які три некопланарні вектори,  $n_i$  – будь-які цілі числа, включаючи від'ємні та нуль. Про вектори  $\mathbf{a}_i$  кажуть, що вони породжують ґратку або є *основними векторами*. Суттєвим є те, що ґратка

Браве має не тільки абсолютно однаковий просторовий порядок, але й те, що орієнтації векторів тут лишаються незмінними незалежно від того, яка точка приймається за вихідну. Вибір основних векторів ґратки Браве не є однозначним. На рис. 1.2 показано трійку основних векторів для об'ємноцентрованої кубічної ґратки Браве (зазвичай для неї використовують коротке позначення ОЦК ґратка). Вузли ґратки розташовані у вершинах куба та на перетинах його просторових діагоналей. Для отримання такої ґратки необхідно взяти всі лінійні ком-

бінації основних векторів із цілими коефіцієнтами. Наприклад, для точки  $P$  маємо  $P = -\mathbf{a}_1 - \mathbf{a}_2 + 2\mathbf{a}_3$ . Рис. 1.3 ілюструє найбільш симетричний набір основних векторів для ОЦК ґратки (зеленим кольором позначено вузли, розташовані у центрі куба).

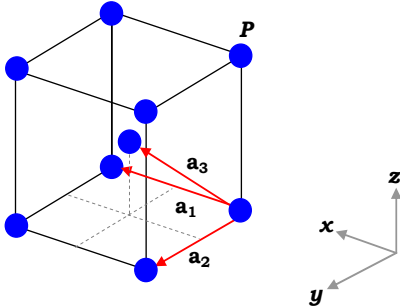


Рис. 1.2. Основні вектори для ОЦК ґратки

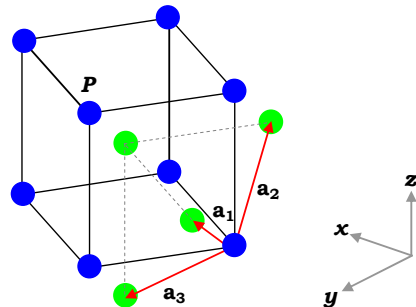


Рис. 1.3. Основні вектори для ОЦК ґратки (симетризований набір)

У цьому випадку для точки  $P$  ґратки можна записати  $P = 2\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3$ . На практиці набір основних векторів намагаються вибирати найбільш симетричним. Наприклад, за основні вектори вибирають:

♦ для простої кубічної ґратки

$$a_x, a_y, a_z; \tag{1.2}$$

♦ для об'ємноцентрованої ґратки (рис. 1.3)

$$\mathbf{a}_1 = \frac{a}{2}(\mathbf{y} + \mathbf{z} - \mathbf{x}), \quad \mathbf{a}_2 = \frac{a}{2}(\mathbf{z} + \mathbf{x} - \mathbf{y}), \quad \mathbf{a}_3 = \frac{a}{2}(\mathbf{x} + \mathbf{y} - \mathbf{z}); \tag{1.3}$$

♦ для гранецентрованої кубічної ґратки (ГЦК ґратки)

$$\mathbf{a}_1 = \frac{a}{2}(\mathbf{y} + \mathbf{z}), \quad \mathbf{a}_2 = \frac{a}{2}(\mathbf{z} + \mathbf{x}), \quad \mathbf{a}_3 = \frac{a}{2}(\mathbf{x} + \mathbf{y}). \tag{1.4}$$

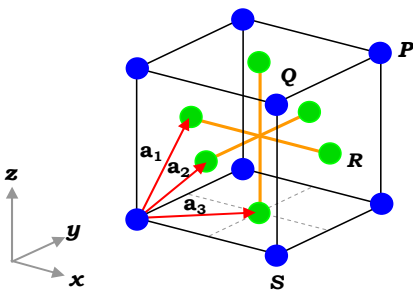


Рис. 1.4. Основні вектори для гранецентрованої ґратки Браве

Рис. 1.4 ілюструє трійку основних векторів (1.4) для гранецентрованої ґратки Браве, яка утворюється вузлами, що розташовані у вершинах куба (показано синім) та на перетині діагоналей усіх граней куба (показано зеленим). Нижче, як приклад, наведено спосіб отримання точок  $P, Q, R, S$  ґратки за допомогою цих векторів:  $P = \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3$ ,  $Q = 2\mathbf{a}_2$ ,  $R = \mathbf{a}_2 + \mathbf{a}_3$ ,  $S = -\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3$ .

Паралелепіпед, побудований на трьох векторах  $\mathbf{a}_i$ , називається *елементарною коміркою* ґратки. Якщо основні вектори  $\mathbf{a}_i$  вибрано так, що трансляція ґратки представляється як

$$\sum_i n_i \mathbf{a}_i$$

із цілими значеннями  $n_i$ , то елементарна комірка, побудована на векторах  $\mathbf{a}_i$ , називається *примітивною*. Якщо примітивна комірка містить один атом, то ґратка називається *простою*, якщо більше одного атома, – ґратка є *складною* або *ґраткою із базисом*. У загальному випадку елементарна комірка у формі паралелепіпеда не має симетрії кристалічної ґратки. Із нескінченної кількості варіантів вибору примітивної комірки бажано вибирати таку, яка б відповідала симетрії кристалічної ґратки. Із цим завданням можна впоратись за допомогою побудови *комірки Вігнера–Зейтца*. Візьмемо деякий атом ґратки  $O$  і проведемо з нього відрізки до найближчих атомів, а через середини відрізків – перпендикулярні до них площини. Перетинання площин утворює деякий багатогранник, що утримує всередині точку  $O$ . Такий багатогранник називається *коміркою Вігнера–Зейтца*. Подібними комірками можна щільно заповнити весь простір кристала. На рис. 1.5 подано комірку Вігнера–Зейтца для двовимірних ґраток Браве. Шість боків комірки розсікають навпіл відрізки прямих, що з'єднують центральну точку із шістьма сусідніми (лінії показано пунктиром). У двовимірному випадку комірка Вігнера–Зейтца будь-якої ґратки, крім прямокутної, завжди є шестикутником. У тривимірному випадку ситуація є складнішою.

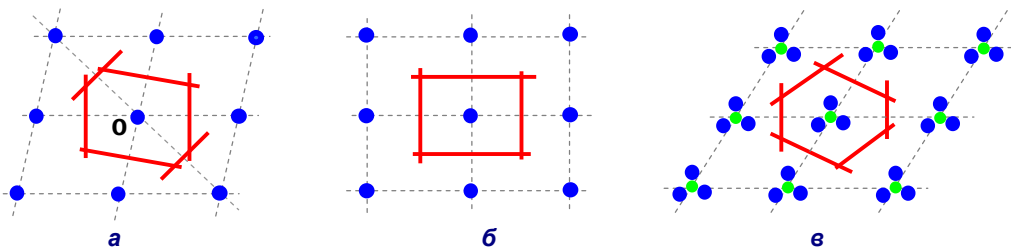


Рис. 1.5. Комірка Вігнера–Зейтца для двовимірних ґраток Браве: а, б – прості ґратки; в – складна ґратка (ґратка із базисом)

На рис. 1.6 подано комірку Вігнера–Зейтца для ОЦК ґратки. Шестикутні грані розсікають навпіл відрізки прямих, що з'єднують центральну точку із вершинами куба. Квадратні грані розсікають навпіл відрізки прямих, що з'єднують центральну точку із центральними точками кожної із шести сусідніх комірок.

На рис. 1.7 показано комірку Вігнера–Зейтца для ГЦК ґратки Браве. Кожна із 12 граней, що утворюють комірку, перпендикулярна прямій, яка з'єднує центральну точку куба із центром ребра та розсікає цей відрізок навпіл. Число найближчих атомів до даного, що знаходяться

від нього на одній відстані  $d$ , називається *координаційним числом*  $Z$ . У будь-якій простій ґратці це число є однаковим для всіх її вузлів. У простій кубічній ґратці координаційне число  $Z = 6$ , і відстань  $d$  є довжиною ребра куба  $a$ . В об'ємноцентрованій ґратці  $Z = 8$ , а  $d = \sqrt{3}a/2$ . У гранецентрованій кубічній ґратці  $Z = 12$ ,  $d = \sqrt{2}a/2$ .

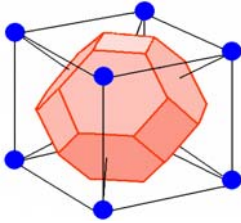


Рис. 1.6. Комірка Вігнера–Зейтца для ОЦК ґратки

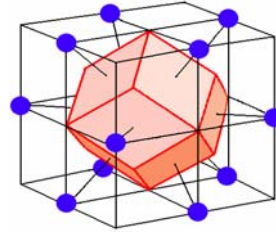


Рис. 1.7. Комірка Вігнера–Зейтца для ГЦК ґратки

### 1.1.2. Класифікація ґраток Браве та кристалічних структур

З погляду симетрії ґратка Браве задається всіма жорсткими операціями (за яких зберігаються відстані між всіма точками ґратки), що переводять ґратку саму у себе. Сукупність таких операцій утворює *групу симетрії* або *просторову групу* ґратки Браве. До таких операцій належать усі трансляції на вектори ґратки, повороти, відбивання та інверсії. Будь-яку операцію симетрії ґратки Браве можна побудувати із трансляції на вектор  $\mathbf{R}$  ґратки та жорсткої операції, що лишає нерухомою щонайменше одну точку ґратки. Продемонструємо це на прикладі простої кубічної ґратки, яка переходить сама у себе за повороту на кут  $90^\circ$  навколо вісі, що проходить через центр кубічної елементарної комірки (на рис. 1.8 позначено червоною точкою). Така операція не залишає жодної точки ґратки нерухомою (див. верхній рядок рисунку). З іншого боку, цей результат можна отримати завдяки складній операції трансляції на вектор ґратки Браве та повороту навколо вісі, на якій знаходиться точка 1 ґратки (див. нижній рядок рисунку), тобто з елемента точкової симетрії, який залишає щонайменше одну точку ґратки нерухомою. Таким чином, повна група симетрії ґратки Браве складається лише з:

- 1) трансляції на вектори ґратки Браве;
- 2) операцій, що лишають нерухомою деяку точку ґратки;
- 3) операцій, які можна побудувати з елементів 1 та 2.

При вивченні симетрії ґратки часто розглядають не всю просторову групу ґратки Браве, а тільки операції, що залишають нерухомою деяку її точку. Цю підмножину повної групи симетрії ґратки Браве називають *точковою групою ґратки Браве*. Існують сім різних точкових груп, симетрії яких можуть відповідати симетрії ґратки Браве. Це означає, що будь-яка кристалічна структура належить до однієї із семи кристалічних систем (сингоній) залежно від того, яка із точкових груп є групою її ґратки Браве. До таких кристалічних систем належать (у дужках позначено кількість типів ґраток даної сингонії, а на рисунках червоними точками – прямі кути):

- ♦ **кубічна система** (3) (рис. 1.9), що містить три ґратки Браве (проста, об'ємноцентрована, гранецентрована), точкова група симетрії яких збігається із групою симетрії куба. У простій кубічній ґратці атоми кристала розташовані у вершинах кубів, з яких складається кристал. Об'ємноцентрована кубічна ґратка, крім атомів у вершинах куба, має ще атом, розташований на перетині об'ємних діагоналей куба. У гранецентрованій ґратці атоми розташовуються у вершинах куба та на перетині діагоналей усіх граней куба;

- ♦ **тетрагональна система** (2). Для зниження симетрії куба можна взяти його за протилежні грані та витягнути у пряму призму із квадратною основою та висотою, що не дорівнює сторонам квадрата. Вісь уздовж цього напрямку називають *c*-віссю (рис. 1.10). Група симетрії такого об'єкта належить до тетрагональної групи. При цьому залежно від того, над якою кубічною ґраткою виконується просторова операція, можна отримати кристалічні структури тетрагональної групи різного типу. Наприклад, при розтягуванні простої кубічної ґратки можна отримати просту тетрагональну ґратку Браве, а при розтягуванні об'ємно- або гранецентрованої ґратки – лише одну ґратку тетрагональної системи – *центровану тетрагональну ґратку*.

Два способи представлення центрованої тетрагональної ґратки Браве (вигляд вздовж вісі *c*) подано на рис. 1.11. Сині точки позначають атоми, що лежать в атомній площині, перпендикулярній вісі *c*, а зелені – знаходяться у паралельній площині на відстані *c/2* від першої. Якщо сині точки з'єднати так, як показано на рис. 1.11 а, то вони утворять

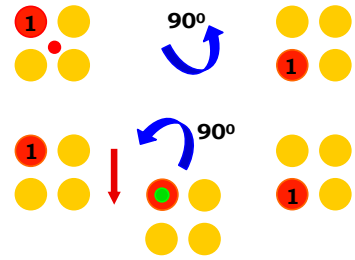


Рис. 1.8. Проста кубічна ґратка

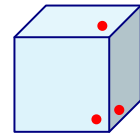


Рис. 1.9. Кубічна система

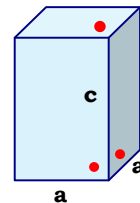


Рис. 1.10. Тетрагональна система

просту квадратну сітку. Видно, що центрована тетрагональна ґратка утворюється деформацією ОЦК ґратки. Якщо ж з'єднати точки ґратки так, щоб сині – лежали у вузлах центрованої квадратної ґратки (рис. 1.11 б), то зрозуміло, що центровану тетрагональну ґратку можна утворити також із ГЦК ґратки;

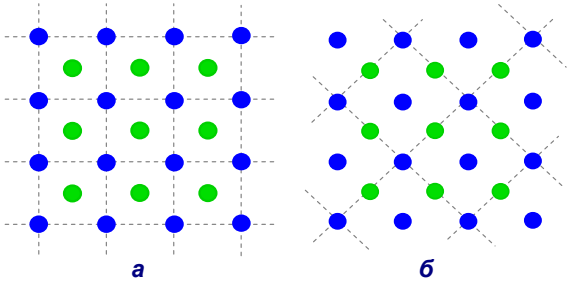


Рис. 1.11. Способи представлення центрованої тетрагональної ґратки Браве

♦ **ромбічна система** (4). Для зниження тетрагональної симетрії можна перетворити квадратні грані об'єкта на прямокутники тетрагональної системи (рис. 1.11). У результаті буде отримано об'єкт із трьома, взаємно перпендикулярними ребрами різної довжини (рис. 1.12). Група симетрії такого об'єкта називається ромбічною. За допомогою витягування простої тетрагональної ґратки вздовж однієї із  $a$ -вісей (див. перехід від рис. 1.13 а до рис. 1.13 б) можна отримати *просту ромбічну ґратку*. Якщо просту тетрагональну ґратку витягувати вздовж діагоналі квадрата основи (вздовж напрямку  $\alpha$ - $\alpha$ , рис. 1.13 а), то можна отримати іншу ґратку Браве із ромбічною точковою групою симетрії – *базоцентровану ромбічну ґратку* (див. перехід від рис. 1.13 а до рис. 1.13 в).

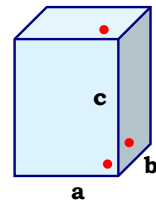


Рис. 1.12. Ромбічна система

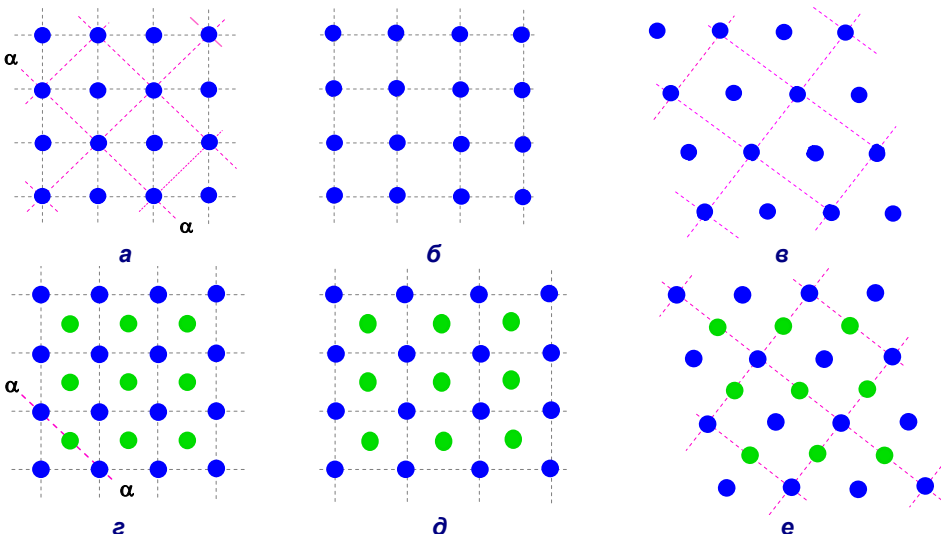


Рис. 1.13. Схема утворення ромбічної ґратки із тетрагональної

Таким самим шляхом у два способи можна знизити точкову симетрію центрованої тетрагональної ґратки та перетворити її на ромбічну. Якщо розтягувати її вздовж сторони квадрата (рис. 1.13 а), то можна отримати *об'ємноцентровану ромбічну ґратку* Браве (рис. 1.13 д), а якщо – вздовж діагоналі квадратів (або, що те ж саме, уздовж прямої  $a-a$  (рис. 1.13 в), – *гранецентровану ромбічну ґратку* (рис. 1.13 з);

♦ **моноклінна система** (2). Знизити симетрію ромбічної системи можна за допомогою перетворення прямокутників, перпендикулярних  $c$ -вісі (рис. 1.12), на довільні паралелограми. У результаті буде отримано об'єкт із моноклінною симетрією (рис. 1.14). Деформуючи таким чином просту ромбічну ґратку, можна отримати *просту моноклінну ґратку* Браве. Такі деформації щодо

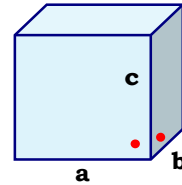


Рис. 1.14. Моноклінна система

базоцентрованої ромбічної ґратки Браве також дають просту моноклінну просторову групу. Але якщо таким чином деформувати гранецентровану або об'ємноцентровану ромбічні ґратки, то виникає центрована моноклінна ґратка Браве (рис. 1.15), в якій атоми розташовуються на вершинах паралелепіпеда (показано синім) і в точці перетину його об'ємних діагоналей (показано зеленим);

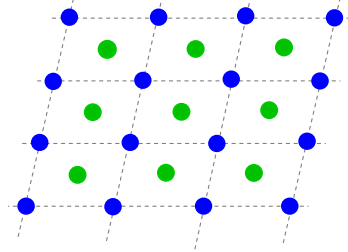


Рис. 1.15. Центрована моноклінна ґратка Браве. Вигляд уздовж  $c$ -вісі

♦ **триклінна система** (1). Якщо ж нахилити  $c$ -вісь щодо площини основи, то отримаємо фігуру, подану на рис. 1.16. Збурюючи таким чином будь-яку із моноклінних ґраток Браве, отримаємо триклінну ґратку Браве;

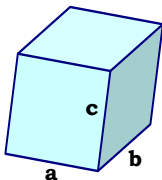


Рис. 1.16. Триклінна ґратка Браве

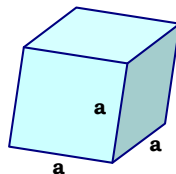


Рис. 1.17. Тригональна система

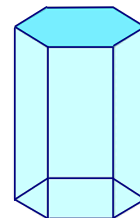


Рис. 1.18. Проста гексагональна ґратка Браве

♦ **тригональна система** (1). Тригональна точкова група описує симетрію об'єкта, що утворюється при розтягуванні куба вздовж об'ємної діагоналі (рис. 1.17);

♦ **гексагональна система** (1) *Проста гексагональна ґратка* Браве (із симетрією правильної шестикутної призми) є єдиною у гексагональній системі (рис. 1.18).

## ■ 1.2. Пряма та обернена ґратки кристала

Із періодичності розташування вузлів (атомів) у кристалі випливає, що за зміщення кристала у цілому на вектор

$$\mathbf{a}_n = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (1.5)$$

кристал збігається сам із собою. Це означає, що величини, які характеризують кристал (щільність електронів або електростатичний потенціал), мають бути просторово-періодичними функціями, тобто

$$V(\mathbf{r}) = V(\mathbf{r} + \mathbf{a}_n). \quad (1.6)$$

Побудуємо на векторах  $\mathbf{a}_i$  систему координат з осями  $\xi_1, \xi_2, \xi_3$ . У загальному випадку це є косокутною системою координат. Розкладемо періодичну функцію  $V(\mathbf{r})$  у цій системі координат у ряд Фур'є

$$V(\mathbf{r}) = \sum_{k_1, k_2, k_3} V_{k_1, k_2, k_3} e^{2\pi i \left( \frac{k_1 \xi_1}{a_1} + \frac{k_2 \xi_2}{a_2} + \frac{k_3 \xi_3}{a_3} \right)} \quad (1.7)$$

Для подальшого аналізу за допомогою матриці переходу  $\alpha_{ij}$  можна перейти від косокутної до прямокутної системи координат

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (1.8)$$

Такий перехід простіше продемонструвати на прикладі двовимірної ґратки (рис. 1.19). Запишемо координати точки  $M$ , що визначається вектором

$$\mathbf{OM} = \mathbf{OB} + \mathbf{BM}, \quad (1.9)$$

$$|\mathbf{OB}| = \xi_1, \quad |\mathbf{BM}| = \xi_2, \quad (1.10)$$

$$\mathbf{OM}_{x_1} = x_1 = \xi_1 \cos \varphi + \xi_2 \sin \varphi, \quad (1.11)$$

$$\mathbf{OM}_{x_2} = x_2 = \xi_1 \sin \varphi + \xi_2 \cos \varphi, \quad (1.12)$$

звідки при розв'язанні системи рівнянь (1.11)–(1.12) отримаємо:

$$\xi_1 = \alpha_{11} x_1 + \alpha_{12} x_2, \quad (1.13)$$

$$\xi_2 = \alpha_{21} x_1 + \alpha_{22} x_2, \quad (1.14)$$

де позначено

$$\alpha_{11} = \frac{\cos \varphi}{\cos(\varphi + \varphi)}, \quad \alpha_{12} = \frac{\sin \varphi}{\cos(\varphi + \varphi)}, \quad (1.15)$$

$$\alpha_{21} = -\frac{\sin \varphi}{\cos(\varphi + \varphi)}, \quad \alpha_{22} = \frac{-\cos \varphi}{\cos(\varphi + \varphi)}. \quad (1.16)$$



Таким чином, будь-яке розкладення типу (1.7), записане у довільній системі координат, можна записати у прямокутній системі координат

$$V(\mathbf{r}) = \sum_{b_1 b_2 b_3} V_{b_1 b_2 b_3} e^{i(b_1 x_1 + b_2 x_2 + b_3 x_3)}, \quad (1.17)$$

де  $b_i = \sum_{j=1}^3 k_j \alpha_{ji} / \alpha_j$ . Якщо розглядати числа  $b_i$

як компоненти деякого вектора  $\mathbf{b}$ , то розкладення (1.17) можна записати у вигляді

$$V(\mathbf{r}) = \sum_{\mathbf{b}} V_{\mathbf{b}} e^{i(\mathbf{b}\mathbf{r})}. \quad (1.18)$$

Оскільки функція  $V(\mathbf{r})$  має бути періодичною на ґратці, запишемо

$$V(\mathbf{r}) = V(\mathbf{r} + \mathbf{a}_n) = \sum_{\mathbf{b}} V_{\mathbf{b}} e^{i\mathbf{b}(\mathbf{r} + \mathbf{a}_n)} = \sum_{\mathbf{b}} V_{\mathbf{b}} e^{i\mathbf{b}\mathbf{r}} e^{i\mathbf{b}\mathbf{a}_n}. \quad (1.19)$$

Звідси випливає, що експонента  $e^{i\mathbf{b}\mathbf{a}_n}$  має дорівнювати одиниці. Це означає, що (1.19) вимагає виконання умови

$$\mathbf{b}\mathbf{a}_n = n_1 \mathbf{b}\mathbf{a}_1 + n_2 \mathbf{b}\mathbf{a}_2 + n_3 \mathbf{b}\mathbf{a}_3 = 2\pi N, \quad (1.20)$$

де  $N$  – довільне ціле число. Оскільки (1.20) має виконуватись для будь-яких цілих чисел  $n_i$ , то лишається єдина можливість задовольнити це рівняння, а саме покласти

$$\mathbf{b}\mathbf{a}_1 = 2\pi g_1, \quad \mathbf{b}\mathbf{a}_2 = 2\pi g_2, \quad \mathbf{b}\mathbf{a}_3 = 2\pi g_3. \quad (1.21)$$

Зрозуміло, що можна запропонувати безліч шляхів забезпечення співвідношень (1.21), але необхідно визначити єдиний шлях. При цьому треба зважати, що вектор  $\mathbf{b}$  має визначатись параметрами ґратки. Тому для однозначного визначення  $\mathbf{b}$  розкладемо його за трьома некопланарними векторами, що побудовані як векторні добутки основних векторів ґратки

$$\mathbf{b} = \zeta_1 (\mathbf{a}_2 \times \mathbf{a}_3) + \zeta_2 (\mathbf{a}_3 \times \mathbf{a}_1) + \zeta_3 (\mathbf{a}_1 \times \mathbf{a}_2). \quad (1.22)$$

Підставляючи це розкладення до (1.21), маємо

$$2\pi g_1 = \zeta_1 \mathbf{a}_1 (\mathbf{a}_2 \times \mathbf{a}_3), \quad 2\pi g_2 = \zeta_2 \mathbf{a}_2 (\mathbf{a}_3 \times \mathbf{a}_1), \quad 2\pi g_3 = \zeta_3 \mathbf{a}_3 (\mathbf{a}_1 \times \mathbf{a}_2), \quad (1.23)$$

звідси  $\zeta_i = \frac{2\pi}{\Omega_0} g_i$ , де  $\Omega_0 = \mathbf{a}_1 (\mathbf{a}_2 \times \mathbf{a}_3)$  – об'єм елементарної комірки. Отже

$$\mathbf{b} = g_1 \mathbf{b}_1 + g_2 \mathbf{b}_2 + g_3 \mathbf{b}_3, \quad (1.24)$$

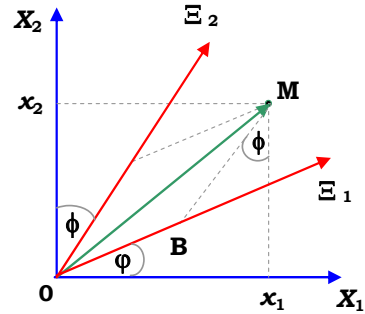


Рис. 1.19. Перетворення системи координат

де 
$$\mathbf{b}_1 = \frac{2\pi(\mathbf{a}_2 \times \mathbf{a}_3)}{\Omega_0}, \mathbf{b}_2 = \frac{2\pi(\mathbf{a}_3 \times \mathbf{a}_1)}{\Omega_0}, \mathbf{b}_3 = \frac{2\pi(\mathbf{a}_1 \times \mathbf{a}_2)}{\Omega_0}. \quad (1.25)$$

Ці вектори називаються векторами *оберненої ґратки*, і, як безпосередньо випливає із визначення, мають розмірність оберненої довжини. Вектори  $\mathbf{b}$  із (1.24) утворюють ґратку з основною коміркою, що побудована на векторах  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ . Така ґратка називається *оберненою до вихідної прямої ґратки* та є інваріантним геометричним об'єктом, властивості якого відіграють визначальну роль у фізиці твердого тіла.

Розглянемо *властивості оберненої ґратки*.

1. Оскільки векторний добуток є вектором, що перпендикулярний до векторів, які його утворюють, маємо

$$\mathbf{a}_i \cdot \mathbf{b}_k = 2\pi\delta_{ik}, \quad (1.26)$$

тобто

$$\mathbf{b}_1 \perp \mathbf{a}_2 \text{ та } \mathbf{a}_3,$$

$$\mathbf{b}_2 \perp \mathbf{a}_1 \text{ та } \mathbf{a}_3,$$

$$\mathbf{b}_3 \perp \mathbf{a}_2 \text{ та } \mathbf{a}_1.$$

2. Визначимо площину прямої ґратки Браве як таку, що проходить щонайменше через будь-які три вузли ґратки Браве, не розташовані на одній прямій. Тоді кожен вектор оберненої ґратки є перпендикулярним до деякої множини площин прямої ґратки. Дійсно, виберемо будь-який

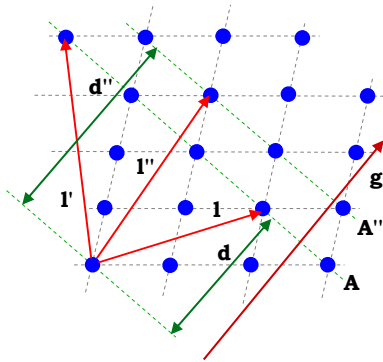


Рис. 1.20. Кристалічні площини та вектори оберненої ґратки

вектор оберненої ґратки  $\mathbf{g}$  і вектор прямої ґратки  $\mathbf{l}$  та утворимо їхній скалярний добуток. Згідно із (1.20) маємо

$$\mathbf{g} \cdot \mathbf{l} = 2\pi N. \quad (1.27)$$

Це означає, що довжина проекції вектора  $\mathbf{l}$  на напрямок вектора  $\mathbf{g}$

$$d = \frac{2\pi N}{|\mathbf{g}|}. \quad (1.28)$$

Така властивість притаманна нескінченній кількості точок прямої ґратки. Дійсно, припустимо, що вектор ґратки  $\mathbf{l}'$  (рис. 1.20) визначається цілими числами

$$l'_1 = l_1 - mn_3, \quad l'_2 = l_2 - mn_3, \quad l'_3 = l_3 + m(n_1 + n_2), \quad (1.29)$$

де  $m$  – ціле. Маємо

$$\mathbf{g} = n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2 + n_3 \mathbf{b}_3, \quad \mathbf{l}' = (l_1 - mn_3) \mathbf{a}_1 + (l_2 - mn_3) \mathbf{a}_2 + (l_3 + m(n_1 + n_2)) \mathbf{a}_3,$$

$$\mathbf{g} \cdot \mathbf{l}' = \{n_1(l_1 - mn_1) + n_2(l_2 - mn_3) + n_3(l_3 + m(n_1 + n_2))\}.$$

Використовуючи (1.27), отримаємо

$$\mathbf{g} \cdot \mathbf{l} = \mathbf{g} \cdot \mathbf{l}' = 2\pi N, \quad (1.30)$$

отже вектор  $\mathbf{l}'$  має ту саму проекцію на напрямок  $\mathbf{g}$ , що й  $\mathbf{l}$ . Таким чином, вектор  $\mathbf{l}$  визначає точку, що лежить у площині, нормальній до  $\mathbf{g}$ , і відстоїть від початку координат на відстані  $d$ . Іншими словами, якщо цій площині належить хоча б одна точка ґратки, то на ній лежить нескінченна множина таких точок: ми побудували одну із площин ґратки.

3. Якщо компоненти вектора  $\mathbf{g}$  не мають загального множника, то його довжина обернено пропорційна до відстані між сусідніми площинами ґратки, що перпендикулярні вектору  $\mathbf{g}$ . Якщо числа  $(n_1, n_2, n_3)$  не мають загального множника, то завжди можна знайти вектор  $\mathbf{l}''$  такий, що

$$\mathbf{g} \cdot \mathbf{l}'' = 2\pi(N+1). \quad (1.31)$$

Звідси робимо висновок, що площина ґратки, якій належить вектор  $\mathbf{l}''$ , розташована на відстані

$$d'' = \frac{2\pi(N+1)}{|\mathbf{g}|} \quad (1.32)$$

від початку координат. Тобто, відстань між площинами  $\{\mathbf{l}''\}$  та  $\{\mathbf{l}\}$

$$d'' - d = \frac{2\pi(N+1)}{|\mathbf{g}|} - \frac{2\pi N}{|\mathbf{g}|} = \frac{2\pi}{|\mathbf{g}|}. \quad (1.33)$$

Для побудови оберненої ґратки можна використати алгебраїчний метод і, застосовуючи (1.25), обчислити вектори  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ , де як на базисі побудувати вектор  $\mathbf{b}$ , згідно із (1.24). Вектори  $\mathbf{b}$  утворюють питому обернену ґратку. Розглянемо тепер важливі приклади побудови оберненої ґратки.

Прямі обчислення із використанням (1.25) показують, що оберненою ґраткою до простої кубічної ґратки зі стороною елементарної комірки  $a$  є проста кубічна ґратка із кубічною елементарною коміркою зі стороною  $2\pi/a$ . Дійсно, якщо базисом кубічної ґратки є  $\mathbf{a}_1 = a\mathbf{x}$ ,  $\mathbf{a}_2 = a\mathbf{y}$  та  $\mathbf{a}_3 = a\mathbf{z}$ , то, згідно із (1.25),  $\mathbf{b}_1 = (2\pi/a)\mathbf{x}$ ,  $\mathbf{b}_2 = (2\pi/a)\mathbf{y}$ , і  $\mathbf{b}_3 = (2\pi/a)\mathbf{z}$ .

Застосовуючи побудову (1.25) до базису гранецентрованої кубічної ґратки зі стороною  $a$  (див. (1.4)) і зважаючи на те, що  $(\mathbf{z} + \mathbf{x}) \times (\mathbf{x} + \mathbf{y}) = \mathbf{y} + \mathbf{z} - \mathbf{x}$ ,  $(\mathbf{x} + \mathbf{y}) \times (\mathbf{y} + \mathbf{z}) = \mathbf{z} + \mathbf{x} - \mathbf{y}$  та  $(\mathbf{y} + \mathbf{z}) \times (\mathbf{z} + \mathbf{x}) = \mathbf{x} + \mathbf{y} - \mathbf{z}$ , отримаємо базисні вектори оберненої ґратки

$$\mathbf{b}_1 = \frac{(4\pi/a)}{2}(\mathbf{y} + \mathbf{z} - \mathbf{x}), \quad \mathbf{b}_2 = \frac{(4\pi/a)}{2}(\mathbf{z} + \mathbf{x} - \mathbf{y}), \quad \mathbf{b}_3 = \frac{(4\pi/a)}{2}(\mathbf{x} + \mathbf{y} - \mathbf{z}),$$

що збігаються за формою із базисом об'ємноцентрованої кубічної ґратки, де стороною умовної кубічної ґратки є  $4\pi/a$ .

Якщо аналогічно виконати побудову, згідно із (1.25), для об'ємноцентрованої кубічної ґратки із базисом (1.3), то отримаємо

$$\mathbf{b}_1 = \frac{(4\pi/a)}{2}(\mathbf{y} + \mathbf{z}), \quad \mathbf{b}_2 = \frac{(4\pi/a)}{2}(\mathbf{z} + \mathbf{x}), \quad \mathbf{b}_3 = \frac{(4\pi/a)}{2}(\mathbf{y} + \mathbf{x}),$$

що за формою збігається із набором основних векторів гранецентрованої кубічної ґратки (див. (1.4)). Тобто оберненою ґраткою для об'ємноцентрованої кубічної ґратки буде гранецентрована кубічна ґратка зі стороною  $4\pi/a$ .

Зауважимо, що часто побудову оберненої ґратки для заданої ґратки Браве варто виконувати із застосуванням геометричних міркувань, що базуються на властивостях оберненої ґратки. Наприклад, для одновимірної ґратки Браве із базисом  $\mathbf{a} = a\mathbf{x}$  можна використати властивість (1.26). У результаті отримаємо, що оберненою ґраткою буде одновимірна ґратка із базисом  $\mathbf{b} = (2\pi/a)\mathbf{x}$ .

Як впливає із властивостей оберненої ґратки, найпростіший спосіб характеристизації площин ґратки Браве полягає у задаванні нормалей до цих площин. *Атомна площина* кристалічної ґратки визначається як площина відповідної ґратки Браве. Для кожного набору атомних площин, що розташовані одна від однієї на відстані  $d$ , існують вектори оберненої ґратки, перпендикулярні до цих площин. Як впливає із (1.33), найменший із них має довжину  $2\pi/d$ . І навпаки, для кожного вектора оберненої ґратки  $\mathbf{G}$  існує набір атомних площин, що перпендикулярні до  $\mathbf{G}$  і розташовані одна від одної на відстані  $d$ . При цьому  $\mathbf{g} = 2\pi/d$  та є найменшим вектором оберненої ґратки, що паралельний до  $\mathbf{G}$ . Таким чином, орієнтація атомної площини однозначно задається вектором нормалі до неї, оскільки для будь-якого набору атомних площин існує вектор оберненої ґратки, нормальний до них. Для визначеності вибирають найменший із таких векторів. Якщо вектор має вигляд

$$\mathbf{K}_0 = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3, \quad (1.34)$$

то числа  $h, k, l$  називають *індексами Міллера* даної площини. Оскільки площина ґратки з індексами Міллера є перпендикулярною до вектора оберненої ґратки (1.34), то можна визначити деяку сталу  $A$  таку, що ця площина визначатиметься рівнянням

$$\mathbf{K}_0 \cdot \mathbf{r} = A. \quad (1.35)$$

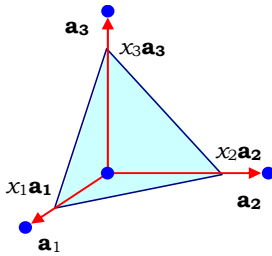


Рис. 1.21. Визначення індексів Міллера

Площина пересікає вісі, направлені за основними векторами  $\mathbf{a}_i$  прямої ґратки у точках  $x_1\mathbf{a}_1$ ,  $x_2\mathbf{a}_2$  та  $x_3\mathbf{a}_3$  (рис. 1.21). Оскільки площина задовольняє (1.35), то цьому рівнянню мають задовольняти й величини  $x_i$ , тобто виконуватись умови  $\mathbf{K}_0 \cdot (x_i\mathbf{a}_i) = A$ . З іншого боку  $\mathbf{K}_0 \cdot \mathbf{a}_1 = 2\pi h$ ,  $\mathbf{K}_0 \cdot \mathbf{a}_2 = 2\pi k$ ,  $\mathbf{K}_0 \cdot \mathbf{a}_3 = 2\pi l$ . Тоді

$$x_1 = \frac{A}{2\pi h}, \quad x_2 = \frac{A}{2\pi k}, \quad x_3 = \frac{A}{2\pi l}, \quad (1.36)$$

тобто відрізки, які відсікає на осях кристала атомна площина, обернено пропорційні до індексів Міллера цієї площини. На основі цього факту у кристалографії індекси Міллера визначають ще як цілі числа, які не мають загального множника та обернено пропорційні відріzkам, що відсікає атомна площина на осях кристала

$$h:k:l = (1/x_1):(1/x_2):(1/x_3). \quad (1.37)$$

Індекси Міллера широко застосовують для визначення орієнтацій у кристалі. І оскільки властивості кристалів сильно залежать від напрямків щодо кристалографічних осей, прийнято загальну схему позначень для орієнтацій кристалічних площин, напрямків і цілих їхніх сімейств. Наприклад, атомні площини зазвичай позначаються індексами Міллера у круглих дужках. На рис. 1.22 подано приклади трьох атомних площин для простої кубічної ґратки та їхні індекси Міллера. Коми між цифрами опускають, і знак мінус (якщо такий є) записують над цифрою. Наприклад, у кубічній системі площину, що відсікає на осях  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$  відрізки  $(2, -4, 1)$ , позначають як  $(2\bar{4}1)$ . Зрозуміло, що для однозначної інтерпретації таких символів необхідно знати, як вибрані осі системи координат, що використовуються. Для нас важливими є кристали кубічної симетрії, оскільки більшість напівпровідників характеризуються кубічною кристалічною структурою. У такому випадку завжди використовують осі простої кубічної комірки.

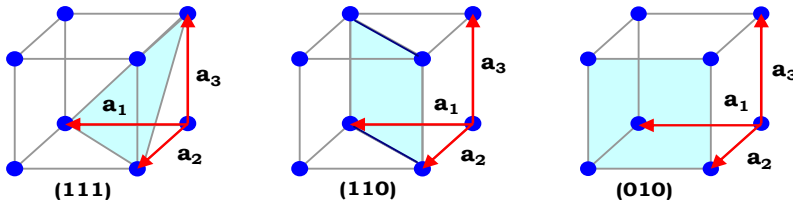


Рис. 1.22. Індекси Міллера атомних площини простої кубічної ґратки

Для напрямків у прямій ґратці використовують аналогічні позначення, але для відмінності від індексів Міллера застосовують квадратні дужки. Оскільки деякий радіус-вектор

$$\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (1.38)$$

відсікає на осях системи координат відрізки  $n_1 \mathbf{a}_1$ ,  $n_2 \mathbf{a}_2$  та  $n_3 \mathbf{a}_3$ , то його напрямок позначається як  $[n_1 n_2 n_3]$ , наприклад просторова діагональ простої кубічної ґратки має напрямок  $[111]$ . Часто використовують позначення, що вказує не тільки на одне сімейство атомних площин, а й на всі інші сімейства, еквівалентні йому через симетрію кристала. Якщо площина  $(hkl)$  еквівалентна цілому набору площин, то весь набір позначають символом  $\{hkl\}$ . Наприклад, у кубічному кристалі площини  $(100)$ ,  $(010)$  та  $(001)$  еквівалентні. Для таких сімейств площин використовують загальне позначення  $\{100\}$ . Подібне правило застосовують і для напрямків. Наприклад, еквівалентні напрямки у кубічному кристалі  $[100]$ ,  $[010]$ ,  $[001]$ ,  $[\bar{1}00]$ ,  $[0\bar{1}0]$  та  $[00\bar{1}]$  позначають символом  $\langle 100 \rangle$ .

Насамкінець зауважимо, що аналогічно до прямого простору елементарна комірка оберненої ґратки не обов'язково є паралелепіпедом. У теорії твердого тіла майже завжди працюють із коміркою Вігнера–Зейтца в оберненій ґратці. Така комірка має спеціальну назву – зона Бриллюена. Такі напівпровідники, як германій і кремній кристалізуються у структуру із ґраткою типу алмазу (рис. 1.23), а арсенід галію – у структуру типу цинкової обманки, що мають симетрію кубічної гранецентрованої ґратки, тому на рис. 1.24 показано зону ГЦК кристалічної структури із позначенням точок і ліній симетрії. Зона Бриллюена обмежується вісьмома правильними шестикутниками, що належать до сімейства площин  $\{111\}$ , і шістьма квадратами, що належать до сімейства площин  $\{002\}$ . Точка  $\Gamma$  лежить у центрі зони Бриллюена. Точками  $W$  позначають вершини, центри шестикутних граней – через  $L$ , а центри квадратних граней – через  $X$ .

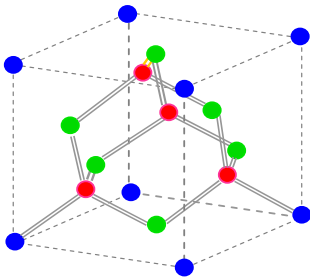


Рис. 1.23. Кристалічна структура типу алмазу

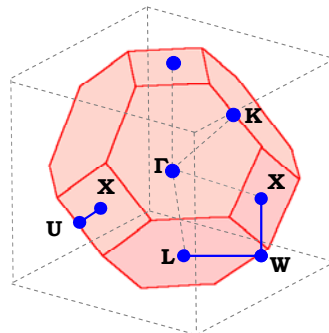
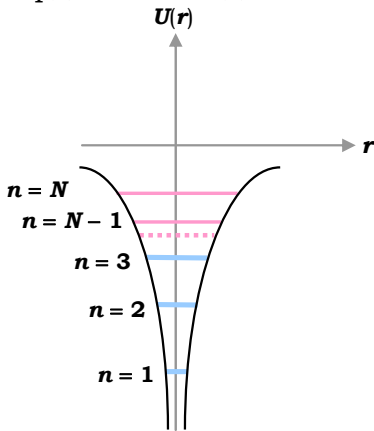


Рис. 1.24. Лінії та точки симетрії у зоні Бриллюена ГЦК структури

### ■ 1.3. Електрони у періодичному потенціалі. Загальні властивості

Тверде тіло іноді називають конденсованим станом речовини, і уточнюють – кристалічним конденсованим станом. Це означає, що при утворенні твердого тіла атоми речовини конденсуються у кристалічну тверду фазу, тобто утворюють за такої конденсації регулярну структуру. Існують й інші можливості конденсації речовини у тверде тіло. Наприклад, якщо за конденсації не виникає далекого порядку, а істотним є ближній порядок, то йдеться про утворення аморфного твердого тіла. Одним із загальновідомих прикладів аморфних твердих



**Рис. 1.25. Енергетичний спектр електронів в ізолюваному  $N$ -електронному атомі**

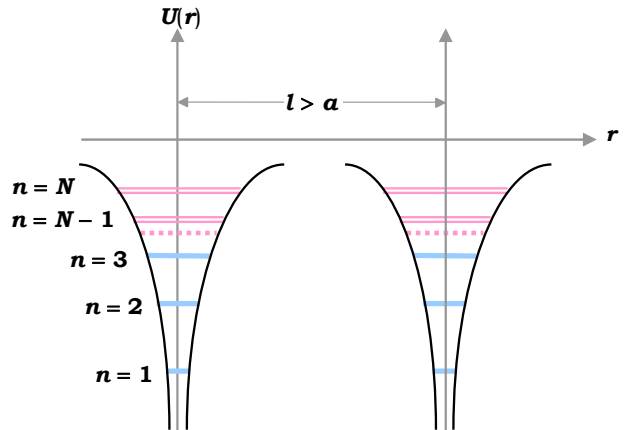
тіл є звичайне скло. Про аморфні напівпровідники йтиметься пізніше, а тепер розглянемо на якісному рівні процес конденсації речовини у кристалічну твердотільну фазу. Наприклад, є окремий атом деякої речовини. Зважаючи на те, що енергетичний спектр електронів в ізолюваному атомі є дискретним, уявимо енергетичну діаграму ізолюваного атома у вигляді, що поданий на рис. 1.25. Вважатимемо, що атом має  $N$  електронів, з яких – два валентних. Тепер уявимо, що два таких однакових атоми конденсуються у молекулу. Зрозуміло, що за рахунок взаємодії їхні електронні стани змінюватимуться, тобто зміниться й структура енергетичних рівнів.

Згідно із принципом Паулі електрони, що заповнюють внутрішні оболонки завдяки ефекту екранування, відчуватимуть дію зовнішнього атома набагато слабкіше за електрони зовнішніх оболонок, тобто валентних електронів. Що трапиться, коли ці два атоми почнуть наближатись один до одного? Оскільки вони однакові, то зрозуміло, що взаємодія відбуватиметься між відповідними виродженими електронними станами. З іншого боку, із курсу квантової механіки відомо, що взаємодія викликає зникнення виродження, тобто можна передбачити,

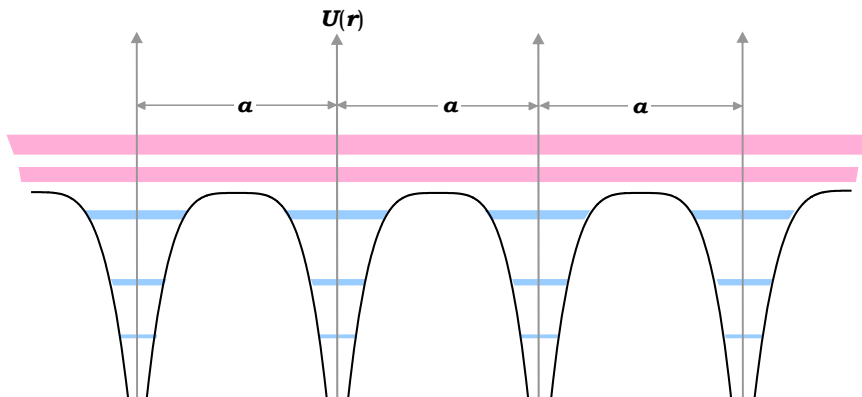
що завдяки взаємодії вироджені електронні стани валентних електронів помітно розщеплюватимуться, стани же внутрішніх електронів визначатимуться незначним розщеплюванням, що ефективно спостерігатиметься як більше або менше розширення рівнів (рис. 1.26).

Тепер уявимо, що у процесі конденсації бере участь дуже велика кількість атомів. Зменшення відстані між ними до розмірів сталої ґратки викличе зниження

потенціальної енергії, і взаємодія між атомами у ґратці приведе до розмиття енергетичних рівнів валентних атомів у майже суцільні смуги, що мають назву енергетичних зон (рис. 1.27). Стани у зонах заповнюватимуться електронами, згідно із принципом Паулі (на одному енергетичному рівні може бути лише один електрон або два – із протилежними спінами). Таким чином, якщо у кристалічне тверде тіло конденсувались атоми з одним валентним електроном, то енергетична зона, в яку розмісться рівень валентних електронів, буде заповнений рівно наполовину. Як результат маємо випадок типового металу.



**Рис. 1.26.** Енергетичний спектр електронів у двох атомах, що розташовані один від одного на відстані, більшій за критичну для утворення хімічного зв'язку



**Рис. 1.27.** Схема виникнення енергетичних зон за конденсації атомів у кристал.

Завдяки взаємодії між атомами потенціальна енергія дещо знижується та утворюються зони колективізованих електронів. Рівні локалізованих електронів також розмиваються в смуги



Якщо у кристалічне тверде тіло конденсувались атоми із двома валентними електронами, то енергетична зона, утворена розмиттям рівня валентних електронів, буде повністю заповнена, і для збудження будь-якого з них потрібно витратити деяку енергію, що дорівнює енергетичній відстані між заповненою зоною і пустою зоною, яка утворилась із першого збудженого рівня валентного електрона в ізолюваному атомі.

Таким чином, із якісного аналізу енергетичної структури електронів у твердих тілах ми доходимо дуже важливого висновку: енергетичний спектр електронів у твердих тілах має зонний характер, тобто у результаті взаємодії атомів, що складають тверде тіло, електронні енергетичні стани утворюють цілі енергетичні смуги – енергетичні зони, розділені смугами недозволенних станів – заборонені зони.

Насправді ж ситуація є набагато складнішою. За допомогою простих наочних моделей неможливо достатньо точно описати енергетичний спектр електронів у кристалах. Для детальнішого вивчення електронних властивостей кристалічних твердих тіл необхідно застосовувати методи квантової теорії. А оскільки кристал являє собою складну систему з атомів ґратки і великої кількості колективізованих електронів, варто використовувати досить складні підходи, що базуються на методах квантової теорії багатьох частинок. Як перший крок до побудови таких підходів сформулюємо надзвичайно важливу теорему, що дозволяє визначати загальний вигляд хвильових функцій частинки, що рухається у періодичному потенціалі.

## ■ 1.4. Теорема Блоха

Для аналізу властивостей (електронних, магнітних, пружних тощо) твердого тіла недостатньо знати функції із симетрією ґратки (напр., кристалічний потенціал або електронні хвильові функції основного стану). Необхідно розглядати й збудження, що порушують точну симетрію структури, яка розглядається. Збудження системи описуються деякими рівняннями, що у будь-якому разі мають бути інваріантними щодо трансляцій ґратки. Розглянемо такий випадок на прикладі електронних збуджень. Нехай  $V(\mathbf{r})$  – потенціальна енергія електрона у точці  $\mathbf{r}$ . Тоді у силу трансляційної симетрії системи для будь-якого вектора ґратки  $\mathbf{l}$  маємо

$$V(\mathbf{r} + \mathbf{l}) = V(\mathbf{r}). \quad (1.39)$$

Хвильова функція електрона  $\Psi(\mathbf{r})$  задовольняє рівнянню Шредингера

$$\left( -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) - E \right) \psi = 0, \quad (1.40)$$

яке має лишатись незмінним за заміни  $\mathbf{r}$  на  $\mathbf{r} + \mathbf{l}$  у всіх операторах, що діють на  $\Psi(\mathbf{r})$ . Для строго розгляду цієї обставини позначимо через  $H(0)$ ,  $|0\rangle$ ,  $H(\mathbf{l})$  та  $|\mathbf{l}\rangle$  гамільтоніан і хвильову функцію до й після трансляції, відповідно. Трансляційна інваріантність вимагає виконання умови

$$H(\mathbf{l}) = H(0). \quad (1.41)$$

Тепер задача на власні значення запишеться як

$$H(0)|0\rangle = E|0\rangle \quad (1.42)$$

або тотожна з нею

$$H(\mathbf{l})|\mathbf{l}\rangle = E|\mathbf{l}\rangle. \quad (1.42 \text{ а})$$

Оскільки всі точки ґратки Браве еквівалентні, то й усі розв'язки (хвильові функції), що відповідають трансляції на вектор ґратки, мають бути еквівалентні один одному. При цьому необхідно враховувати, що стани квантової системи можуть бути виродженими. Розглянемо випадки вироджених і невироджених станів окремо. Спочатку припустимо, що розв'язок  $|0\rangle$  *невироджений*. Розглядатимемо трансляцію на один крок у напрямку основного вектора ґратки Браве  $\mathbf{a}_1$ . Тоді еквівалентність вимагає, щоб функція  $|\mathbf{a}_1\rangle$  відрізнялась від  $|0\rangle$  не більше, як на множник. З фізичного погляду ці стани однакові, тобто має існувати таке число  $\lambda_1$ , за якого у результаті одного кроку у напрямку вектора  $\mathbf{a}_1$  хвильова функція трансльованої системи визначалась би як добуток цього числа на хвильову функцію нетрансльованої системи

$$|\mathbf{a}_1\rangle = \lambda_1 |0\rangle. \quad (1.43)$$

Оскільки хвильова функція є нормованою на одиницю, маємо рівняння для визначення  $\lambda_1$

$$|\lambda_1|^2 = 1. \quad (1.44)$$

Найпростішою функцією, що задовольняє (1.44), є експонента

$$\lambda_1 = e^{ik_1}, \quad (1.45)$$

де  $k_1$  – дійсне число. Аналогічно, для одиничних трансляцій у напрямку інших базисних векторів запишемо

$$|\mathbf{a}_2\rangle = e^{ik_2} |0\rangle \text{ та } |\mathbf{a}_3\rangle = e^{ik_3} |0\rangle. \quad (1.46)$$

Для трансляції на довільний вектор ґратки  $\mathbf{l}$  маємо

$$\begin{aligned} |\mathbf{l}\rangle &= |l_1 \mathbf{a}_1 + l_2 \mathbf{a}_2 + l_3 \mathbf{a}_3\rangle = e^{ik_1} |(l_1 - l) \mathbf{a}_1 + l_2 \mathbf{a}_2 + l_3 \mathbf{a}_3\rangle = \\ &= e^{i(k_1 + k_1)} |(l_1 - 2) \mathbf{a}_2 + l_2 \mathbf{a}_2 + l_3 \mathbf{a}_3\rangle = \dots \end{aligned} \quad (1.47)$$

$$= e^{ik_1 l_1} |0 + l_2 \mathbf{a}_2 + l_3 \mathbf{a}_3\rangle = \dots e^{ik_1 l_1} e^{ik_2 l_2} e^{ik_3 l_3} |0\rangle = e^{i(k_1 l_1 + k_2 l_2 + k_3 l_3)} |0\rangle.$$

Дійсно, при виконанні перетворень (1.47) було зроблено  $l_1$  кроків у напрямку  $\mathbf{a}_1$ ,  $l_2$  кроків – у напрямку  $\mathbf{a}_2$  та  $l_3$  кроків – у напрямку  $\mathbf{a}_3$ , і після кожного кроку функція домножувалася на відповідний множник. Нехай, наприклад вектор

$$\mathbf{k} = k_1 \mathbf{b}_1 + k_2 \mathbf{b}_2 + k_3 \mathbf{b}_3, \quad (1.48)$$

де  $\mathbf{b}_i$ ,  $i = 1, 2, 3$  – трійка векторів оберненої ґратки, що відповідає векторам  $\mathbf{a}_i$ ,  $i = 1, 2, 3$  прямої ґратки. Тоді можна записати

$$|\mathbf{l}\rangle = e^{i\mathbf{k}\cdot\mathbf{l}} |0\rangle. \quad (1.49)$$

Таким чином, ми отримали дуже важливий результат: для будь-якої хвильової функції, що задовольняє рівнянню Шредингера, існує такий вектор  $\mathbf{k}$ , за якого трансляція на вектор ґратки  $\mathbf{l}$  еквівалентна множенню цієї функції на фазовий фактор  $\exp(i\mathbf{k}\cdot\mathbf{l})$ .

У загальнішому випадку вироджених станів також можна сформулювати твердження, аналогічне (1.49). Строгий розгляд загального випадку подано у додатку А, де на прикладі дворазово виродженого стану показано, що для кожної функції  $|0\rangle_i$   $n$ -разово виродженого стану існує хвильовий вектор  $\mathbf{k}_i$  такий, за якого при трансляції на довільний вектор ґратки  $\mathbf{l}$  хвильова функція виродженого стану запишеться як

$$|\mathbf{l}\rangle_i = e^{i\mathbf{k}_i\cdot\mathbf{l}} |0\rangle_i. \quad (1.50)$$

При цьому враховується, що кожен вироджений стан  $|0\rangle_i$  може бути суперпозицією вироджених станів  $|0\rangle_{i_b}$ , тобто

$$|0\rangle_n = \mathbf{S} \times \begin{pmatrix} |0\rangle_1 \\ |0\rangle_2 \\ \vdots \\ |0\rangle_n \end{pmatrix}, \quad (1.51)$$

де  $\mathbf{S}$  – деяка унітарна матриця.

Ці рівняння складають суть такої загальної теореми:

**Будь-який розв'язок, що відповідає виродженому значенню енергії за трансляції системи на вектор ґратки  $\mathbf{l}$ , можна представити у вигляді лінійної комбінації розв'язків, що відповідають тій самій енергії. Коефіцієнтами у такій суперпозиції виступають комплексні експоненти  $\exp(i\mathbf{k}\cdot\mathbf{l})$ .**

Це й є одним із загальних формулювань теореми Блоха.

## ■ 1.5. Циклічні граничні умови Борна–Кармана. Приведення до зони Бриллюена

У випадку електронних хвиль із теореми Блоха випливає, що будь-яку електронну функцію можна характеризувати її хвильовим вектором  $\mathbf{k}$  і записати

$$\psi(\mathbf{r} + \mathbf{l}) = e^{i\mathbf{k}\cdot\mathbf{l}}\psi(\mathbf{r}). \quad (1.52)$$

Такій умові відповідає й хвильова функція вільного електрона

$$\psi^{(0)}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (1.53)$$

Цей факт не є несподіваним, оскільки й в цьому випадку можна розглядати розв'язок рівняння Шредингера із періодичним потенціалом, що всюди дорівнює нулю. Тобто за допомогою *порожньої ґратки* ми просто виконали перевірку, яка часто виявляється корисною при розгляді властивостей твердого тіла. Іноді при формулюванні теореми Блоха хвильову функцію електрона представляють у вигляді

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}). \quad (1.54)$$

Тоді, згідно із (1.54), функція  $u_{\mathbf{k}}(\mathbf{r})$  має бути періодичною, тобто

$$u_{\mathbf{k}}(\mathbf{r} + \mathbf{l}) = u_{\mathbf{k}}(\mathbf{r}). \quad (1.55)$$

Це твердження також часто називають *теоремою Блоха*, а функції  $u_{\mathbf{k}}(\mathbf{r})$  – блохівськими.

Якщо виявиться, що деякому стану відповідає вектор оберненої ґратки  $\mathbf{g}$ , то хвильова функція, відповідна такому стану, буде періодичною

$$\psi_{\mathbf{g}}(\mathbf{r} + \mathbf{l}) = e^{i\mathbf{g}\cdot\mathbf{l}}\psi_{\mathbf{g}}(\mathbf{r}) = \psi_{\mathbf{g}}(\mathbf{r}), \quad (1.56)$$

оскільки  $e^{i\mathbf{g}\cdot\mathbf{l}} = 1$  для всіх  $\mathbf{l}$ . Нехай тепер деякому стану, що описується хвильовою функцією  $\psi_{\mathbf{k}}$ , відповідає хвильовий вектор

$$\mathbf{k} = \mathbf{g} + \mathbf{k}', \quad (1.57)$$

де  $\mathbf{g}$  – деякий вектор оберненої ґратки, а  $\mathbf{k}'$  – новий хвильовий вектор. Тоді маємо

$$\psi_{\mathbf{k}}(\mathbf{r} + \mathbf{l}) = e^{i(\mathbf{g}+\mathbf{k}')\cdot\mathbf{l}}\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}'\cdot\mathbf{l}}\psi_{\mathbf{k}}(\mathbf{r}), \quad (1.58)$$

тобто функція  $\psi_{\mathbf{k}}$  задовольняє теоремі Блоха так, якби їй відповідав хвильовий вектор  $\mathbf{k}'$ . Первинний вибір хвильового вектора  $\mathbf{k}$  не є однозначним, кожній хвильовій функції відповідає вся множина можливих хвильових векторів, що відрізняються один від одного на вектор обер-

неної ґратки. Як однозначно визначити хвильовий вектор, що відповідає даній хвильовій функції? Розглянемо процедуру однозначного визначення хвильового вектора на прикладі одновимірної ґратки. Аналог оберненої ґратки у цьому випадку – набір обернених довжин ґратки

$$g_n = n \frac{2\pi}{a}. \quad (1.59)$$

Стану із даною енергією можна приписати будь-яке хвильове число із набору

$$k = n \frac{2\pi}{a} + k', \quad (1.60)$$

тобто хвильове число, визначене лише із точністю до  $n(2\pi/a)$ . Зазвичай як представник усіх чисел  $k$  вибирають хвильове число із найменшим можливим абсолютним значенням  $|k'|$ , а саме для хвильового числа завжди вибирається значення, що лежить в інтервалі

$$-\frac{\pi}{a} < k' \leq \frac{\pi}{a}. \quad (1.61)$$

Цей інтервал збігається із зоною Бриллюена одновимірної системи та е елементарною коміркою оберненої ґратки.

У випадку трьох вимірів можна виконати ту саму процедуру – вибрати хвильовий вектор в оберненому просторі, а далі – вектор  $\mathbf{g}$  так, щоб модуль вектора  $\mathbf{k}'$  був найменшим із можливих або точка  $\mathbf{k}'$  лежала найближче до початку координат в оберненому просторі, тобто до "нульового" вузла, ніж до будь-якого іншого із вузлів оберненої ґратки. Це рівнозначно твердженню, що ця точка лежить у комірці Вігнера-Зейтца оберненої ґратки, тобто у зоні Бриллюена. Таким чином будь-який вектор  $\mathbf{k}$  можна привести до відповідного вектора у зоні Бриллюена (рис. 1.28). Тому хвильову функцію можна описувати у схемі *приведених зон* або, іншими словами, будь-якій хвильовій функції відповідає приведений хвильовий вектор. Але може існувати багато функцій із тим самим приведеним хвильовим вектором, що відповідають різним енергіям.

З'ясуємо, що відбувається із власними функціями вільного електрона у порожній ґратці. Нехай

$$\psi = e^{i\mathbf{k}\cdot\mathbf{r}} = e^{i(\mathbf{k}-\mathbf{g})\cdot\mathbf{r}} e^{i\mathbf{g}\cdot\mathbf{r}} = e^{i\mathbf{k}'\cdot\mathbf{r}} (e^{i\mathbf{g}\cdot\mathbf{r}}), \quad (1.62)$$

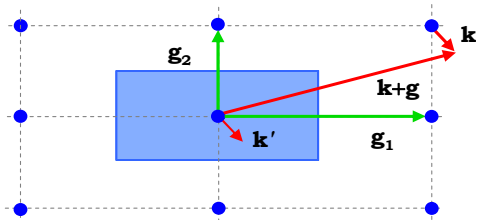


Рис. 1.28. Приведення до зони Бриллюена у двовимірній ґратці

де  $\mathbf{k}'$  – приведені значення істинного хвильового вектора. Іншими словами, ця хвильова функція має стандартний вигляд блохівської, оскільки  $\exp(i\mathbf{g} \cdot \mathbf{r})$  є періодичною функцією у заданій ґратці. І якщо енергія задається формулою

$$E(k) = \frac{\hbar^2 k^2}{2m}, \quad (1.63)$$

то вона виявляється багатозначною функцією  $k'$  у приведеній зоні.

Необхідно зазначити, що основною ідеєю, що приводила нас до того чи іншого результату, була ідея трансляційної інваріантності кристалічної ґратки, яка вважалась необмеженою. Але таким чином доводиться мати справу із нескінченним числом атомів і хвильових функцій. Якщо ж вважати, що ґратка обмежена, то варто брати до уваги наявність меж розподілу й необхідність задоволення граничних умов. Зрозуміло, що така програма не може бути реалізована із багатьох причин, тим паче, що існує математичний прийом, який дозволяє задовільно розв'язувати проблему підрахунку станів, не вводючи додаткових ефектів, пов'язаних із межами розподілу. Цей прийом полягає у застосуванні *циклічних граничних умов*, або *умов Борна–Кармана*.

В одновимірному випадку можна припустити, що кристал складається з  $L$  комірок, які з'єднані так, що утворюють замкнене коло. При цьому хвильова функція має задовольняти умові

$$\psi(x + La) = \psi(x), \quad (1.64)$$

що забезпечує її неперервність у точці з'єднання. З іншого боку, теорема Блоха дає змогу записати

$$\psi_k(x + La) = e^{ikLa} \psi_k(x), \quad (1.65)$$

що означає необхідність виконання умови

$$e^{ikLa} = 1 \text{ або } k = \frac{2\pi m}{La}, \quad (1.66)$$

де  $m$  – ціле число.

В одновимірній приведеній зоні  $-\pi/a < k < \pi/a$ , тому цілі числа з інтервалу

$$-\frac{L}{2} < m \leq \frac{L}{2} \quad (1.67)$$

дадуть усі можливі значення приведеного хвильового числа. Разом існує  $L$  таких значень, що відстоять одне від одного на відстані  $(1/L) \cdot (2\pi/a)$ . Оскільки величина  $L$  є дуже великою, то значення розподілені в оберненому просторі практично безперервно зі сталою щільністю.

Припустимо тепер, що система є циклічною у трьох вимірах, тобто являє кристал із розмірами вздовж трьох базисних напрямків ґратки –  $L_1 \mathbf{a}_1$ ,  $L_2 \mathbf{a}_2$  та  $L_3 \mathbf{a}_3$ . Тоді циклічні умови дають

$$\psi(\mathbf{r} + L_i \mathbf{a}_i) = \psi(\mathbf{r}), \quad i = 1, 2, 3. \quad (1.68)$$

Звідси для блохівських функцій, що характеризуються хвильовим вектором  $\mathbf{k}$ , маємо рівності

$$e^{i\mathbf{k}(L_1 \mathbf{a}_1)} = e^{i\mathbf{k}(L_2 \mathbf{a}_2)} = e^{i\mathbf{k}(L_3 \mathbf{a}_3)} = 1, \quad (1.69)$$

що задовольняються лише за умови

$$\mathbf{k} = k_1 \mathbf{b}_1 + k_2 \mathbf{b}_2 + k_3 \mathbf{b}_3 = \sum_{i=1}^3 \frac{2\pi m_i}{L_i} \mathbf{b}_i, \quad (1.70)$$

де  $m_i$  – цілі числа, а  $\mathbf{b}_i$  – вектори оберненої ґратки. Отже дозволені значення вектора  $\mathbf{k}$  можна отримати за допомогою поділу утворюючих векторів оберненої ґратки на  $L_i$  частин для  $i$ -го напрямку. Таким чином, обернений простір складається із рівномірно розподілених точок, на яких визначено вектори  $\mathbf{k}$ . При обчисленні щільності цих точок варто пам'ятати, що вони можуть накрити всю елементарну комірку оберненої ґратки, якщо цілі числа  $m_i$  пробігають значення  $0 \leq m_i < L$ . Можна розглядати симетричну область

$$-L_i/2 < m_i < L_i/2. \quad (1.71)$$

Таким чином ми отримали б елементарну комірку, яка має вигляд паралелепіпеда із центром на початку координат. Але зручніше вибрати комірку у вигляді комірки Вігнера–Зейтца, тобто зони Бриллюена. Об'єм зони Бриллюена збігається з об'ємом елементарної комірки, що має форму паралелепіпеда. Із цієї причини у ній має міститися та ж сама кількість дозволених значень вектора  $\mathbf{k}$ . Згідно із нерівностями (1.71) це число дорівнює добутку вузлів ґратки вздовж кожної із трьох осей

$$L_1 \times L_2 \times L_3 = N, \quad (1.72)$$

що збігається із кількістю елементарних комірок у всьому макроскопічному кристалі. Таким чином можна сформулювати дуже важливу теорему:

**Зона Бриллюена містить рівно стільки дозволених хвильових векторів, скільки елементарних комірок містить блок кристала.**

Об'єм зони Бриллюена дорівнює  $8\pi^3/v_c$  (див. задачу 1). Якщо в об'ємі кристала міститься  $N$  елементарних комірок, то об'єм елементарної комірки у прямому просторі

$$v_c = V/N, \quad (1.73)$$

і тоді на кожний дозволений вектор  $\mathbf{k}$  у  $\mathbf{k}$ -просторі приходиться об'єм

$$\frac{8\pi^3}{v_c} \frac{1}{N} = \frac{8\pi^3}{V}, \quad (1.74)$$

тобто в одиниці об'єму оберненого простору існують  $V/8\pi^3$  дозволених  $\mathbf{k}$ -векторів. Оскільки величина  $N$  дуже велика, то розподіл дозволених векторів можна розглядати як неперервний. На основі цього твердження часто виконують перехід від підсумовування до інтегрування:

$$\sum_{\mathbf{k}} f(\mathbf{k}) \rightarrow \frac{V}{8\pi^3} \int d\mathbf{k} f(\mathbf{k}). \quad (1.75)$$

Ця формула дуже часто використовується у фізиці твердого тіла.

## ■ 1.6. Задачі

1. Доведіть, що вектори оберненої і прямої ґратки задовольняють співвідношенням

$$\mathbf{b}_1 \times (\mathbf{b}_2 \times \mathbf{b}_3) = \frac{(2\pi)^3}{\mathbf{a}_1 \times (\mathbf{a}_2 \times \mathbf{a}_3)}, \quad (1.76)$$

або що об'єм елементарної комірки в оберненій ґратці обернено пропорційний до об'єму елементарної комірки прямої ґратки:

$$v_r = (2\pi)^3 / \Omega_0.$$

*Розв'язок.* Скористаємося виразом для вектора оберненої ґратки

$$\mathbf{b}_1 = 2\pi \frac{(\mathbf{a}_2 \times \mathbf{a}_3)}{\mathbf{a}_1 \times (\mathbf{a}_2 \times \mathbf{a}_3)}.$$

Запишемо об'єм елементарної комірки оберненої ґратки за допомогою цієї формули

$$\mathbf{b}_1 \times (\mathbf{b}_2 \times \mathbf{b}_3) = 2\pi \frac{(\mathbf{a}_2 \times \mathbf{a}_3) \times (\mathbf{b}_2 \times \mathbf{b}_3)}{\mathbf{a}_1 \times (\mathbf{a}_2 \times \mathbf{a}_3)}. \quad (1.77)$$

Використаємо тепер тотожність Лагранжа

$$(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \times \mathbf{c})(\mathbf{b} \times \mathbf{d}) - (\mathbf{b} \times \mathbf{c})(\mathbf{a} \times \mathbf{d})$$

і вираз  $\mathbf{b}_i \times \mathbf{a}_j = 2\pi \delta_{ij}$ . У результаті маємо

$$(\mathbf{a}_2 \times \mathbf{a}_3) \times (\mathbf{b}_2 \times \mathbf{b}_3) = (\mathbf{a}_2 \times \mathbf{b}_2)(\mathbf{a}_3 \times \mathbf{b}_3) = (2\pi)^2.$$

Тоді із (1.77) отримуємо (1.76).

2. Нехай вектори прямої ґратки  $\alpha_i$  побудовано на векторах оберненої ґратки за таким самим правилом, що й вектори оберненої ґратки через основні вектори ґратки  $\mathbf{a}_i$  Браве. Доведіть, що ці вектори збігаються, тобто  $\alpha_i = \mathbf{a}_i$ .

*Розв'язок.* Необхідно довести, зокрема, що

$$\alpha_1 = 2\pi \frac{\mathbf{b}_2 \times \mathbf{b}_3}{\mathbf{b}_1 \times (\mathbf{b}_2 \times \mathbf{b}_3)} = \mathbf{a}_1. \quad (1.78)$$



Для цього запишемо

$$\mathbf{b}_3 = 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \times (\mathbf{a}_2 \times \mathbf{a}_3)}$$

і підставимо вираз до (1.78). Маємо

$$\alpha_1 = (2\pi)^2 \frac{1}{\mathbf{b}_1 \times (\mathbf{b}_2 \times \mathbf{b}_3)} \times \frac{1}{\mathbf{a}_1 \times (\mathbf{a}_2 \times \mathbf{a}_3)} (\mathbf{b}_2 \times (\mathbf{a}_1 \times \mathbf{a}_2)).$$

Тепер запишемо, що  $(\mathbf{b}_2 \times (\mathbf{a}_1 \times \mathbf{a}_2)) = \mathbf{a}_1(\mathbf{b}_2 \times \mathbf{a}_2) - \mathbf{a}_2(\mathbf{b}_2 \times \mathbf{a}_1) = 2\pi \mathbf{a}_1$ . Користуючись результатом попередньої задачі, маємо

$$\alpha_1 = (2\pi)^2 \frac{1}{\mathbf{b}_1 \times (\mathbf{b}_2 \times \mathbf{b}_3)} \times \frac{2\pi \mathbf{a}_1}{\mathbf{a}_1 \times (\mathbf{a}_2 \times \mathbf{a}_3)} = \mathbf{a}_1.$$

Подібні перетворення можна записати й для інших векторів ґратки. Таким чином ми строго отримали дуже важливий результат:

### Обернена ґратка до оберненої ґратки – вихідна ґратка Браве.

3. Доведіть, що для будь-якого вектора ґратки Браве  $\mathbf{R}$  виконується тотожність

$$\sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} = N\delta_{\mathbf{k},\mathbf{g}}. \quad (1.79)$$

**Розв'язок.** Вектор  $\mathbf{R}$  пробігає всіма  $N$  вузлами ґратки.

$$\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3, \quad 0 \leq n_i < N_i, \quad N_1 N_2 N_3 = N. \quad (1.80)$$

Якщо хвильовий вектор  $\mathbf{k}$  задовольняє граничним умовам Борна–Кармана, то величина суми не зміниться, якщо вектор  $\mathbf{R}$  зсунути на довільний вектор ґратки Браве  $\mathbf{R}'$ , тобто

$$\sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R}+\mathbf{R}')} = e^{i\mathbf{k}\cdot\mathbf{R}'} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}. \quad (1.81)$$

Звідси робимо висновок, що сума має дорівнювати нулю, якщо передекспонента у правій частині формули не є одиницею. Але, як відомо, передекспонента у правій частині (1.81) дорівнює одиниці тільки за умови, що  $\mathbf{k}$  є одним із векторів оберненої ґратки. Нехай цей вектор оберненої ґратки є нулем. Тоді сума в (1.81) просто дорівнюватиме  $N$ , або

$$\sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} = N\delta_{\mathbf{k},\mathbf{0}}. \quad (1.82)$$

З іншого боку будь-який хвильовий вектор можна привести до першої зони Бриллюена, якщо записати його через суму вектора із першої зони Бриллюена  $\mathbf{k}_1$  та відповідного вектора оберненої ґратки:

$$\mathbf{k} = \mathbf{k}_1 + \mathbf{g}. \quad (1.83)$$

Із (1.82) випливає, що для ненульової суми вимагається, щоб  $\mathbf{k}_1 = \mathbf{0}$ . Тоді для довільного вектора із рівняння (1.82) випливає  $\mathbf{k} = \mathbf{g}$ , тобто тотожність (1.79) доведено.

4. Доведіть тотожність

$$\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}} = N\delta_{\mathbf{R},0}, \quad (1.84)$$

якщо  $\mathbf{R}$  – будь-який вектор ґратки Браве, а сума береться за всіма  $\mathbf{k}$  у першій зоні Бриллюена, що задовольняють граничним умовам Борна–Кармана.

**Розв'язок.** Сума не має змінитися, якщо кожен вектор  $\mathbf{k}$  перенести на один і той самий вектор  $\mathbf{k}_0$  із першої зони Бриллюена, який, крім того, задовольняє умовам Борна–Кармана. Дійсно, примітивну комірку, що виникає внаслідок зсуву всієї першої зони на вектор  $\mathbf{k}_0$ , можна повернути до першої зони, якщо розбити її на частини та зсунути на відповідні вектори оберненої ґратки. Оскільки жодний доданок  $e^{i\mathbf{k}\cdot\mathbf{R}}$  не зміниться за зсуву  $\mathbf{k}$  на вектор оберненої ґратки, сума за зсунутою зоною тотожно дорівнює сумі за висхідною зоною Бриллюена

$$\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}} = e^{i\mathbf{k}_0\cdot\mathbf{R}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}}, \quad (1.85)$$

тобто сума у лівій частині має дорівнювати нулю, якщо тільки експонента не дорівнює одиниці для всіх векторів  $\mathbf{k}_0$ . Єдиний вектор  $\mathbf{R}$  ґратки Браве, для якого це можливо, це  $\mathbf{R} = 0$ . Але тоді сума просто дає  $N$ .

5. Побудуйте примітивну комірку та зону Бриллюена для двовимірної ґратки Браве, що утворюється вузлами сітки правильних шестикутників.

6. Покажіть, що в кристалі можуть бути тільки поворотні осі симетрії другого, третього, четвертого та шостого порядків.

**Указівка.** Поверніть кристал навкруги вісі, що проходить через деякий вузол ґратки Браве, і скористайтеся властивостями ґратки.

7. Дайте відповідь на запитання:

- Як будується зона Бриллюена?
- Чому дорівнює її об'єм?
- Які властивості симетрії кристалічної ґратки визначають структуру зони Бриллюена?
- Скільки станів (дозволених векторів  $\mathbf{k}$ ) міститься у зоні Бриллюена?

## Список літератури

1. Займан Дж. Принципы теории твердого тела. – М.: Мир, 1974.
2. Ашкрофт Н., Мермин Н. Физика твердого тела: В 2 т. – М.: Мир, 1979.
3. Давыдов А.С. Теория твердого тела. – М.: Наука, 1976.
4. Киттель Ч. Введение в физику твердого тела. – М.: Наука, 1978.
5. Neamen D. Semiconductor Physics and Devices. – N.Y.: McGraw-Hill, 2003.
6. Kasap S.O. Principles of Electronic Materials and Devices. – N.Y.: McGraw-Hill, 2005.

## Розділ 2

# ПРИНЦИПИ КВАНТОВОЇ ТЕОРІЇ БАГАТЬОХ ЧАСТИНОК У ФІЗИЦІ ТВЕРДОГО ТІЛА

Подібно до того, як в атомній фізиці основним завданням є визначення станів електронів в атомі та обчислення дозволених рівнів енергії, що можуть мати електрони в атомі, одним із головних завдань фізики твердого тіла є визначення енергетичного спектра та стаціонарних станів електронів у кристалі. Оскільки тверде тіло є багатоелектронною системою, то у загальному випадку таке завдання технічно не може бути точно розв'язано, але окремі важливі результати можливо отримати за певних наближень і низки припущень. Ці наближення й моделі мають принциповий характер та є основою фізики твердого тіла, зокрема фізики напівпровідників. Із цієї причини ми змушені нижче навести досить громіздкі викладки та детальні обчислення, сподіваючись, що наші зусилля буде винагороджено чітким розумінням того факту, що фізика напівпровідників базується на міцному підмурку добре обґрунтованих моделей і наближень.

### ■ 2.1. Адіабатичне наближення

Кристал являє собою єдину систему ядер та електронів. Якщо координати електронів позначити через  $\mathbf{r}_1, \mathbf{r}_2, \dots$ , а ядер – через  $\mathbf{R}_1, \mathbf{R}_2$ , то хвильова функція, очевидно залежатиме від координат усіх частинок

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots; \mathbf{R}_1, \mathbf{R}_2, \dots) = \Psi(\mathbf{r}_e; \mathbf{R}_\alpha). \quad (2.1)$$

Гамільтоніан такої системи містить усі види енергії:

♦ **кінетичну енергію електронів**

$$T_e = - \sum_i \frac{\hbar^2}{2m} \Delta_i; \quad (2.2)$$

♦ **кінетичну енергію ядер**

$$T_z = - \sum_i \frac{\hbar^2}{2M_\alpha} \Delta_\alpha; \quad (2.3)$$

♦ **енергію попарної взаємодії електронів**

$$U_e = \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|}; \quad (2.4)$$

## ♦ енергію попарної взаємодії ядер

$$U_Z = \frac{1}{2} \sum_{\alpha \neq \beta} \frac{Z_\alpha Z_\beta e^2}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}; \quad (2.5)$$

## ♦ енергію взаємодії електронів з ядрами

$$U_{ez} = -\frac{1}{2} \sum_{\alpha, i} \frac{Z_\alpha e^2}{|\mathbf{R}_\alpha - \mathbf{r}_i|}; \quad (2.6)$$

## ♦ енергію електронів та ядер у зовнішньому полі

$$V = V(\mathbf{r}_i; \mathbf{R}_\alpha). \quad (2.7)$$

Тобто гамільтоніан кристала у зовнішньому полі має вигляд

$$H = T_e + T_Z + U_e + U_Z + U_{ez} + V, \quad (2.8)$$

а рівняння Шредингера для кристала, що складається із  $Z$  атомів, містить  $3(Z + 1)N$  змінних і має у стаціонарному випадку стандартний вигляд

$$H\Psi = E\Psi. \quad (2.9)$$

Розв'язання рівняння (2.9) неможливе з багатьох причин, принципів і суто технічних ( $N$  – дуже велике число). Як видно з рівнянь, параметрами задачі є маси електронів та ядер, і тому припустимо, що першим кроком до спрощення задачі може бути врахування істотної різниці у масі частинок, що складають кристал – легких електронів і важких ядер. Зрозуміло, що відгук електронів та ядер на однакове збурення буде різний перш за все через різну інерційність частинок. Наближення, що враховує істотно різний характер руху легких електронів і важких ядер у кристалі для розв'язку (2.9), називається *адіабатичним*. Основна ідея наближення полягає в тому, що для електронів, які здібні швидко змінювати свій рух під дією сили (оскільки вони є легкими порівняно з важкими ядрами атомів), важливим є миттєве положення ядер, на рух яких у свою чергу має впливати усереднений рух електронів. Почнемо із грубого припущення нерухомих ядер, тобто вважатимемо, що електрони рухаються у замороженій ґратці із координатами вузлів  $\mathbf{R}_\alpha = \mathbf{R}_\alpha^0$ . У цьому випадку кінетична енергія ядер занулюється, а потенціальна енергія взаємодії ядер стає константою і відповідним вибором початку відліку енергії може також занулитися. Таким чином, маємо гамільтоніан електронів у замороженій ґратці:

$$H_e = T_e + U_e + U_{ez}. \quad (2.10)$$

Хвильова функція електронів  $\Psi_e(\mathbf{r}_i; \mathbf{R}_\alpha^0)$ , що задовольняє рівнянню

$$H_e \Psi_e = E \Psi_e \quad (2.11)$$

є нормованою на одиницю для будь-якого набору координат ядер

$$\int d\tau_e \Psi_e^*(\mathbf{r}_1, \dots; \mathbf{R}_\alpha^0) \Psi_e(\mathbf{r}_1, \dots; \mathbf{R}_\alpha^0) = 1, \quad d\tau_e = d\mathbf{r}_1 d\mathbf{r}_2 \dots, \quad (2.12)$$

тобто залежить від координат ядер як від параметрів. У свою чергу від координат ядер залежатиме й енергія електронів, що рухаються у полі замороженої ґратки

$$E_e(\mathbf{R}_\alpha^0) = \int d\tau_e \Psi_e^* H_e \Psi_e. \quad (2.13)$$

Врахуємо тепер рух ядер. Для цього запишемо ядерну частину гамільтоніана як

$$H_Z = T_Z + U_Z + E_e(R_\alpha) \quad (2.14)$$

і припустимо, що динаміка ядер описується хвильовою функцією, залежною тільки від координат ядер  $\Phi_Z(\dots \mathbf{R}_\alpha \dots)$ . Гамільтоніан кристала за відсутності зовнішнього поля запишемо як

$$H = H_e + H_Z - E_e. \quad (2.15)$$

Оскільки гамільтоніан (2.15) є адитивним щодо електронної та ядерної підсистем, хвильову функцію кристала можна представити як добуток електронної хвильової функції (тобто власної функції гамільтоніана (2.10), обчисленої для довільного розподілу координат ядер) і хвильової функції ядер

$$\Psi(\dots, \mathbf{r}_i, \dots; \dots \mathbf{R}_\alpha, \dots) = \Psi_e(\dots, \mathbf{r}_i, \dots; \dots \mathbf{R}_\alpha, \dots) \Phi_Z(\dots \mathbf{R}_\alpha, \dots). \quad (2.16)$$

Підставимо цю функцію до рівняння Шредингера та скористаємося тим, що, як випливає із (2.10),  $\Phi_Z H_e \Psi_e = \Phi_Z E_e \Psi_e$ . Маємо

$$H\Psi = \Phi_Z H_e \Psi_e + H_Z \Psi_e \Phi_Z - E_e \Psi_e \Phi_Z. \quad (2.17)$$

З огляду на (2.11), отримаємо

$$H\Psi = H_Z \Psi_e \Phi_Z = E \Psi_e \Phi_Z = E\Psi. \quad (2.18)$$

Знайдемо результат дії оператора Лапласа на хвильову функцію (2.16) із врахуванням того, що  $\Psi_e$  залежить від координат ядер

$$\Delta_\alpha \Psi = \left[ \frac{\partial^2}{\partial X_\alpha^2} + \frac{\partial^2}{\partial Y_\alpha^2} + \frac{\partial^2}{\partial Z_\alpha^2} \right] \Psi,$$

де  $\mathbf{R}_\alpha = (X_\alpha, Y_\alpha, Z_\alpha)$  – координата  $\alpha$ -го ядра. У результаті отримуємо

$$\Delta_\alpha \Psi_e \Phi_Z = \Delta_\alpha (\Psi_e \Delta_\alpha \Phi_Z + \Phi_Z \Delta_\alpha \Psi_e) = \Psi_e \Delta_\alpha \Phi_Z + \Phi_Z \Delta_\alpha \Psi_e + 2(\Delta_\alpha \Psi_e \cdot \Delta_\alpha \Phi_Z). \quad (2.19)$$

Тоді із (2.18) маємо

$$\sum_\alpha \left( -\frac{\hbar^2}{2M} \right) \left[ \Psi_e \Delta_\alpha \Phi_Z + \Phi_Z \Delta_\alpha \Psi_e + 2\Delta_\alpha \Psi_e \cdot \Delta_\alpha \Phi_Z \right] + U_Z \Psi_e \Phi_Z + E_e \Psi_e \Phi_Z = E \Psi_e \Phi_Z. \quad (2.20)$$

Подіємо на це рівняння інтегральним оператором

$$\hat{I} = \int d\tau_e \Psi_e^* \dots \quad (2.21)$$

і врахуємо, що електронна хвильова функція нормована на одиницю  $\int d\tau_e \Psi_e^* \Psi_e = 1$  (див. (2.12)). У результаті отримуємо

$$\left\{ \sum_{\alpha} \left( -\frac{\hbar^2}{2M} \right) \Delta_{\alpha} + U_Z + E_e \right\} \Phi_Z + \sum_{\alpha} \left( -\frac{\hbar^2}{2M} \right) \left[ \Phi_Z \int d\tau_e \Psi_e^* \Delta_{\alpha} \Psi_e + 2\Delta_{\alpha} \Phi_Z \int d\tau_e \Psi_e^* \Delta_{\alpha} \Psi_e \right] = E \Phi_Z. \quad (2.22)$$

Оскільки у фігурних дужках стоїть ядерна частина гамільтонана, перепишемо рівняння у вигляді

$$H_Z \Phi_Z = E \Phi_Z + \sum_{\alpha} \left( \frac{\hbar^2}{2M} \right) \left[ \Phi_Z \int d\tau_e \Psi_e^* \Delta_{\alpha} \Psi_e + 2\nabla_{\alpha} \Phi_Z \int d\tau_e \Psi_e^* \Delta_{\alpha} \Psi_e \right]. \quad (2.23)$$

Діючи на це рівняння оператором

$$\hat{I}_Z = \int d\tau_Z \Phi_Z^* \dots \quad (2.24)$$

і враховуючи, що ядерна хвильова функція є нормованою на одиницю, маємо для енергії кристала

$$\int d\tau_Z \Phi_Z^* H_Z \Phi_Z = E + \delta E, \quad (2.25)$$

де  $\delta E$  – дефіцит енергії, який, згідно із (2.23), визначається формулою

$$\delta E = \sum_{\alpha} \left( \frac{\hbar^2}{2M} \right) \left[ \int d\tau_Z \Phi_Z^* \Phi_Z \int d\tau_e \Psi_e^* \Delta_{\alpha} \Psi_e + 2 \int d\tau_Z \Phi_Z^* \nabla_{\alpha} \Phi_Z \int d\tau_e \Psi_e^* \nabla_{\alpha} \Psi_e \right]. \quad (2.26)$$

Для оцінювання  $\delta E$  згадаємо, що електрон-електронна взаємодія обернено пропорційна відстані від одного електрона до іншого, тобто для грубої оцінки такою взаємодією, порівняно із кінетичною енергією електронів, можна знехтувати. Таким чином,  $H_e$  перетворюється на гамільтоніан системи електронів, які не взаємодіють, отже за хвильову функцію електронів можна взяти комбінацію атомних хвильових функцій, що залежать від модуля  $|\mathbf{r}_i - \mathbf{R}_{\alpha}|$ . Тоді можна записати

$$\Delta_{\alpha} \Psi_e(|\mathbf{r}_i - \mathbf{R}_{\alpha}|) = \left( \frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2} \right) \Psi_e(|\mathbf{r}_i - \mathbf{R}_{\alpha}|) \rightarrow, \quad (2.27)$$

$$\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \Psi_e(|\mathbf{r}_i - \mathbf{R}_{\alpha}|) = \Delta_i \Psi_e(|\mathbf{r}_i - \mathbf{R}_{\alpha}|).$$

Згідно із (2.27) перший доданок у правій частині (2.26) перепишемо наступним чином

$$\sum_{\alpha} \left( \frac{\hbar^2}{2M_{\alpha}} \right) \int d\tau_e \Psi_e^* \Delta_{\alpha} \Psi_e = - \sum_{\alpha, i} \left( \frac{m}{M_{\alpha}} \right) \int d\tau_e \Psi_e^* \left( - \frac{\hbar^2}{2m} \Delta_i \right) \Psi_e = - \sum_{\alpha, i} \left( \frac{m}{M_{\alpha}} \right) \langle T_i \rangle, \quad (2.28)$$

де  $\langle T_i \rangle$  – середня кінетична енергія електрона. Зрозуміло, що у кристалі, який містить  $N$  електронів та  $Z$  ядер, сума (2.28) дає просто  $NZ(m/V) \langle T_i \rangle$ . Порівняно з енергією кристала, що складає за порядком величини кінетичну енергію електронів та ядер, ця величина є малою за мірою малості відношення  $m/M$  (напр., для германію таке відношення  $\sim 10^{-5}$ ). Другий доданок у (2.26) оцінюємо так:

$$\begin{aligned} & \sum_{\alpha} \left( \frac{\hbar^2}{M_{\alpha}} \right) \int d\tau_z \Phi_z^* \nabla_{\alpha} \Phi_z \int d\tau_e \Psi_e^* \nabla_{\alpha} \Psi_e \cong \\ & - \frac{1}{M} \sum_{\alpha, i} \int d\tau_z \Phi_z^* (-i\hbar \nabla_{\alpha}) \Phi_z \int d\tau_e \Psi_e^* (-i\hbar \nabla_i) \Psi_e = - \frac{1}{M} NZ \langle p_{\alpha} \rangle \langle p_i \rangle. \end{aligned} \quad (2.29)$$

Оскільки у термодинамічній рівновазі середня кінетична енергія електронів дорівнює середній кінетичній енергії ядер (у випадку класичної статистики, тобто т. зв. невідродженого напівпровідника):

$$\left\langle \frac{p_i^2}{2m} \right\rangle = \left\langle \frac{p_{\alpha}^2}{2M} \right\rangle \quad \text{і} \quad \langle p_i \rangle^2 = \frac{8}{3\pi} \langle p_i^2 \rangle, \quad \text{то} \quad \langle p_i \rangle \approx \sqrt{m/M} \langle p_{\alpha} \rangle.$$

Таким чином ми з'ясували, що другий доданок у (2.26) є величиною порядку  $\sqrt{m/M}$  повної енергії кристала. Ця величина, наприклад для германію, складає  $\sim 0.003$ . Отже, адіабатичне наближення полягає у нехтуванні поправки  $\delta E$ , і для визначення енергії кристала із досить великою точністю можна користуватися рівняннями

$$H_e \Psi_e = E_e \Psi_e, \quad (2.30)$$

$$H_z \Phi_z = E \Phi_z, \quad (2.31)$$

які показують, що:

**В адіабатичному наближенні хвильова функція електронів визначається миттєвою конфігурацією ядер (доданок  $U_{ez}$  в електронному гамільтоніані). У той же час хвильова функція ядер визначається усередненим полем електронів (доданок  $E_e$  в ядерній частині гамільтоніана).**

Таким чином, наближення, що приводить до рівнянь (2.30) і (2.31), виключає прямий обмін енергією між електронною та ядерною підсистемами, тому й має назву адіабатичного.

Наступним кроком може бути припущення, що електрони під час руху взаємодіють із рухомими ядрами. Оскільки ядра в основному стані знаходяться у мінімумі потенціальної енергії, то відхилення їхніх координат від рівноважного значення  $\mathbf{R}_a^{(0)}$  носитиме коливальний характер. При цьому такі зміщення мають бути невеликими, тобто координата  $a$ -го ядра представляється у вигляді  $\mathbf{R}_a = \mathbf{R}_a^{(0)} + \mathbf{r}_a$ , де зміщення ядер припускається набагато меншим за сталу ґратки,  $|\mathbf{r}_a| \ll a$ . Взаємодія електронів із такими коливаннями ядер, що враховує динамічну взаємодію електронів і ґратки, називається електрон-фононною взаємодією й буде розглянута детальніше у розд. 6.

## ■ 2.2. Одноелектронне наближення

У попередньому розділі ми з'ясували, що із досить великою точністю ( $\sim \sqrt{m/M}$ ) рух електронів та ядер може розглядатись розрізнено, отже значно спростили задачу визначення властивостей кристалів. Але таке спрощення ще не дозволяє розв'язати поставлену задачу. Дійсно, (2.30) не може бути розв'язано точно, оскільки в ньому є доданок, що визначає попарну взаємодію всіх електронів. Цей факт робить розв'язок задачі нездійсненним. Однак за припущення, що електрони не взаємодіють між собою, а взаємодіють із полем ядер, яке заекрановане всіма іншими електронами (крім того, що розглядається), це рівняння розпадається на систему рівнянь для кожного електрона. Кожне таке рівняння можна легко розв'язати. Таким чином виникає ідея врахувати взаємодію всіх електронів так, щоб у підсумку розглядати електрон, що знаходиться в деякому полі, яке утворюється всіма іншими електронами системи. Оскільки стан усіх електронів, крім  $i$ -го, залежить від того, як він рухається, то результуюче поле, що діє на  $i$ -й електрон, залежатиме від стану  $i$ -го електрона, тобто буде *самоузгодженим*. Покажемо, як можна провести процедуру визначення самоузгодженого поля. Позначимо потенціальну енергію  $i$ -го електрона у полі всіх інших (крім  $i$ -го) електронів як  $Q_i$ . Ця енергія залежить тільки від координат  $i$ -го електрона. Припустимо, що ми знайшли таке поле. Тоді електронний гамільтоніан кристала запишеться у вигляді суми одноелектронних гамільтоніанів



$$H_e = \sum_i H_i^e, \quad (2.32)$$

де гамільтоніаном  $i$ -го електрона є

$$H_i^e = -\frac{\hbar^2}{2m}\Delta_i + Q_i(\mathbf{r}_i) + U_Z(\mathbf{r}_i). \quad (2.33)$$

Тут  $U_Z(\mathbf{r}_i)$  – поле всіх ядер. Тепер хвильова функція системи електронів розпадається на добуток хвильових функцій окремих електронів

$$\Psi_e(\mathbf{r}_1, \mathbf{r}_2, \dots) = \prod_i \psi_i(\mathbf{r}_i), \quad (2.34)$$

а енергія системи дорівнює сумі енергій окремих електронів

$$E_e = \sum_i E_i, \quad (2.35)$$

які визначаються з рівняння

$$H_i^e \psi_i = E_i \psi_i. \quad (2.36)$$

Для визначення самоузгодженого поля запишемо два рівняння Шредингера для електронної підсистеми – точно і таке, де замість потенціалу парного електрон-електронної взаємодії записано самоузгоджене поле  $Q_i(\mathbf{r}_i)$

$$\sum_i \left( -\frac{\hbar^2}{2m} \Delta_i \right) \Psi_e + \frac{1}{2} \sum_{i \neq j} U_{ij} \Psi_e + \sum_i U_Z(\mathbf{r}_i) \Psi_e = E_e \Psi_e, \quad (2.37)$$

$$\sum_i \left( -\frac{\hbar^2}{2m} \Delta_i \right) \Psi_e + \sum_i Q_i(\mathbf{r}_i) \Psi_e + \sum_i U_Z(\mathbf{r}_i) \Psi_e = E_e \Psi_e. \quad (2.38)$$

Подіємо на ці рівняння оператором (2.21) і віднімемо від першого отриманого рівняння друге. Маємо

$$\frac{1}{2} \int d\mathbf{r}_i \Psi_e^* \sum_{i \neq j} U_{ij} \Psi_e - \int d\mathbf{r}_i \Psi_e^* \sum_i Q_i(\mathbf{r}_i) \Psi_e = 0. \quad (2.39)$$

Використовуючи (2.34), перепишемо (2.39) у вигляді

$$\begin{aligned} & \sum_i \int d\mathbf{r}_1 d\mathbf{r}_2 \dots \psi_1^*(\mathbf{r}_1) \dots \psi_i^*(\mathbf{r}_i) Q_i(\mathbf{r}_i) \psi_i(\mathbf{r}_i) \dots \psi_1(\mathbf{r}_1) = \\ & = \sum_i \int d\mathbf{r}_i \psi_i^*(\mathbf{r}_i) Q_i(\mathbf{r}_i) \psi_i(\mathbf{r}_i) = \frac{1}{2} \sum_{i \neq j, j} \int d\mathbf{r}_1 d\mathbf{r}_2 \dots \psi_1^*(\mathbf{r}_1) U_{ij}(|\mathbf{r}_i - \mathbf{r}_j|) \psi_1(\mathbf{r}_1) \dots \\ & = \sum_i \int d\mathbf{r}_i \psi_i^*(\mathbf{r}_i) \left\{ \frac{1}{2} \sum_{j \neq i} \int d\mathbf{r}_j \psi_j^*(\mathbf{r}_j) U_{ij}(|\mathbf{r}_i - \mathbf{r}_j|) \psi_j(\mathbf{r}_j) \right\} \psi_i(\mathbf{r}_i). \end{aligned}$$

Оскільки функції  $\psi_j(\mathbf{r})$  становлять повну ортонормовану систему, то звідси отримуємо, що вираз у фігурних дужках відіграє роль самоузгодженого поля.

$$Q_i(\mathbf{r}_i) = \frac{1}{2} \sum_{j \neq i} \int d\mathbf{r}_j \psi_j^*(\mathbf{r}_j) U_{ij}(|\mathbf{r}_i - \mathbf{r}_j|) \psi_j(\mathbf{r}_j), \quad (2.40)$$

або, виписуючи в явному вигляді потенціал електрон-електронної взаємодії  $U_{ij}$ ,

$$Q_i(\mathbf{r}_i) = \frac{1}{2} \sum_{j \neq i} \int d\mathbf{r}_j |\psi_j(\mathbf{r}_j)|^2 \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (2.41)$$

Оскільки  $e|\psi_j(\mathbf{r}_j)|^2 d\mathbf{r}_j$  є елементом заряду, що розташований в інфінітезимально малому об'ємі, який оточує точку  $\mathbf{r}_j$  і визначає потенціал цього заряду у точці  $\mathbf{r}_i$ , інтеграл у сумі (2.41) є енергією взаємодії  $i$ -го електрона із зарядом  $j$ -го електрона, що "розмазаний" по всьому кристалу. За уважного розгляду (2.41) бачимо, що для визначення самоузгодженого поля необхідно знати хвильову функцію електронів, яку можна отримати із рівняння Шредингера, що вміщує невизначене самоузгоджене поле. Таким чином знаходимо самоузгоджене рівняння (рівняння Хартрі) для визначення функцій  $\psi_i(\mathbf{r})$

$$-\frac{\hbar^2}{2m} \Delta \psi_i(\mathbf{r}_i) + \left[ \frac{1}{2} \sum_{j \neq i} \int |\psi_j(\mathbf{r}_j)|^2 \frac{e(ed\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \psi_i(\mathbf{r}_i) + U_i(\mathbf{r}_i; \mathbf{R}_\alpha) \psi_i(\mathbf{r}_i) = E_i \psi_i(\mathbf{r}_i). \quad (2.42)$$

Це рівняння зазвичай розв'язують так: визначають затравочні функції  $\psi_j^{(0)}(\mathbf{r}_j)$  (напр., хвильові функції вільних електронів), обчислюють потенціал  $Q_i(\mathbf{r}_i)$  за (2.41). Значення потенціалу підставляють до (2.42) і визначають функції  $\psi_j^{(1)}(\mathbf{r}_j)$  і т.д. При виконанні такої ітераційної процедури  $K$  разів можна отримати розв'язок рівняння із заданою точністю. Але постає питання про збіжність процедури. До того ж, у (2.42) не враховується, що електрони є фермі-частинками (тобто принцип Паулі, згідно з яким хвильова функція електронів має бути антисиметричною щодо перестановки двох електронів з урахуванням їхніх координат і спіну). Функція (2.34) не задовольняє цій вимозі. Для врахування цієї обставини необхідно замість (2.34) використовувати іншу функцію, що є суперпозицією одноелектронних функцій, наприклад так званий детермінант Слетера

$$\Psi_e(\mathbf{q}_1, \mathbf{q}_2, \dots) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \Psi_1(\mathbf{q}_1) & \Psi_1(\mathbf{q}_2) & \dots \\ \Psi_2(\mathbf{q}_1) & \Psi_2(\mathbf{q}_2) & \dots \\ \dots & \dots & \dots \end{vmatrix}, \quad (2.43)$$

де  $N$  – число електронів,  $\mathbf{q}_i = (x_i, y_i, z_i, s_i)$ . Як випливає із властивостей визначників

$$\Psi_e(\dots, \mathbf{q}_i, \dots, \mathbf{q}_j, \dots) = -\Psi_e(\dots, \mathbf{q}_j, \dots, \mathbf{q}_i, \dots). \quad (2.44)$$

Рівняння типу (2.42) у термінах функцій (2.43) називається *рівнянням Хартрі–Фока*. Таким чином, наближення Хартрі при визначенні самоузгодженого поля не враховує статистику електронів. Це наближення було розвинене Фоком за допомогою застосування специфічної симетрії хвильових функцій. Такий підхід має назву метода Хартрі–Фока та враховує, що електрони характеризуються напівцілим спіном, підлягають статистиці Фермі та принципу Паулі. Продемонструємо відмінність між наближеннями Хартрі та Хартрі–Фока на прикладі системи, що складається із двох електронів. У цьому випадку хвильова функція записується у вигляді

$$\Phi = \frac{1}{\sqrt{2}} \{ \varphi_1(\mathbf{r}_1, s_1) \varphi_2(\mathbf{r}_2, s_2) - \varphi_1(\mathbf{r}_2, s_2) \varphi_2(\mathbf{r}_1, s_1) \}, \quad (2.45)$$

і середня енергія системи задається формулою

$$\begin{aligned} E &= \sum_{s_1, s_2} \int d\mathbf{r}_1 d\mathbf{r}_2 \Phi^* H \Phi = \sum_{s_1, s_2} \sum_{i=1}^2 \int d\mathbf{r}_i \varphi_i^*(\mathbf{r}_1, s_1) H_i \varphi_i(\mathbf{r}_1, s_1) + \\ &+ \frac{1}{2} \sum_{s_1, s_2} \sum_{i, j=1}^2 \int d\mathbf{r}_1 d\mathbf{r}_2 |\varphi_i(\mathbf{r}_1, s_1)|^2 \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|^2} |\varphi_j(\mathbf{r}_2, s_2)|^2 - \\ &- \frac{1}{2} \sum_{s_1, s_2} \sum_{i, j=1}^2 \int d\mathbf{r}_1 d\mathbf{r}_2 \varphi_i^*(\mathbf{r}_1, s_1) \varphi_j(\mathbf{r}_1, s_1) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|^2} \varphi_i(\mathbf{r}_2, s_2) \varphi_j^*(\mathbf{r}_2, s_2). \end{aligned} \quad (2.46)$$

Головна відмінність між рівняннями Хартрі та Хартрі–Фока полягає у тому, що коли другий доданок у (2.42) при обчисленні енергії дає усереднену енергію кулонівської взаємодії електронів

$$E_C = \sum_i \int d\mathbf{r}_i \psi_i^*(\mathbf{r}_i) \left[ \frac{1}{2} \sum_{j \neq i} \int |\psi_j(\mathbf{r}_j)|^2 \frac{e(ed\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \psi_i(\mathbf{r}_i), \quad (2.47)$$

то у наближенні Хартрі–Фока виникає третій доданок, який називається *обмінною енергією*. У доданку сума залежить від спінового стану електронів, тобто ця взаємодія відбиває суто квантову природу і не має

класичних електростатичних аналогів. Обмінні інтеграли можна зіставити із кулонівською взаємодією з обмінною щільністю зарядів

$$\rho_{ij}(\mathbf{r}_1) = -e\varphi_i^*(\mathbf{r}_1)\varphi_j(\mathbf{r}_1) \text{ та } \rho_{ij}(\mathbf{r}_2) = -e\varphi_i^*(\mathbf{r}_2)\varphi_j(\mathbf{r}_2).$$

Оскільки ці величини можуть бути комплексними, то про щільність зарядів може йтися дуже умовно. Однак, можна казати, що два електрони із паралельними спінами, які частково знаходяться в  $i$ -му та  $j$ -му станах і взаємодіють, нібито обмінюються місцями.

Раніше ми вже з'ясували, що завдяки трансляційній інваріантності у кристалі хвильова функція має задовольняти співвідношенню  $\Psi(\mathbf{r} + \mathbf{b}) = e^{i\mathbf{k}\cdot\mathbf{b}}\Psi(\mathbf{r})$ , тобто вектор  $\mathbf{k}$  є характеристикою даної хвильової функції  $\Psi(\mathbf{r}) = \Psi_{\mathbf{k}}(\mathbf{r})$ . Оскільки  $H\Psi_{\mathbf{k}}(\mathbf{r}) = E\Psi_{\mathbf{k}}(\mathbf{r})$ , то енергія має бути функцією хвильового вектора, тобто  $E = E(\mathbf{k})$ . Така залежність називається законом дисперсії електронів, і її визначення для конкретних систем є важливим завданням фізики твердого тіла. Але при визначенні закону дисперсії виникає питання про врахування того факту, що електрон у кристалі поводить себе як квантова частинка, тобто виявляє властивості як хвилі (хвильова функція залежить від хвильового вектора), так і частинки. А рух частинки має фундаментальну характеристику – імпульс. Як бути із визначенням імпульсу за руху електрона в істотно неоднорідному середовищі – кристалічній ґратці?

### ■ 2.3. Квазіімпульс

Із курсу теоретичної механіки відомо, що закон збереження імпульсу є наслідком трансляційної інваріантності простору (у полі постійної потенційної енергії), закон збереження енергії випливає з однорідності часу, закон збереження моменту імпульсу – з ізотропії простору (виконується у полі із центральною симетрією). Усі ці приклади є наслідками загальної *теорема Емі Нетер*: для будь-якої фізичної системи, рівняння руху якої можуть бути отримані із варіаційного принципу, кожному однопараметричному перетворенню, що залишає варіаційний функціонал інваріантом, відповідає один закон збереження. У квантовій теорії аналог цього твердження формулюється так: величина, що не залежить від часу явно, зберігається, якщо її оператор комутує з оператором Гамільтона системи.

Аналогічно у кристалі трансляційній симетрії потенціального поля ґратки відповідає деяка величина, що зберігається за руху частинки у цьому полі. Така величина називається *квазіімпульсом*. Нехай квазіімпульсу відповідає оператор  $\hat{P}$ . Закон збереження квазіімпульсу формально запишемо як

$$\hat{P}H - H\hat{P} = 0. \quad (2.48)$$

Звідси випливає, що за руху електрона у полі ґратки власні функції операторів  $\hat{P}$  та  $H$  мають збігатися, отже має існувати зв'язок між їхніми власними значеннями, тобто функціональна залежність

$$E = E(P). \quad (2.49)$$

Нехай гамільтоніан електрона в кристалі має вигляд  $\hat{h} = t + U$ , де  $T$  – кінетична енергія електрона,  $U$  – кристалічний потенціал. Оскільки оператор  $-i\hbar\nabla$  не комує з гамільтоніаном (тобто імпульс не зберігається у полі ґратки)

$$\frac{d\mathbf{p}}{dt} = \frac{1}{i\hbar} [\mathbf{p}H - H\mathbf{p}] = -\nabla U = \mathbf{F}, \quad (2.50)$$

то оператор квазіімпульсу не може бути представлений як оператор імпульсу, домножений на деяку функцію. У той же час при  $\nabla U \rightarrow 0$  оператори імпульсу та квазіімпульсу мають збігатися, тобто є сенс шукати оператор квазіімпульсу у вигляді

$$\hat{\mathbf{P}} = -i\hbar\nabla + i\hbar\gamma(\mathbf{r}), \quad (2.51)$$

де  $\gamma \rightarrow 0$ , якщо  $\nabla U \rightarrow 0$ . Зважаючи на те, що хвильова функція  $\Psi_{\mathbf{k}}(\mathbf{r})$  є власною функцією оператора квазіімпульсу

$$\hat{\mathbf{P}}\Psi_{\mathbf{k}}(\mathbf{r}) = \mathbf{P}\Psi_{\mathbf{k}}(\mathbf{r}), \quad (2.52)$$

і, крім того, її можна представити як блохівську  $e^{i\mathbf{k}\cdot\mathbf{r}}\phi_{\mathbf{k}}(\mathbf{r})$ , маємо

$$\begin{aligned} -i\hbar i\mathbf{k}\Psi_{\mathbf{k}}(\mathbf{r}) + e^{i\mathbf{k}\cdot\mathbf{r}}(-i\hbar\nabla\phi_{\mathbf{k}}(\mathbf{r})) + i\hbar\gamma\Psi_{\mathbf{k}}(\mathbf{r}) = \\ = \hbar\mathbf{k}\Psi_{\mathbf{k}}(\mathbf{r}) + i\hbar[\gamma - \nabla\ln\phi_{\mathbf{k}}(\mathbf{r})]\Psi_{\mathbf{k}}(\mathbf{r}) = \mathbf{P}\Psi_{\mathbf{k}}(\mathbf{r}). \end{aligned} \quad (2.53)$$

Із цієї формули можна записати

$$\mathbf{P} = \hbar\mathbf{k}, \quad (2.54)$$

якщо  $\gamma = \nabla\ln\phi_{\mathbf{k}}(\mathbf{r}). \quad (2.55)$

Таким чином, оператор квазіімпульсу має вигляд

$$\hat{\mathbf{P}} = -i\hbar\nabla + i\hbar(\nabla\ln\phi_{\mathbf{k}}(\mathbf{r})), \quad (2.56)$$

тобто від потенціалу ґратки  $U$  цей оператор залежить завдяки залежності хвильової функції  $\phi_{\mathbf{k}}(\mathbf{r})$  від потенціалу. Нехай тепер кристал знаходиться у деякому полі  $V(\mathbf{r})$ , тоді гамільтоніан системи

$$H' = H + V(\mathbf{r}). \quad (2.57)$$

Оскільки оператор квазіімпульсу комує з оператором  $H_0$ , а другий доданок у (2.56), що є оператором множення, – з  $V(\mathbf{r})$ , то

$$\frac{d\hat{\mathbf{P}}}{dt} = [H', \hat{\mathbf{P}}] = [V(\mathbf{r}), \hat{\mathbf{P}}] = \frac{1}{i\hbar} \{\hat{\mathbf{P}}V - V\hat{\mathbf{P}}\} = -\nabla V(\mathbf{r}) = \mathbf{F}_{ext}, \quad (2.58)$$

тобто квазіімпульс змінюється під дією неперіодичної частини потенціального поля. Це означає, що за будь-яких порушень ідеальної структури ґратки має відбуватися розсіювання електронних хвиль. Наприклад, такими порушеннями структури є теплові коливання ґратки. Розсіювання на них електронів є однією із причин виникнення електричного опору.

Знайдемо тепер рівняння, якому задовольняє хвильова функція  $\varphi_{\mathbf{k}}(\mathbf{r})$ . Для цього зауважимо, що

$$\nabla\psi_{\mathbf{k}} = i\mathbf{k}\psi_{\mathbf{k}} + e^{i\mathbf{k}\cdot\mathbf{r}}\nabla\varphi_{\mathbf{k}}; \quad \Delta\psi_{\mathbf{k}} = i\mathbf{k}[i\mathbf{k}\psi_{\mathbf{k}} + e^{i\mathbf{k}\cdot\mathbf{r}}\nabla\varphi_{\mathbf{k}}] + i\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{r}}\nabla\varphi_{\mathbf{k}} + e^{i\mathbf{k}\cdot\mathbf{r}}\Delta\varphi_{\mathbf{k}}.$$

Рівняння Шредингера тепер запишемо як

$$-\frac{\hbar^2}{2m}[-\mathbf{k}^2\varphi_{\mathbf{k}} + 2i\mathbf{k}\nabla\varphi_{\mathbf{k}} + \Delta\varphi_{\mathbf{k}}] + U\varphi_{\mathbf{k}} = E(\mathbf{k})\varphi_{\mathbf{k}} \quad (2.59)$$

або

$$-\frac{\hbar^2}{2m}\Delta\varphi_{\mathbf{k}} + \left[ \frac{\hbar^2\mathbf{k}^2}{2m} - i\frac{\hbar^2}{m}\mathbf{k}\nabla + U \right] \varphi_{\mathbf{k}} = E(\mathbf{k})\varphi_{\mathbf{k}}. \quad (2.60)$$

Оскільки енергія є дійсною функцією хвильового вектора

$$E^*(\mathbf{k}) = E(\mathbf{k}), \quad (2.62)$$

то рівняння Шредингера для комплексно-спряженої функції  $\varphi_{\mathbf{k}}^*(\mathbf{r})$  запишемо як

$$H^*\psi_{\mathbf{k}}^* = E\psi_{\mathbf{k}}^* \quad (2.62)$$

або

$$-\frac{\hbar^2}{2m}\Delta\varphi_{\mathbf{k}}^* + \left[ \frac{\hbar^2\mathbf{k}^2}{2m} + i\frac{\hbar^2}{m}\mathbf{k}\nabla + U \right] \varphi_{\mathbf{k}}^* = E(\mathbf{k})\varphi_{\mathbf{k}}^*. \quad (2.63)$$

Якщо тепер записати (2.60) для від'ємних значень хвильового вектора, то отримаємо

$$-\frac{\hbar^2}{2m}\Delta\varphi_{-\mathbf{k}} + \left[ \frac{\hbar^2\mathbf{k}^2}{2m} + i\frac{\hbar^2}{m}\mathbf{k}\nabla + U \right] \varphi_{-\mathbf{k}} = E(-\mathbf{k})\varphi_{-\mathbf{k}}. \quad (2.64)$$

Легко побачити, що (2.64) і (2.63) збігаються для хвильової функції, яка задовольняє умові Шварца

$$(\varphi_{\mathbf{k}}(\mathbf{r}))^* = \varphi_{-\mathbf{k}}(\mathbf{r}). \quad (2.65)$$

Таким чином, можна стверджувати, що енергія є парною функцією хвильового вектора

$$E(\mathbf{k}) = E(-\mathbf{k}), \quad (2.66)$$

тобто поблизу точки  $\mathbf{k} = 0$  енергія залежить у крайньому випадку від  $k^2$ . У просторі хвильових векторів рівняння

$$E(\mathbf{k}) = \text{const} \quad (2.67)$$

визначає поверхні постійної енергії – ізоенергетичні поверхні. Топологія ізоенергетичних поверхонь визначає багато властивостей твердих тіл, зокрема напівпровідників.

## ■ 2.4. Ізоенергетичні поверхні

Як ми бачили у попередніх розділах, велику роль у визначенні електронних станів твердого тіла відіграє обернений простір, тобто простір хвильових векторів, що характеризують хвильові функції електронів у періодичному потенціалі. Зручною топологічною характеристикою електронних станів кристала є поверхні сталої енергії або ізоенергетичні поверхні, що визначаються в оберненому просторі із рівняння  $E(\mathbf{k}) = E$ . У напівпровідниках такі поверхні відіграють велику роль поблизу точок екстремумів енергії.

Нехай  $\mathbf{k}^0$  є точкою екстремуму енергії. Зрозуміло, що завдяки симетрії кристалічної ґратки таких точок екстремуму може бути кілька. Наприклад, ми бачили, що енергія є парною функцією хвильового вектора, тобто точка  $\mathbf{k}^0$  також буде точкою екстремуму. Оскільки парність енергії має виявлятися і для компонентів хвильового вектора, тобто

$$E(k_x, k_y, k_z) = E(-k_x, k_y, k_z) = E(-k_x, -k_x, k_z) = \dots, \quad (2.68)$$

то можна стверджувати, що кількість екстремальних точок визначатиметься групою симетрії ґратки. Наприклад, у кубічній ґратці, що має 24 елементи симетрії, екстремальних еквівалентних точок може бути 24. Розкладаючи енергію як функцію хвильового вектора у просторі хвильових векторів (в оберненому просторі), маємо<sup>1</sup>

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \left. \frac{\partial E}{\partial k_i} \right|_{\mathbf{k}_0} (k_i - k_i^0) + \frac{1}{2} \left. \frac{\partial^2 E}{\partial k_i \partial k_j} \right|_{\mathbf{k}_0} (k_i - k_i^0)(k_j - k_j^0) + \dots \quad (2.69)$$

Оскільки  $\mathbf{k}^0$  є точкою екстремуму, то перша похідна енергії за хвильовим вектором має дорівнювати нулю, тобто

<sup>1</sup> Тут і далі використовується правило, згідно з яким за повторюваним індексом відбувається підсумовування.

$$\left. \frac{\partial E}{\partial k_i} \right|_{\mathbf{k}_0} = 0,$$

і ряд (2.69) починається з квадратичних членів

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{1}{2} \left. \frac{\partial^2 E}{\partial k_i \partial k_j} \right|_{\mathbf{k}_0} (k_i - k_i^0)(k_j - k_j^0) + \frac{1}{6} \left. \frac{\partial^3 E}{\partial k_i \partial k_j \partial k_l} \right|_{\mathbf{k}_0} (k_i - k_i^0)(k_j - k_j^0)(k_l - k_l^0) + \dots \quad (2.70)$$

Задля залишення у розкладенні (2.70) тільки членів другого порядку необхідно виконання нерівності:

$$\left. \frac{\partial^2 E}{\partial k_i \partial k_j} \right|_{\mathbf{k}_0} \gg \frac{1}{3} \left. \frac{\partial^3 E}{\partial k_i \partial k_j \partial k_l} \right|_{\mathbf{k}_0} (k_l - k_l^0). \quad (2.71)$$

Зрозуміло, що завжди можна назвати малу область поблизу екстремальної точки для виконання умови (2.71). Тензор другого рангу за допомогою вибору осей координат завжди можна привести до діагонального вигляду, тобто недіагональні компоненти тензора дорівнюватимуть нулю

$$\left. \frac{\partial^2 E}{\partial k_i \partial k_j} \right|_{\mathbf{k}^0, i \neq j} = 0. \quad (2.72)$$

У цьому випадку рівняння ізоенергетичної поверхні має вигляд

$$E = E(\mathbf{k}_0) + \frac{1}{2} \left. \frac{\partial^2 E}{\partial k_i^2} \right|_{\mathbf{k}_0} (k_i - k_i^0)^2. \quad (2.73)$$

Оскільки розкладання (2.73) проводиться поблизу точки екстремуму, то всі три коефіцієнти

$$\left. \frac{\partial^2 E}{\partial k_x^2} \right|_{\mathbf{k}_0}, \left. \frac{\partial^2 E}{\partial k_y^2} \right|_{\mathbf{k}_0} \text{ та } \left. \frac{\partial^2 E}{\partial k_z^2} \right|_{\mathbf{k}_0}$$

мають однаковий знак. Отже, у даному наближенні ізоенергетичні поверхні є *еліпсоїдами*.

Домножимо другий доданок у (2.73) на  $\hbar^2/\hbar^2$  і візьмемо до уваги, що  $\hbar k_i = P_i$ , де  $P_i$  –  $i$ -й компонент квазіімпульсу. У результаті другий доданок у правій частині рівняння матиме вигляд

$$\frac{1}{2} \left. \frac{\partial^2 E}{\partial P_i^2} \right|_{\mathbf{P}_0} (P_i - P_i^0)^2. \quad (2.74)$$

Оскільки розмірність квазіімпульсу збігається із розмірністю імпульсу, то похідна



$$\left. \frac{1}{2} \frac{\partial^2 E}{\partial P_i^2} \right|_{\mathbf{P}_0}$$

має розмірність оберненої маси. Вважаючи систему ізотропною, введемо позначення

$$\left. \frac{1}{2} \frac{\partial^2 E}{\partial P_i^2} \right|_{\mathbf{P}_0} = \frac{1}{m^*}. \quad (2.75)$$

Оскільки залежність енергії від хвильового вектора є такою самою, що й залежність енергії від квазіімпульсу, то отримаємо із (2.73)

$$E = E(\mathbf{P}_0) + \frac{(\mathbf{P} - \mathbf{P}_0)^2}{2m^*}. \quad (2.76)$$

Таким чином, ми бачимо, що рух електрона поблизу екстремуму енергії аналогічний руху вільної частинки із масою  $m^*$ , яку називають ефективною. Такий підхід, що базується на ідеї розгляду руху електрона у кристалі зі значеннями енергії поблизу екстремумів як руху деякої вільної частинки із масою  $m^*$ , широко застосовується у фізиці напівпровідників.

Запишемо (2.76) у компонентах квазіімпульсу

$$E = E_0 + \frac{(P_x - P_x^0)^2}{2m^*} + \frac{(P_y - P_y^0)^2}{2m^*} + \frac{(P_z - P_z^0)^2}{2m^*} \quad (2.77)$$

і введемо позначення 
$$R = \sqrt{2m^*(E - E_0)}. \quad (2.78)$$

У результаті отримаємо, що для ізотропного кристала ізоенергетичними поверхнями в оберненому просторі є сфери радіуса  $R$ , оскільки рівняння ізоенергетичної поверхні є рівнянням сфери:

$$(P_x - P_x^0)^2 + (P_y - P_y^0)^2 + (P_z - P_z^0)^2 = R^2. \quad (2.79)$$

Зрозуміло, що у загальному випадку розкладення енергії у ряд за квазіімпульсом поблизу точки екстремуму повинно мати вигляд

$$E = E(\mathbf{P}_0) + \left. \frac{1}{2} \frac{\partial^2 E}{\partial P_i \partial P_j} \right|_{\mathbf{P}_0} (P_i - P_i^0)(P_j - P_j^0). \quad (2.80)$$

Введемо позначення 
$$(m_{ij}^*)^{-1} = \left. \frac{\partial^2 E}{\partial P_i \partial P_j} \right|_{\mathbf{P}_0}. \quad (2.81)$$

Аналогічно до (2.75), можна стверджувати, що розмірність компонент тензора (2.81) є розмірністю оберненої маси. Належним вибором системи координат тензор (2.81) можна привести до діагонального вигляду

$$(m_{ij}^*)^{-1} = \begin{pmatrix} \frac{1}{m_x^*} & 0 & 0 \\ 0 & \frac{1}{m_y^*} & 0 \\ 0 & 0 & \frac{1}{m_z^*} \end{pmatrix}, \quad (2.82)$$

де введено позначення  $(m_i^*)^{-1} = (m_{ii}^*)^{-1} = \frac{\partial^2 E}{\partial P_i^2}$ . (2.83)

Тензор, обернений до тензора (2.82), називається тензором ефективної маси. Тепер рівняння ізоенергетичної поверхні можна записати так

$$E = E_0 + \frac{(P_x - P_x^0)^2}{2m_x^*} + \frac{(P_y - P_y^0)^2}{2m_y^*} + \frac{(P_z - P_z^0)^2}{2m_z^*}. \quad (2.84)$$

Зауважимо, що у випадку, коли екстремум є мінімумом, ефективна маса є позитивною, а коли екстремум є максимумом – ефективна маса є негативною. Якщо ввести позначення

$$a = \sqrt{2m_x^*(E - E_0)}, \quad b = \sqrt{2m_y^*(E - E_0)}, \quad c = \sqrt{2m_z^*(E - E_0)}, \quad (2.85)$$

то (2.84) можна записати як рівняння еліпсоїда у канонічній формі

$$\frac{(P_x - P_x^0)^2}{a^2} + \frac{(P_y - P_y^0)^2}{b^2} + \frac{(P_z - P_z^0)^2}{c^2} = 1, \quad (2.86)$$

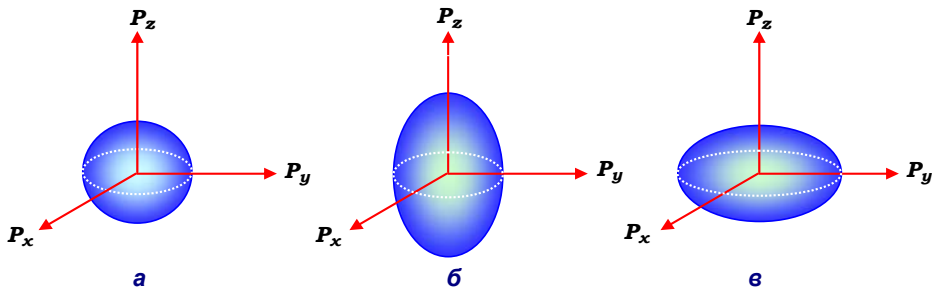


Рис. 2.1. Ізоенергетичні поверхні поблизу точки екстремуму:

$a - m_x^* = m_y^* = m_z^*$ ; б, в -  $m_x^* = m_y^* = m_{\perp}^*$ ; б -  $m_{\perp}^* < m_z^*$ ; в -  $m_{\perp}^* > m_z^*$

із напіввісями  $a$ ,  $b$  та  $c$  (рис. 2.1). Видно, що довжина напіввісі пропорційна квадратному кореню із відповідного компонента ефективної маси. Якщо всі три компоненти тензора ефективної маси однакові, то ізоенергетичні поверхні являють собою сфери, що описуються рівнянням (2.77); якщо однаковими є два (напр.,  $m_x^* = m_y^* = m_{\perp}^*$ ) компоненти

тензора ефективної маси, то ізоенергетичні поверхні являють собою еліпсоїди обертання навкруги вісі  $OZ$  і описуються формулою

$$E = E_0 + \frac{(P_x - P_x^*)^2 + (P_y - P_y^*)^2}{2m_{\perp}^*} + \frac{(P_z - P_z^*)^2}{2m_{\parallel}^*}, \quad (2.87)$$

де  $m_{\perp}^*$  – поперечна, а  $m_{\parallel}^*$  – повздовжня ефективна маса. Якщо  $m_{\parallel}^* > m_{\perp}^*$ , то еліпсоїд буде витягнутий уздовж вісі обертання, причому тим більше, чим більшим є відношення  $m_{\parallel}^*/m_{\perp}^*$ . У зворотному випадку  $m_{\parallel}^* < m_{\perp}^*$  еліпсоїд буде стиснутий уздовж вісі обертання. Необхідно пам'ятати, що за сильнішого розрізнення компонент тензора ефективної маси сильніше витягнутий буде еліпсоїд. За мірою віддалення від екстремальної точки необхідно враховувати більшу кількість членів у розкладанні енергії. Еліпсоїди енергії деформуватимуться у все складніші фігури.

Розкладання (2.73) є основою потужного методу визначення енергетичної структури електронних станів у напівпровідниках – методу ефективної маси, який детально буде розглянуто у розд. 4.

## ■ 2.5. Задачі

1. Доведіть на прикладі детермінанта Слетера для двох і трьох електронів, що хвильова функція (2.80) нормована на одиницю.

2. Побудуйте ізоенергетичні поверхні (2.73) для різних співвідношень між компонентами тензора ефективної маси: а)  $m_y^* = m_{\perp}^*/2$ ; б)  $m_z^* = 3m_{\perp}^*$ ; в)  $m_z^* = m_{\perp}^*/4$ .

3. Обчисліть швидкості електронів у різних напрямках для ізоенергетичних поверхонь, що задані у попередній задачі. Чому швидкість електрона залежить від напрямку руху?

## Список літератури

1. Харрисон У. Теория твердого тела. – М.: Мир, 1972.
2. Анималу А. Квантовая теория кристаллических твердых тел. – М.: Мир, 1981.
3. Киреев П.С. Физика полупроводников. – М.: Высшая шк., 1975.
4. Цидильковский И.М. Электроны и дырки в полупроводниках. – М.: Наука, 1972.
5. Таулес Д. Квантовая механика многих частиц. – М.: Мир, 1975.
6. Марч Н., Янг У., Сампантхар С. Проблема многих тел в квантовой механике. – М.: Мир, 1969.
7. Ли М.-Ф. Modern Semiconductor Quantum Physics. – Singapore: World Scientific, 1995.



## Розділ 3

# ЕЛЕКТРОН У ПЕРІОДИЧНОМУ ПОТЕНЦІАЛІ

Визначення зв'язку між квазіімпульсом частинки та її енергією є основною задачею фізики твердого тіла, для розв'язання якої необхідно розглянути рух електронів у періодичному потенціалі атомів, що утворюють тверде тіло. Ми детально розглянули основні ідеї та наближення, що можуть бути використані при розв'язанні цієї складної задачі, – адіабатичне та одноелектронне наближення. У подальшому ми завжди працюватимемо у рамках цих наближень.

### ■ 3.1. Рівняння Шредингера для електрона у періодичному потенціалі

Характерною рисою руху електрона у кристалі є наявність періодичного кристалічного потенціалу, що діє на електрон. При цьому, зважаючи на складність задачі, зазвичай використовують одноелектронне та адіабатичне наближення, у рамках яких задача визначення закону дисперсії електрона у кристалі зводиться до розв'язку рівняння Шредингера для електрона у періодичному полі

$$\left( -\frac{\hbar^2}{2m} \Delta + U(\mathbf{r}) \right) \Psi(\mathbf{r}) = E \Psi(\mathbf{r}) \quad (3.1)$$

із кристалічним потенціалом, що задовольняє умові трансляційної інваріантності на вектор ґратки

$$U(\mathbf{r}) = U(\mathbf{r} + \mathbf{n}). \quad (3.2)$$

Зрозуміло, що у загальному випадку аналітичний розв'язок рівняння (3.1) неможливий, оскільки самоузгоджене визначення кристалічного потенціалу вимагає додаткових наближень. Як відомо з курсу квантової механіки, такі рівняння можна розв'язувати за допомогою методів квантової теорії збурень. При цьому важливою обставиною є той стан електрона, який береться за нульовий крок ітераційної процедури теорії збурень. Ясно, що у випадку руху електронів у полі кристалічного потенціалу за нульового наближення можна вибирати як стани електронів в ізольованому атомі, так і стани вільного електрона. За послідовного виконання ітераційної процедури обидва підходи нададуть однаковий результат. Перший випадок відповідає наближенню квазізв'язаного, а другий – квазівільного електрона. Розглянемо детально обидва наближення для впевненості у тому, що вони дають однаково якісну картину поведінки електрона у кристалі.

## ■ 3.2. Електрон у полі потенціалу типу гребінця Дірака

Для вивчення особливостей поведінки електронної підсистеми у полі періодичного кристалічного потенціалу спочатку розглянемо найпростіші моделі одновимірного кристала. Варто зауважити, що така модель не є просто плодом теоретичної уяви. У наш час широко вивчаються екзотичні колись системи –

металоорганічні координаційні сполуки, зокрема фталоціаніни металів. Органічні молекули, що мають просторову форму рамки, усередині якої знаходиться атом металу (рис. 3.1), на поверхні води збираються у молекулярні пакети довжиною у десятки молекул. При цьому атоми металу всередині молекул збираються в одновимірний ланцюжок. Від геометрії конкретної молекули, з яких утворюються такі стопки, залежить форма потенціалу одновимірної системи. Якщо атоми металу віддають електрон зовнішньої орбіталі в електронну підсистему, то електрони внутрішніх оболонок можуть бути

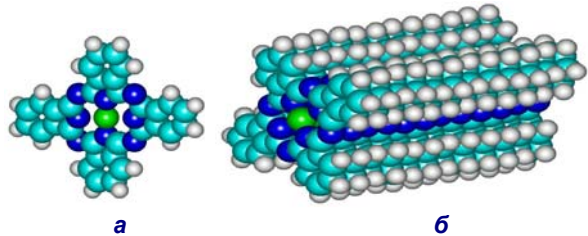


Рис. 3.1. Молекула (а) та утворення молекулярного пакета (б) фталоціаніну металу.

У молекулі атом металу знаходиться всередині органічного оточення та тримається за рахунок т. зв. координаційного зв'язку. При утворенні молекулярного пакета атоми металу розташовані на відстані порядку сталої ґратки в об'ємному зразку та утворюють одновимірний кристал

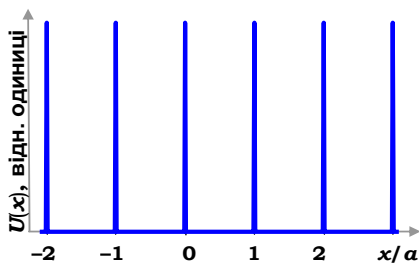


Рис. 3.2. Потенціальний гребінець Дірака

сильно зв'язані на іоні та значно екрануватимуть кулонівське поле іона. Іншими словами, потенціал такого кристалічного вузла може бути представлений як значно локалізований репульсивний потенціал із великою висотою бар'єру, у граничному випадку – потенціал, що має дельта-подібну форму (рис. 3.2).

Таким чином, можна розглянути електронні стани у періодичному потенціалі типу гребінця Дірака

$$U(x) = \frac{\hbar^2 Q}{m} \sum_{n=-\infty}^{\infty} \delta(x + na), \quad (3.3)$$

де  $Q$  – константа розмірності довжини,  $m$  – маса електрона,  $a$  – стала ґратки. Оскільки потенціал дорівнює нулю при  $x \neq na$ , то в інтервалі  $0 < x < a$  хвильову функцію можна представити у вигляді

$$\psi(x) = Ae^{i\beta x} + Be^{-i\beta x} \equiv \psi_1(x). \quad (3.4)$$

Тут 
$$\beta = \sqrt{2mE/\hbar^2}, \quad (3.5)$$

а  $E$  – енергія електрона. Якщо змінна  $x$  належить інтервалу  $(0, a)$ , то й змінна  $x-a$  з інтервалу  $a < x < 2a$  – інтервалу  $(0, a)$ . Згідно із теоремою Блоха в інтервалі  $a < x < 2a$  хвильову функцію можна записати як

$$\psi(x) = \psi_1(x-a)e^{ika} \equiv \psi_2(x). \quad (3.6)$$

Для знаходження коефіцієнтів  $A$  та  $B$ , що визначають хвильову функцію електрона, функції  $\psi_1(x)$  і  $\psi_2(x)$  та їхні похідні мають бути "зшиті" у точці  $x = a$ . Зважаючи на те, що потенціал є розривним, похідні хвильових функцій можна обчислити через інтегрування рівняння Шредингера в інфінітезимально малому інтервалі поблизу точки  $a$ ,  $a - \varepsilon < x < a + \varepsilon$ ,  $\varepsilon \rightarrow 0$

$$\int_{a-\varepsilon}^{a+\varepsilon} \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} - \frac{\hbar^2 Q}{m} \delta(x-a) \right] \Psi(x) dx = \int_{a-\varepsilon}^{a+\varepsilon} dx E \Psi(x). \quad (3.7)$$

Оскільки хвильова функція є неперервною, то права частина (3.7) дорівнює нулю. При інтегруванні лівої частини отримуємо рівняння

$$-\frac{\hbar^2}{2m} \frac{d\psi}{dx} \Big|_{x=a} - \frac{\hbar^2 Q}{m} \psi(a) = 0. \quad (3.8)$$

Звідси маємо 
$$\psi'(a) = -2Q\psi(a). \quad (3.9)$$

Таким чином, умови "зшивання" хвильових функцій та їхніх похідних запишемо у вигляді

$$\psi_1(a) = \psi_2(a), \quad \psi'_1(a) = \psi'_2(a) = -2Q\psi_1(a). \quad (3.10)$$

Після підстановки до умов "зшивання" (3.10) виразів для хвильових функцій (3.4) і (3.6) отримуємо

$$f_+ A + f_- B = 0, \quad (3.11)$$

$$\left( f_+ - 2i \frac{Q}{\beta} e^{i\beta a} \right) A - \left( f_- + 2i \frac{Q}{\beta} e^{-i\beta a} \right) B = 0, \quad (3.12)$$

де позначено 
$$f_{\pm} = e^{\pm i\beta a} - e^{ika}. \quad (3.13)$$

Для існування ненульового розв'язку системи лінійних рівнянь (3.11)–(3.12), тобто розв'язку задачі у формі (3.4), необхідне виконання умови

$$f_+ \left( f_- + 2i \frac{Q}{\beta} e^{-i\beta a} \right) = -f_- \left( f_+ - 2i \frac{Q}{\beta} e^{i\beta a} \right). \quad (3.14)$$

Спростуючи (3.14), отримуємо умову існування ненульових розв'язків системи (3.12)

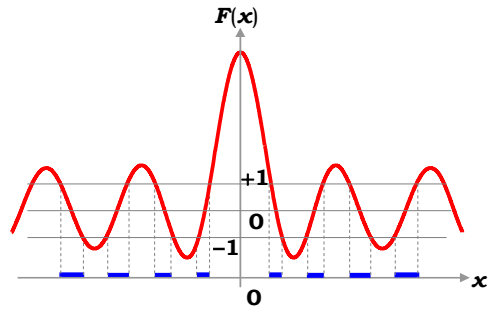
$$\cos \beta a + (Q/\beta) \sin \beta a = \cos ka. \quad (3.15)$$

Оскільки  $\beta a = \sqrt{2mEa^2/\hbar^2}$ , то рівняння (3.15) дає зв'язок між дозволе-ними значеннями енергії  $E$  та хвильовим вектором  $k$ . З іншого боку, у правій частині рівняння стоїть косинус, який за модулем не може перевищувати одиницю, тому із рівняння можна визначити області дозволених значень енергії

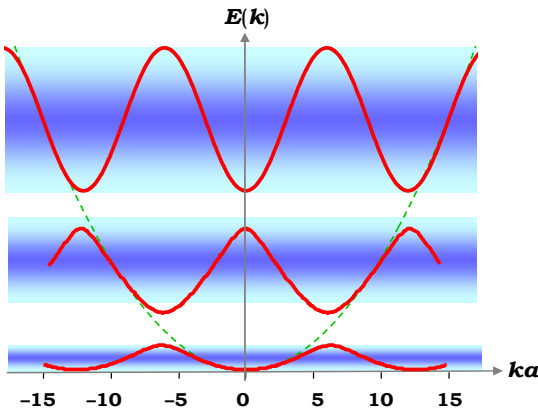
$$\left| \cos\beta a + \frac{Q}{\beta} \sin\beta a \right| \leq 1. \tag{3.16}$$

Поведінку функції  $F(x) = \cos x + \frac{Qa}{x} \sin x$

та області дозволених значень енергії подано на рис. 3.3. Зауважимо, що області дозволених значень енергії, які називаються енергетичними зонами, завдяки наявності фактора  $1/x$  перед синусом збільшуються зі зростанням  $x$ . Отриманий результат показує, що за руху електрона у полі періодичного потенціалу виникають області (зони) дозволених значень енергії, відокремлені одна від одної зонами заборонених значень.



**Рис. 3.3.** Функція  $F(x)$  та області дозволених значень енергії (позначено синім). Між цими областями розташовані області заборонених значень енергії



**Рис. 3.4.** Закон дисперсії електрона в полі потенціалу типу гребінця Дірака

Характерний вигляд закону дисперсії у полі потенціалу типу гребінця Дірака подано на рис. 3.4. Граничний випадок  $Q \rightarrow 0$  дає закон дисперсії вільного електрона (штрихова лінія). Видно, що друга зона ширша за першу, а третя – ширша за другу. Таким чином, розсіювання електрона на періодично розташованих сильних репульсивних точкових розсіювачах приводить до того, що закон дисперсії вільного електрона  $E = \hbar^2 k^2 / 2m$  набуває

значних змін: виникають зони дозволених і заборонених значень енергії. У кожній дозволений зоні закон дисперсії є періодичним із таким періодом, що визначається просторовою структурою потенціалу.



### ■ 3.3. Наближення майже вільного електрона

Для розуміння багатьох процесів у напівпровідниках велике значення мають наближені методи обчислення  $E(\mathbf{k})$ . Для цього зазвичай застосовується теорія збурень. Існують два наближення при обчисленні енергії електрона у періодичному полі, що відрізняються вибором нульового кроку. Вважаючи вільну частинку нульовим кроком у методі послідовних наближень і розглядаючи періодичне поле як збурення, отримуємо наближення (модель) *квазівільного електрона*. Розглянемо цей підхід детальніше. Нехай гамільтоніаном електрона у періодичному кристалічному потенціалі  $U(\mathbf{r})$  є

$$H = T + U(\mathbf{r}). \quad (3.17)$$

Вважатимемо, що гамільтоніан незбуреної системи  $H_0 = T$ , де  $T$  – гамільтоніан вільної частинки, тобто

$$-\frac{\hbar^2}{2m} \Delta \psi^0(\mathbf{r}) = E^0 \psi^0(\mathbf{r}). \quad (3.18)$$

Розв'язком (3.18) є хвиля де-Бройля

$$\psi_{\mathbf{k}}^0(\mathbf{r}) = A e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (3.19)$$

а власними значеннями – неперервний спектр енергії

$$E^0(\mathbf{k}) = \frac{\hbar^2 k^2}{2m}. \quad (3.20)$$

Вважаючи хвильову функцію нормованою на одиницю, при інтегруванні за деякою обмеженою областю  $G$  можна визначити нормувальний коефіцієнт  $A$

$$\psi_{\mathbf{k}}^0 = \frac{1}{\sqrt{G}} e^{i\mathbf{k}\cdot\mathbf{r}}, \quad \rightarrow \quad \frac{1}{G} \int d\mathbf{r} e^{-i\mathbf{k}'\cdot\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{r}} = \delta_{\mathbf{k},\mathbf{k}'}. \quad (3.21)$$

Як відомо із теорії збурень, для обчислення енергії і хвильової функції у методі послідовних збурень необхідно знати матричні елементи

$$U_{\mathbf{k}'\mathbf{k}} = \int d\mathbf{r} (\psi_{\mathbf{k}'}^0)^* U(\mathbf{r}) \psi_{\mathbf{k}}^0. \quad (3.22)$$

Обчислимо їх, попередньо розклавши потенціал у ряд Фур'є

$$\begin{aligned} U_{\mathbf{k}'\mathbf{k}} &= \frac{1}{G} \int d\mathbf{r} e^{-i\mathbf{k}'\cdot\mathbf{r}} \sum_{\mathbf{b}} U_{\mathbf{b}} e^{i\mathbf{b}\cdot\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{\mathbf{b}} U_{\mathbf{b}} \frac{1}{G} \int d\mathbf{r} e^{i(\mathbf{k}+\mathbf{b}-\mathbf{k}')\cdot\mathbf{r}} = \\ &= \sum_{\mathbf{b}} U_{\mathbf{b}} \delta_{\mathbf{k}',\mathbf{k}+\mathbf{b}} = \begin{cases} 0, & \mathbf{k}' \neq \mathbf{k} + \mathbf{b} \\ U_{\mathbf{b}}, & \mathbf{k}' = \mathbf{k} + \mathbf{b} \end{cases}, \end{aligned} \quad (3.23)$$

де вектор оберненої ґратки  $\mathbf{b} = l_1 \mathbf{b}_1 + l_2 \mathbf{b}_2 + l_3 \mathbf{b}_3$ ,  $\mathbf{b}_i = \mathbf{e}_i (2\pi/a_i)$ ,  $a_i$  – стала ґратки вздовж  $i$ -ї вісі. Тобто ненульовими матричними елементами є діагональні матричні елементи, що дорівнюють Фур'є-трансформантам

потенціалу, і вектор  $\mathbf{k}' - \mathbf{k}$  дорівнює будь-якому вектору оберненої ґратки. Поправка у першому порядку теорії збурень дорівнює усередненій за область  $G$  потенціальної енергії і не залежить від імпульсу

$$E^{(1)}(\mathbf{k}) = \langle U \rangle = \frac{1}{G} \int d\mathbf{r} U(\mathbf{r}). \quad (3.24)$$

У другому порядку теорії збурень поправка до енергії задається величиною

$$E^{(2)}(\mathbf{k}) = \sum_{\mathbf{k}'} \frac{|U_{\mathbf{k}'\mathbf{k}}|^2}{E^0(\mathbf{k}) - E^0(\mathbf{k}')}. \quad (3.25)$$

Оскільки поправка першого порядку викликає лише зміщення всіх рівнів енергії на однакову величину  $\langle U \rangle$ , то можна вибрати початок відліку енергії так, щоб  $\langle U \rangle = 0$ . Тоді, з урахуванням (3.23), енергія електрона задається формулою

$$E(\mathbf{k}) = E^0(\mathbf{k}) + \sum_{\mathbf{b}} \frac{|U_{\mathbf{b}}|^2}{E^0(\mathbf{k}) - E^0(\mathbf{k} + \mathbf{b})}. \quad (3.26)$$

Зрозуміло, що коли різниця  $|E^0(\mathbf{k}) - E^0(\mathbf{k} + \mathbf{b})|$  набагато більша за  $|U_{\mathbf{b}}|^2$ , то поправка до енергії несуттєва, і збурення незначно впливає на хвильову функцію. Нехай за деякого значення  $\mathbf{b} = -\mathbf{g}$ ,  $E^0(\mathbf{k}) \rightarrow E^0(\mathbf{k} - \mathbf{g})$ . Отже знаменник у (3.26) стає інфінітезимально малою величиною. При цьому умови застосування методу збурень не виконуються, і формула (3.26) втрачає сенс. За точної рівності  $E^0(\mathbf{k}) = E^0(\mathbf{k} - \mathbf{g})$  стан  $E^0(\mathbf{k})$  є виродженим, оскільки одному й тому самому значенню енергії відповідають дві різні хвильові функції  $\psi_{\mathbf{k}}^0(\mathbf{r})$  і  $\psi_{\mathbf{k}-\mathbf{g}}^0(\mathbf{r})$ . Це означає, що вже у нульовому наближенні необхідно враховувати обидві хвильові функції, тобто виродження. Згідно із теорією збурень вироджених станів хвильова функція у нульовому наближенні повинна мати вигляд

$$\psi^0(\mathbf{r}) = \alpha \psi_{\mathbf{k}}^0(\mathbf{r}) + \beta \psi_{\mathbf{k}-\mathbf{g}}^0(\mathbf{r}). \quad (3.27)$$

Підставимо її до рівняння

$$(H_0 + U)\psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (3.28)$$

і врахуємо, що функції  $\psi_{\mathbf{k}}^0(\mathbf{r})$  і  $\psi_{\mathbf{k}-\mathbf{g}}^0(\mathbf{r})$  є власними функціями незбуреного гамільтоніана. Використовуючи позначення  $\psi_{\mathbf{k}}^0(\mathbf{r}) = \psi_1$ ,  $\psi_{\mathbf{k}-\mathbf{g}}^0(\mathbf{r}) = \psi_2$ ,  $E^0(\mathbf{k}) = E_1$  та  $E^0(\mathbf{k} - \mathbf{g}) = E_2$ , запишемо (3.28) у вигляді рівняння

$$\alpha E_1 \psi_1 + \beta E_2 \psi_2 + U(\alpha \psi_1 + \beta \psi_2) = E(\alpha \psi_1 + \beta \psi_2). \quad (3.29)$$

Подіємо на нього інтегральними операторами

$$I_1 = \int d\mathbf{r} \psi_1^* \dots \quad \text{і} \quad I_2 = \int d\mathbf{r} \psi_2^* \dots$$

У результаті маємо два рівняння, що складають однорідну систему лінійних рівнянь

$$\begin{cases} (E_1 + U_{11} - E)\alpha + U_{12}\beta = 0, \\ U_{21}\alpha + (E_2 + U_{22} - E)\beta = 0. \end{cases} \quad (3.30)$$

Для нетривіального розв'язку цієї системи її детермінант має дорівнювати нулю

$$\begin{vmatrix} (E_1 + U_{11} - E) & U_{12} \\ U_{21} & (E_2 + U_{22} - E) \end{vmatrix} = 0. \quad (3.31)$$

Враховуючи, що ми вибрали такий початок відліку енергії, за якого  $U_{11} = U_{22} = \langle U \rangle = 0$ , із (3.31) маємо

$$E = \frac{1}{2}(E_1 + E_2) \pm \sqrt{\frac{1}{4}(E_1 - E_2)^2 + U_{12}U_{21}}. \quad (3.32)$$

Видно, що у точках, де  $E_1 = E_2$ , енергія має розрив. Для точок

$$E(\mathbf{k}) = E^0(\mathbf{k}) \pm |U_{\mathbf{g}}|. \quad (3.33)$$

Величина розриву дорівнює  $2|U_{\mathbf{g}}|$ . Умовою розриву є  $E_1 = E_2$  або в явному вигляді

$$|\mathbf{k}| = |\mathbf{k} - \mathbf{g}|. \quad (3.34)$$

Це рівняння можна переписати у зручнішому вигляді

$$\mathbf{g} \cdot (\mathbf{k} - \mathbf{g}/2) = 0, \quad (3.35)$$

тобто вектори  $\mathbf{g}$  та  $\mathbf{k} - \mathbf{g}/2$  є ортогональними. Ясно, що умові (3.34) задовольняють усі хвильові вектори  $\mathbf{k}$ , проведені із початку  $\mathbf{g}$  на площину, що перпендикулярна вектору  $\mathbf{g}$  і проходить через його середину. Вектори  $\mathbf{k} - \mathbf{g}/2$  лежать у цій площині та перпендикулярні вектору  $\mathbf{g}$ .

На рис. 3.5 за сутністю подано побудову Евальда, що описує когерентну бреггівську дифракцію електронів із хвильовим вектором  $\mathbf{k}$  у напрямку  $\mathbf{k}' = \mathbf{k} - \mathbf{g}$ . Для знаходження всіх можливих векторів  $\mathbf{k}$ , які задовольняють (3.34), необхідно задати всі вектори  $\mathbf{g}$ . Для цього треба побудувати обернену ґратку із базисом  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ , і зі всіх вузлів ґратки провести вектори  $\mathbf{g}$  у початок координат, де  $\mathbf{k} = 0$ . Сукупність площин, що проходять через середини векторів  $\mathbf{g}$  та є нормальними до них, визначають усі ті вектори  $\mathbf{k}$ , які задовольняють умовам (3.34)  $|\mathbf{k} - \mathbf{g}| = |\mathbf{k}|$ . Але це рівняння визначає зону Бриллюена для кри-

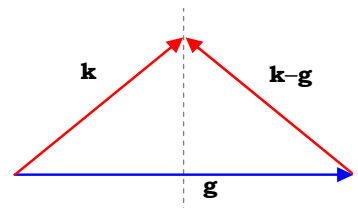


Рис. 3.5. Побудова Евальда

стала (воно еквівалентне рівнянню  $E^0(\mathbf{k} - \mathbf{g}) = E^0(\mathbf{k})$ ), тобто розрив енергії відбувається на границі зони Бриллюена. Вважатимемо, що  $|\mathbf{k}| = 2\pi/\lambda$  ( $\lambda$  – довжина хвилі де Бройля), а  $|\mathbf{g}| = 2\pi n/d$ , де  $n$  – ціле число,  $d$  – відстань між сусідніми атомними площинами. Якщо ввести кут  $\vartheta$  між вектором  $\mathbf{k}$  і нормаллю до атомної площини (що паралельна вектору  $\mathbf{g}$ ), то можна переписати (3.35) як

$$\frac{4\pi^2}{\lambda} \frac{n}{d} \cos\vartheta - 2\pi^2 \frac{n^2}{d^2} = 0.$$

$$\text{Звідси маємо} \quad 2d \cos\vartheta = n\lambda \quad (3.36)$$

– умову Вульфа–Брегга для інтерференції рентгенівських променів, якщо  $\vartheta$  є кутом падіння. Таким чином, виродження спостерігатиметься для станів (що лежать на границях зони Бриллюена), для яких має спостерігатися бреггівське відбиття. Отже причиною виникнення енергетичної щілини є інтерференція між падаючою та відбитою хвилями де Бройля. У напрямках, нормальних до атомних площин, виникають стоячі хвилі. Біжучі хвилі можуть існувати лише вздовж атомних площин. Зрозуміло, що збурення буде суттєвим і в деяких областях поблизу граней зони Бриллюена. Нехай  $\mathbf{k}_0$  – хвильовий вектор на грані зони Бриллюена:  $\mathbf{g} \cdot (\mathbf{k}_0 - \mathbf{g}/2) = 0$  і

$$E(\mathbf{k}_0) = \frac{\hbar^2 \mathbf{k}_0^2}{2m} \pm |U_{\mathbf{g}}|. \quad (3.37)$$

Запишемо (3.32) у вигляді

$$E(\mathbf{k}) = \frac{\hbar^2 \mathbf{k}^2}{2m} + \frac{\hbar^2}{2m} \left[ \mathbf{g} \cdot \left( \mathbf{k} - \frac{\mathbf{g}}{2} \right) \right] \pm \sqrt{\left( \frac{\hbar^2}{2m} \right)^2 \left[ \mathbf{g} \cdot \left( \mathbf{k} - \frac{\mathbf{g}}{2} \right) \right]^2 + |U_{\mathbf{g}}|^2}. \quad (3.38)$$

Закон дисперсії (3.38) обчислено у так званому наближенні однієї бреггівської площини, тобто враховано розсіювання електрона тільки однією бреггівською площиною, що характеризується хвильовим вектором  $\mathbf{g}/2$ . Зрозуміло, що фізично коректнішою є модель, де враховується розсіювання електрона парою бреггівських площин, які характеризуються хвильовими векторами  $\pm \mathbf{g}/2$ . У такому випадку закон дисперсії буде симетричним щодо заміни знака  $\mathbf{k}$  на  $-\mathbf{k}$ , а це відповідає вимозі симетрії моделі.

Приклад закону дисперсії подано на рис. 3.6, де через  $E^{(+)}$  та  $E^{(-)}$  позначено енергії, що відповідають дну другої зони та стелі першої зони, відповідно. При цьому перша зона відділена від другої щілиною  $E^{(+)} - E^{(-)} = 2|U_{\mathbf{g}}|$ . Закон дисперсії вільного електрона показано на рисунку пунктирною кривою. Порівнюючи рис. 3.6 із законом дисперсії

(рис. 3.4), бачимо, що модель гребінця Дірака адекватна методу квазівільного електрона. Проаналізуємо тепер поведінку закону дисперсії електрона в кристалі поблизу особливих точок (3.37). Зазначимо: якщо вважати, що за законом (3.24) електронні стани деякого твердого тіла розподілені таким чином, що всі рівні з енергіями  $E < E^{(-)} = \hbar^2 \kappa_0^2 (2m)^{-1} - |U_{\mathbf{g}}|$  заповнені, а з інтервалу  $E > E^{(+)} = \hbar^2 \kappa_0^2 (2m)^{-1} + |U_{\mathbf{g}}|$  порожні, то  $E^{(+)}$  та  $E^{(-)}$  відповідають дну зони провідності та стелі валентної зони. Дійсно, для того, щоб електрони могли брати участь в електричному струмі, необхідно, щоб вони мали змогу переходити під дією електричного поля на рівні із більшою енергією. Але всі енергетичні стани у зоні  $E < E^{(-)}$  заповнені електронами, які у вільних атомах були валентними (звідси й назва – валентна зона), і для участі у транспортних процесах електрон має перейти в найнижчий вільний стан у зоні  $E > E^{(+)}$ . Отже тільки у верхній зоні можуть існувати стани електронів, що беруть участь у створенні струму. Тому така зона називається зоною провідності. Перепишемо (3.38) у такому вигляді

$$E(\mathbf{k}) - E(\mathbf{k}_0) = \frac{\hbar^2 \mathbf{k}^2}{2m} - \frac{\hbar^2 \mathbf{k}_0^2}{2m} + \frac{\hbar^2}{2m} \left[ \mathbf{g} \cdot \left( \mathbf{k} - \frac{\mathbf{g}}{2} \right) \right] \pm \sqrt{\left( \frac{\hbar^2}{2m} \right)^2 \left[ \mathbf{g} \cdot \left( \mathbf{k} - \frac{\mathbf{g}}{2} \right) \right]^2 + |U_{\mathbf{g}}|^2} \mp |U_{\mathbf{g}}| = f(\mathbf{k}). \quad (3.39)$$

Розкладемо цей вираз у ряд Тейлора поблизу точки  $\mathbf{k}_0$ , де  $f(\mathbf{k}_0) = 0$ . Для цього треба знайти похідну  $dE(\mathbf{k})/d\mathbf{k}$ .

$$\frac{dE(\mathbf{k})}{d\mathbf{k}} = \frac{\hbar^2 \mathbf{k}}{m} + \frac{\hbar^2 \mathbf{g}}{2m} \pm \left( \frac{\hbar^2}{2m} \right)^2 \frac{[\mathbf{g} \cdot (\mathbf{k} - \mathbf{g}/2)] \mathbf{g}}{\sqrt{\left( \frac{\hbar^2}{2m} \right)^2 [\mathbf{g} \cdot (\mathbf{k} - \mathbf{g}/2)] + |U_{\mathbf{g}}|^2}}. \quad (3.40)$$

Швидкість електронів  $\mathbf{v} = \frac{1}{\hbar} \frac{dE(\mathbf{k})}{d\mathbf{k}}$ . Для станів, що лежать далеко від граней зони Бриллюена, швидкість майже дорівнює швидкості вільного електрона

$$\mathbf{v} = \frac{\hbar \mathbf{k}}{m}, \quad (3.41)$$

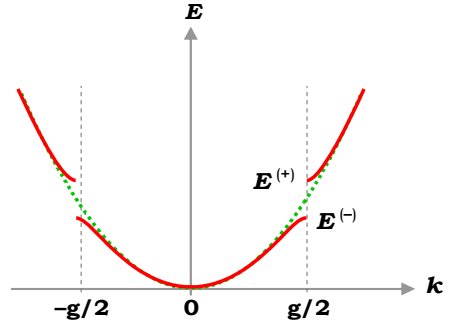


Рис. 3.6. Закон дисперсії електрона у моделі двох брегівських площин

тобто має такий самий напрямок, що й хвильовий вектор. Але, згідно із (3.40), для станів, де  $|\mathbf{k}| \approx |\mathbf{g}|$ , напрямок швидкості не збігається з напрямком хвильового вектора. А на самій границі зони Бриллюена, де  $\mathbf{k} = \mathbf{k}_0$  і  $\mathbf{g} \cdot (\mathbf{k}_0 - \mathbf{g}/2) = 0$ , маємо для швидкості електронів

$$\mathbf{v} = \frac{\hbar}{m} (\mathbf{k}_0 - \mathbf{g}/2). \quad (3.42)$$

Тобто вектор швидкості лежить у площині, перпендикулярній хвильовому вектору. Оскільки похідна  $dE/d\mathbf{k}|_{\mathbf{k}_0}$  лежить у площині грані зони Бриллюена, то ізоенергетичні поверхні мають бути нормальними до границь зони Бриллюена, де вони зазнають розриву. Для станів  $\mathbf{k}_0 = \mathbf{g}/2$  швидкість дорівнює нулю, відтак для станів на границі зони Бриллюена рух у напрямку, нормальному до атомної площини, неможливий. Для таких точок похідна  $dE/d\mathbf{k}|_{\mathbf{k}_0}$  також обертається на нуль. Занулення першої похідної означає, що у точках енергія має екстремум. Використовуючи (3.40), знайдемо другу похідну енергії за хвильовим вектором

$$\frac{d^2 E(\mathbf{k})}{d\mathbf{k}^2} = \frac{\hbar^2}{m} \pm \left( \frac{\hbar^2}{2m} \right)^2 \frac{\mathbf{g}^2 |U_{\mathbf{g}}|^2}{\left[ \left( \frac{\hbar^2}{2m} \right)^2 [\mathbf{g} \cdot (\mathbf{k} - \mathbf{g}/2)]^2 + |U_{\mathbf{g}}|^2 \right]^{3/2}}. \quad (3.43)$$

Введемо позначення 
$$\frac{d^2 E(\mathbf{k})}{\hbar^2 d\mathbf{k}^2} = \frac{1}{m^*}, \quad (3.44)$$

звідки випливає, що  $m^*$  – стала розмірності маси. Тепер із (3.43) та (3.44) знайдемо, що для точки екстремуму

$$\frac{1}{m^*} = \frac{1}{m} \left( 1 \pm \frac{\hbar^2 \mathbf{g}^2}{4m |U_{\mathbf{g}}|} \right) \quad (3.45)$$

або 
$$m^* = \frac{m}{1 \pm \frac{\hbar^2 \mathbf{g}^2}{4m |U_{\mathbf{g}}|}}. \quad (3.46)$$

Оскільки  $\hbar^2 \mathbf{g}^2 / (2m) > 2 |U_{\mathbf{g}}|$ , то можна стверджувати, що позитивне значення величини  $m^*$  відповідає енергії  $\hbar^2 \mathbf{k}_0^2 / 2m + |U_{\mathbf{g}}|$ , тобто є мінімумом енергії, а негативне значення  $m^*$  – енергії  $\hbar^2 \mathbf{k}_0^2 / 2m - |U_{\mathbf{g}}|$ : у цій точці спостерігається максимум.

Обговоримо тепер методи представлення залежності енергії зонного електрона від хвильового вектора. Для спрощення розглядатимемо одновимірний випадок, коли можливе лише дворазове виродження рівнів. За відсутності взаємодії між електроном і кристалічним потен-

ціалом залежність енергії електрона від хвильового вектора  $E(k)$  має вигляд парабол, які зсунуті на вектор оберненої ґратки  $g = 2\pi/a$  ( $a$  – стала ґратки), що відповідає станам вільного електрона. У першому порядку за слабким одновимірним періодичним потенціалом  $U_g$  крива  $E(k)$  лишається незмінною всюди, за винятком околів точок  $k = ng/2$ ,  $n = \pm 1, \pm 2, \dots$ , що відповідають бреггівським площинам у тривимірному випадку. У цих точках взаємодія призведе до зняття виродження та виникнення енергетичної щілини величиною  $2U_g$ . У результаті із врахуванням усіх бреггівських площин ми отримаємо закон дисперсії електрона у схемі розширених зон (рис. 3.7 а), де кожна гілка спектра  $E_i^{(+)}$  та  $E_i^{(-)}$  визначена тільки з одного боку  $i$ -тої бреггівської площини.

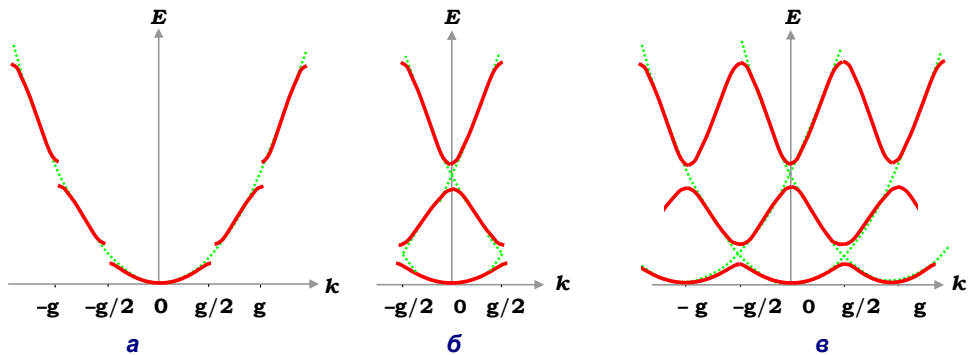


Рис. 3.7. Представлення енергії зонного електрона (метод квазівільного електрона) у схемах зон: а – розширених; б – приведених; в – повторюваних

Якщо необхідно задавати всі енергетичні стани хвильовими векторами з першої зони Бриллюєна, то можна перенести всі частини кривих (рис. 3.7 а) у першу зону, зміщуючи їх на відповідний вектор оберненої ґратки. Як результат отримуємо закон дисперсії електрона у схемі приведених зон (рис. 3.7 б). Іноді при представленні закону дисперсії зонного електрона важливо підкреслити періодичність опису в  $k$ -просторі. Тоді криві, представлені на рис. 3.7 б, можна періодично продовжити на всю  $k$ -вісь, трансляючи їх на вектори оберненої ґратки й отримати представлення у схемі повторюваних зон (рис. 3.7 в). Усі ці представлення енергії електрона є еквівалентними та широко використовуються у фізиці твердого тіла, зокрема фізиці напівпровідників.

### ■ 3.4. Утворення енергетичних зон

Запишемо тепер розкладення у ряд Тейлора енергії (3.38) поблизу довільної точки  $\mathbf{k}_0$ , що лежить у площині грані зони Бриллюена. Використовуючи (3.42), (3.44) та (3.45), знайдемо

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{1}{2} \frac{\hbar^2}{m} \left( 1 \pm \frac{\hbar^2 \mathbf{g}^2}{4m |U_{\mathbf{g}}|} \right) (\mathbf{k} - \mathbf{k}_0)^2 + \dots, \quad (3.47)$$

де 
$$E(\mathbf{k}_0) = \frac{\hbar^2 \mathbf{k}_0^2}{2m} \pm |U_{\mathbf{g}}|. \quad (3.48)$$

Енергія (3.47) має дві гілки, що відповідають знакам "+" та "-" у (3.47) і (3.48). Позначимо гілки  $E^{(+)}\mathbf{k}$  та  $E^{(-)}\mathbf{k}$ , кожна з яких є неперервною на границі зони Бриллюена. Закон дисперсії поблизу границі зони Бриллюена подано на рис. 3.8 у так званій схемі повторюваних зон. Рівняння ж (3.48), взагалі кажучи, записані у схемі розширених зон. Це означає, що кожен із виразів можна застосовувати тільки по один бік грані зони: вираз для енергії  $E^{(-)}\mathbf{k}$  – із внутрішнього боку зони Бриллюена (область, де знаходиться початок координат  $\mathbf{k} = 0$ ), а  $E^{(+)}\mathbf{k}$  – із зовнішнього боку. Необхідно звернути увагу, що наявність розриву енергії у точках  $\mathbf{k}_0$  ще не означає розривів в енергетичному спектрі  $E(\mathbf{k})$ . Дійсно, для деяких точок біля границі зони Бриллюена із внутрішнього та зовнішнього її боків  $\mathbf{k}_1$  і  $\mathbf{k}_2$ , відповідно, які розташовані близько одна до одної і  $|\mathbf{k}_1| > |\mathbf{k}_2|$ , цілком може статись, що

$$\frac{\hbar^2 \mathbf{k}_1^2}{2m} - |U_{\mathbf{g}}| > \frac{\hbar^2 \mathbf{k}_2^2}{2m} + |U_{\mathbf{g}}|, \quad (3.49)$$

тобто енергетичні смуги  $E^{(+)}\mathbf{k}$  та  $E^{(-)}\mathbf{k}$  в околі грані зони Бриллюена перекриваються, отже енергетичний спектр електронів у цьому випадку є неперервним. Таким чином на прикладі моделі квазівільного електрона показано, що енергетичний спектр електронних станів у періодичному потенціалі набуває характерних рис: завдяки інтерференції хвиль де Бройля електронів (перерозсіювання на регулярному потенціалі) виникають широкі смуги енергії (енергетичних зон), розділені смугами недозволених станів. При цьому утворюються області

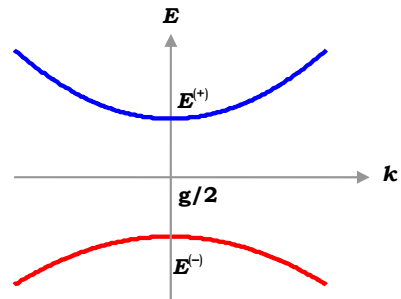


Рис. 3.8. Ілюстрація закону дисперсії поблизу границі Брюллуена



поблизу точок локальних мінімумів і максимумів енергії, де залежність енергії від квазіімпульсу є квадратичною. В оберненому просторі ці точки лежать на гранях зони Бриллюена та характеризуються нульовою груповою швидкістю електронів у напрямку відповідного вектора оберненої ґратки. Зауважимо, що розкладення типу (3.47) можна виконати поблизу екстремальних точок будь-якого закону дисперсії. Цей факт лежить в основі методу ефективної маси, який широко використовується у фізиці напівпровідників.

### ■ 3.5. Електрон у полі потенціалу Кроніґа–Пенні

Ми розглянули особливості формування зонної структури в моделі квазівільного електрона, де зовнішні електрони при утворенні кристала вважались слабозв'язаними з атомом. При розгляді одновимірної моделі такого кристала ми використали модель потенціалу гребінця Дірака, де зонна структура формувалась шляхом розсіювання майже вільних електронів на регулярній ґратці нескінченно вузьких і високих бар'єрів. Розглянемо тепер ситуацію, коли кристал утворюють атоми з електронами, що сильно зв'язані на атомах. Непоганою моделлю такої ситуації в одновимірному випадку можна вважати електрон у потенціалі Кроніґа–Пенні (рис. 3.9), що є набором еквідистантних прямокутних бар'єрів. Треба зауважити, що така модель добре описує властивості електронів у так званих надґратках – штучних системах гетеропереходів, що утворюють набір квантових ям та є одним з яскравих прикладів сучасних високих технологій квантово-розмірних структур.

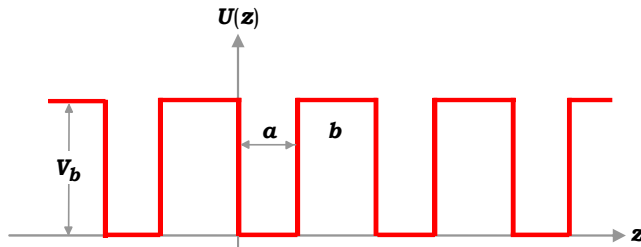


Рис. 3.9. Потенціал Кроніґа–Пенні

Зважаючи на те, що потенціальна енергія  $U_b(\mathbf{z})$  є періодичною функцією  $z$  із періодом  $d$ ,  $d = a + b$ , аналітично потенціал Кроніґа–Пенні можна записати у вигляді

$$V_b(z) + \sum_n V_b(z - nd), \quad (3.50)$$

$$\text{де } V_b(z) = \begin{cases} 0, & n(a+b) \leq z \leq n(a+b)+a \\ V_b, & n(a+b)-b \leq z \leq n(a+b) \end{cases}. \quad (3.51)$$

Розв'язки рівняння Шредингера відомі для кожного із доменів – це області ями та бар'єра. При цьому окремо необхідно розглядати випадки, коли енергія електрона перевищує або не перевищує величину потенціального бар'єра. Для значень енергії  $E < V_b$  в області  $a \leq z \leq a+b$  хвильова функція поперечного руху має вигляд двох еванесцентних хвиль, що затухають углиб бар'єра назустріч одна одній

$$\chi_1(z) = A \exp[k_b z] + B \exp[-k_b z]. \quad (3.52)$$

В області ями – це дві плоскі хвилі, що розповсюджуються назустріч одна одній, тобто при  $0 \leq z \leq a$  маємо

$$\chi_2(z) = C \exp[ik_w z] + D \exp[-ik_w z]. \quad (5.53)$$

У цих формулах позначено

$$k_w = \sqrt{2m\epsilon/\hbar^2} \quad \text{та} \quad k_b = \sqrt{2m(V_b - \epsilon)/\hbar^2}. \quad (3.54)$$

При цьому енергію можна записати як

$$\epsilon = V_b + \frac{\hbar^2 k_b^2}{2m} = \frac{\hbar^2 k_w^2}{2m}. \quad (3.55)$$

Використаємо той факт, що потенціал є періодичним за  $z$ . Оскільки гамільтоніан є періодичним, періодичними будуть і його власні функції та значення. Для довільного хвильового вектора  $q$  маємо

$$\chi_q(z + nd) = e^{iqnd} \chi_q(z) \quad (3.56)$$

$$\text{або} \quad \chi_q(z) = e^{iqz} u_q(z), \quad \text{де} \quad u_q(z+d) = u_q(z), \quad (3.57)$$

тобто доходимо висновку: розв'язок рівняння Шредингера задовольняє теоремі Блоха. Застосуємо тепер циклічні граничні умови до блохівської функції, вважаючи число  $N$  періодів потенціалу великим:

$$\chi_q(z + Nd) = \chi_q(z). \quad (3.58)$$

Для того, щоб співвідношення (3.58) виконувалось для хвильової функції (3.57), необхідно

$$qNd = 2p\pi, \quad (3.59)$$

де  $p$  – ціле число. Очевидно, що вектор  $q$  має бути дійсним. Крім того, існує  $N$  дозволених незалежних значень  $q$ . Відстанню між двома дозволеними величинами  $q \in 2\pi/Nd$ . Без обмежень у загальному підході можемо розглянути значення  $q$  з інтервалу  $[-\pi/d, +\pi/d]$ , який зазвичай

є першою зоною Бриллюена в оберненому просторі задачі. Запишемо ці умови у компактному вигляді додаткової хвильової функції:

$$\chi_3(z) = \chi_1(z+d)e^{iqd}. \quad (3.60)$$

Умови неперервності для хвильової функції  $\chi_q$  та її похідної  $d\chi_q/dz$  на межах розподілу мають вигляд

$$\begin{cases} \chi_1(a) - \chi_2(a) + [\chi_1'(a) - \chi_2'(a)]/k_b = 0 \\ \chi_2(0) - \chi_3(0) + [\chi_2'(0) - \chi_3'(0)]/k_b = 0 \\ \chi_1(a) - \chi_2(a) - [\chi_1'(a) - \chi_2'(a)]/k_b = 0 \\ \chi_2(0) - \chi_3(0) - [\chi_2'(0) - \chi_3'(0)]/k_b = 0 \end{cases}. \quad (3.61)$$

За явного запису цих умов отримаємо систему із чотирьох рівнянь для визначення невідомих коефіцієнтів  $A$ ,  $B$ ,  $C$ ,  $D$ , що формулюють хвильові функції (3.51) та (3.52)

$$\begin{cases} 2e^{k_b a} A + 0 - I_+ e^{ik_w a} C - I_- e^{-ik_w a} D = 0 \\ -2e^{k_b d + iqd} A + 0 + I_+ C + I_- D = 0 \\ 0 + 2e^{-k_b a} B - I_- e^{ik_w a} C - I_+ e^{-ik_w a} D = 0 \\ 0 - 2e^{-k_b d + iqd} B - I_- C + I_+ D = 0 \end{cases}. \quad (3.62)$$

або в матричному вигляді

$$\begin{pmatrix} 2e^{k_b a} & 0 & -I_+ e^{ik_w a} & -I_- e^{-ik_w a} \\ -2e^{k_b d + iqd} & 0 & I_+ & I_- \\ 0 & 2e^{-k_b a} & -I_- e^{ik_w a} & -I_+ e^{-ik_w a} \\ 0 & -2e^{-k_b d + iqd} & I_- & I_+ \end{pmatrix} \cdot \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = 0, \quad (3.63)$$

де введено позначення  $I_{\pm} = 1 \pm i(k_w/k_b)$ . Для нетривіального розв'язку цієї системи рівнянь необхідно виконання умови

$$\det \begin{pmatrix} e^{k_b a} & 0 & -I_+ e^{ik_w a} & -I_- e^{-ik_w a} \\ -e^{k_b d + iqd} & 0 & I_+ & I_- \\ 0 & e^{-k_b a} & -I_- e^{ik_w a} & -I_+ e^{-ik_w a} \\ 0 & -e^{-k_b d + iqd} & I_- & I_+ \end{pmatrix} = 0. \quad (3.64)$$

Звідси знайдемо, що дозволені блохівські стани мають задовольняти рівнянню ( $\xi = k_w/k_b$ )

$$\cos(qd) = \cos(k_w a) \operatorname{ch}(k_b b) - \frac{1}{2}(-\xi + 1/\xi) \sin(k_w a) \operatorname{sh}(k_b b). \quad (3.65)$$

У граничному випадку необмежено товстого бар'єру права частина (3.65) розходиться як  $\exp(k_b b)$ , коли коефіцієнт перед цією експонентою не є нулем, тобто не виконується умова

$$\cos(k_w a) - \frac{1}{2}(-\xi + 1/\xi) \sin(k_w a) = 0. \quad (3.66)$$

Це рівняння збігається із рівнянням зв'язаних станів відомої квантово-механічної задачі про ізольовану квантову яму

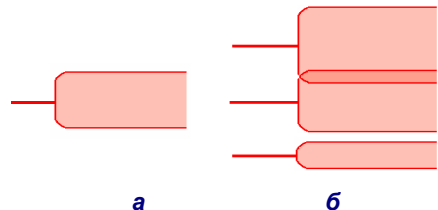
$$\cot(k_w a) = \frac{1}{2} \left( \frac{k_w}{k_b} - \frac{k_b}{k_w} \right). \quad (3.67)$$

Таким чином, зони за енергій електронів, менших від висоти потенціалу бар'єра  $V_b$ , з'являються як гібридизація зв'язаних станів ізольованих квантових ям завдяки тунельному зв'язку через бар'єр. Коли бар'єри є широкими (і/або високими), але кінцевими, то добуток  $k_b b$  є великим і зони – вузькими. Ширина зон пропорційна інтегралу передачі й швидко (експоненціально) зменшується зі зростанням  $b$  за фіксованого  $k_b$ . Функція  $F(\epsilon)$  у правій частині (3.65) значно змінюється зі зміною  $\epsilon$ , й енергетичний сегмент, де  $|F(\epsilon)| < 1$ , є вузьким. Для отримання наближених, але зручних виразів для дисперсійних співвідношень позначимо через  $\epsilon_1, \epsilon_2, \dots, \epsilon_n$  зв'язані стани ізольованої квантової ями. Розкладаючи  $F(\epsilon)$  поблизу деякої енергії  $\epsilon_s$ , наприклад  $\epsilon_j$  у першому порядку, за  $\epsilon - \epsilon_j$  маємо для дисперсійних співвідношень для  $j$ -ї зони

$$\epsilon_j(q) = \epsilon_j + s_j + 2t_j \cos(qd), \quad (3.68)$$

де 
$$s_j = \frac{-F(E_j)}{[F'(E)]_{E=E_j}}, \quad 2t_j = \frac{1}{[F'(E)]_{E=E_j}}. \quad (3.69)$$

Таким чином, (3.69) показує як формуються зони спектру електронів у потенціалі Кроніга-Пенні, рух яких у ямі локалізований. Наочно це подано на рис. 3.10: картина якісно збігається із тією, яку ми розглядали у розд. 2 при



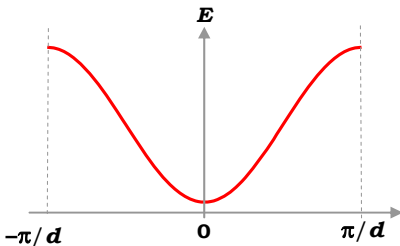
аналізі особливостей електронних станів за конденсації атомів у тверде тіло.

Розглянемо випадок, коли  $\epsilon > V_b$ . Необхідні умови існування ненульових розв'язків системи рівнянь (3.62) дають енергетичний спектр електронів у періодичному потенціалі

$$\cos(qd) = \cos(k_w a) \cos(k_b b) - \frac{1}{2} \left( \xi + \frac{1}{\xi} \right) \sin(k_w a) \sin(k_b b). \quad (3.70)$$

Це рівняння є дисперсійним і пов'язує енергію із хвильовим вектором  $q$ . Для даної величини  $q$  існує нескінченна множина розв'язків рівняння (3.70). Для розрізнення розв'язків індекс зони позначає енергію  $E$  та хвильову функцію  $\chi$ . Одновимірний блохівський стан залежить від двох квантових чисел  $n$  та  $q$  і позначається як  $|nq\rangle$ . Дозволені зони енергії розділяються щільною. Для енергій, що відповідають енергетичним щільностям, не існує дозволених станів. Величина щільності між двома зонами  $|nq\rangle$  і  $|n+1, q\rangle$  зменшується за зростання  $n$ . Для великих значень енергії ( $\epsilon \gg V_b$ ),  $k_b \sim k_w$  і (3.70) зводиться до  $k_w = q + 2\pi/d$ . Потенціал тепер майже не відчувається електронами із високими енергіями, тобто довжина хвилі електрона  $2\pi/k_w$  стає набагато меншою, ніж характерний період потенціалу  $d$ , і таким чином ефекти дифракції зникають. Зауважимо, що енергії, за яких  $k_w L = p\pi$ , завжди відповідають дозволеним станам. Дійсно, так само, як у випадку, коли енергія електрона в ямі менша від висоти бар'єра, при  $\epsilon > V_b$  можна розв'язок (3.70) записати у вигляді, аналогічному (3.68)

$$\epsilon_i(q) = \epsilon_i + s_i - 2|t_i| \cos qd. \quad (3.71)$$



**Рис. 3.11. Закон дисперсії  $j$ -ї зони електрона у потенціалі Кроніга-Пенні ( $E > V_b$ )**

**Рис. 3.10. Формування енергетичних зон у системі потенціальних ям: а – розщеплення одиночного рівня в окремій ямі; б – зони, утворені із різних енергетичних рівнів окремої ями**

Дисперсійне співвідношення (3.71) схематично подано на рис. 3.11. Дисперсійна крива  $\epsilon_i(q)$  виявляє невеликий зсув  $s_i$  щодо рівня  $\epsilon_i$ , який розвивається у неперервну зону  $2t_i \cos qd$  завширшки  $4|t_i|$ , причому ширина зони не залежить від числа ям (головне, щоб це число було досить великим). Останній доданок у (3.71) дає дисперсійний закон для  $i$ -ї енергетичної зони.

### ■ 3.6. Утворення енергетичних зон у теорії квазізв'язаного електрона

Розглянемо тепер підхід до визначення енергії електрона в кристалі, що за нульового наближення набуває станів електронів одиничного атома. Цей підхід схожий на той, що ми використали при розгляді формування зонної структури електрона у потенціалі Кроніга–Пенні. Для розв'язку рівняння Шредингера для електрона у періодичному полі

$$\left(-\frac{\hbar^2}{2m}\Delta + U(\mathbf{r})\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (3.72)$$

$$U(\mathbf{r}) = U(\mathbf{r} + \mathbf{n}) \quad (3.73)$$

за нульове наближення візьмемо стани електронів в ізолюваному атомі. Якщо гамільтоніан ізолюваного атома позначити як

$$H_a = -\frac{\hbar^2}{2m}\Delta + V_a(\mathbf{r}), \quad (3.74)$$

(тут  $V_a(\mathbf{r})$  – потенціальна енергія електрона в ізолюваному атомі), то хвильова функція електрона в атомі, що задовольняє рівнянню

$$H_a\psi_a(\mathbf{r}) = E_a\psi_a(\mathbf{r}), \quad (3.75)$$

відповідає деякому стану, локалізованому на атомі. Така хвильова функція, починаючи із деякої відстані від центра атома, різко (експоненціально) зменшується. Хвильову функцію електрона в кристалі утворимо з атомних хвильових функцій

$$\psi^0(\mathbf{r}) = \sum_{\mathbf{m}} c_{\mathbf{m}} \psi_a(\mathbf{r} - \mathbf{m}). \quad (3.76)$$

Тут підсумовування ведеться за вузлами ґратки, де розташовані ядра атомів. Коефіцієнти  $c_{\mathbf{m}}$  вибираються так, щоб хвильова функція (3.76) задовольняла умові теореми Блоха

$$\psi^0(\mathbf{r} + \mathbf{n}) = e^{i\mathbf{k}\cdot\mathbf{n}}\psi^0(\mathbf{r}), \quad (3.77)$$

тобто  $c_{\mathbf{m}} = e^{i\mathbf{k}\cdot\mathbf{m}}$ . Неважко впевнитись у цьому:

$$\psi^0(\mathbf{r} + \mathbf{n}) = \sum_{\mathbf{m}} e^{i\mathbf{k}\cdot\mathbf{m}} \psi_a(\mathbf{r} - \mathbf{m} + \mathbf{n}) = e^{i\mathbf{k}\cdot\mathbf{n}} \sum_{\mathbf{m}} e^{i\mathbf{k}\cdot(\mathbf{m}-\mathbf{n})} \psi_a(\mathbf{r} - (\mathbf{m} - \mathbf{n})).$$

Звідси, переходячи до підсумовування за новим набором хвильових векторів  $\mathbf{q} = \mathbf{m} - \mathbf{n}$ , отримаємо

$$\psi^0(\mathbf{r} + \mathbf{n}) = e^{i\mathbf{k}\cdot\mathbf{n}} \sum_{\mathbf{p}} e^{i\mathbf{k}\cdot\mathbf{p}} \psi_a(\mathbf{r} - \mathbf{p}).$$

Атомні хвильові функції нормовані на одиницю. Це означає, що їхня комбінація  $\psi^0$  не нормована на одиницю. Чому дорівнюватиме нормувальний коефіцієнт для цієї функції?

$$\int d\mathbf{r} (\psi^0(\mathbf{r}))^* \psi^0(\mathbf{r}) = \sum_{\mathbf{m}, \mathbf{n}} e^{i\mathbf{k}\cdot(\mathbf{m}-\mathbf{n})} \int d\mathbf{r} \psi_a^*(\mathbf{r} - \mathbf{m}) \psi_a(\mathbf{r} - \mathbf{n}). \quad (3.78)$$

Перейдемо до нових змінних:  $\mathbf{r}' = \mathbf{r} - \mathbf{m}$ ,  $d\mathbf{r}' = d\mathbf{r}$ . Тоді інтеграл у (3.78) запишемо у вигляді

$$\int d\mathbf{r} \psi_a^*(\mathbf{r} - \mathbf{m}) \psi_a(\mathbf{r} - \mathbf{n}) = \int d\mathbf{r}' \psi_a^*(\mathbf{r}') \psi_a(\mathbf{r}' - (\mathbf{n} - \mathbf{m})). \quad (3.79)$$

Він залежить не від позиції ядер  $\mathbf{m}$  та  $\mathbf{n}$ , а тільки від відстані між ними. За фіксованого значення  $\mathbf{m}$  величина  $\mathbf{q} = \mathbf{n} - \mathbf{m}$  пробігає ті самі значення, що й  $\mathbf{n}$ . Таким чином, із (3.78) із врахуванням (3.79) отримаємо

$$\sum_{\mathbf{m}, \mathbf{n}} e^{i\mathbf{k}\cdot(\mathbf{m}-\mathbf{n})} \int d\mathbf{r} \psi_a^*(\mathbf{r} - \mathbf{m}) \psi_a(\mathbf{r} - \mathbf{n}) = \sum_{\mathbf{m}} \sum_{\mathbf{q}} e^{i\mathbf{k}\cdot\mathbf{q}} D_{\mathbf{q}} = ND, \quad (3.80)$$

$$\text{де введено позначення } D_{\mathbf{q}} = \int d\mathbf{r}' \psi_a^*(\mathbf{r}') \psi_a(\mathbf{r}' - \mathbf{q}). \quad (3.81)$$

$$\text{Таким чином, маємо } \int d\mathbf{r} (\psi^0(\mathbf{r}))^* \psi^0(\mathbf{r}) = ND. \quad (3.82)$$

Представимо потенціальну енергію атома у ґратці як

$$U(\mathbf{r}) = \sum_{\mathbf{n}} V_a(\mathbf{r} - \mathbf{n}) + W(\mathbf{r}), \quad (3.83)$$

де  $W(\mathbf{r})$  – самоузгоджена частина потенціалу, що виникає завдяки взаємодії атомів при формуванні кристалічної ґратки. Вона є періодичною, але всюди обмеженою функцією.  $W(\mathbf{r})$  є негативною функцією на відстанях, коли атоми притягуються, і позитивною (репульсивною) – коли відштовхуються (напр., коли кристал сильно стискується під дією зовнішнього тиску).

Як відомо з курсу квантової механіки, хвильова функція в нульовому наближенні дозволяє обчислити енергію у першому наближенні.

$$\left( -\frac{\hbar^2}{2m} \Delta + \sum_{\mathbf{n}} V_a(\mathbf{r} - \mathbf{n}) + W(\mathbf{r}) \right) \psi^0(\mathbf{r}) = E^1 \psi^0(\mathbf{r}). \quad (3.84)$$

Діючи на це рівняння оператором  $I_a = \int d\mathbf{r} (\psi^0(\mathbf{r}))^* \dots$  і зважаючи на (3.82), отримаємо

$$\begin{aligned} & \int d\mathbf{r}(\psi^0(\mathbf{r}))^* \left[ -\frac{\hbar^2}{2m}\Delta + \sum_{\mathbf{n}} V_a(\mathbf{r}-\mathbf{n}) + W(\mathbf{r}) \right] \psi^0(\mathbf{r}) = \\ & = N \left\{ \sum_{\mathbf{p}} e^{i\mathbf{k}\cdot\mathbf{p}} \int d\mathbf{r}' \psi_a^*(\mathbf{r}') \left[ -\frac{\hbar^2}{2m}\Delta' + \sum_{\mathbf{n}\neq\mathbf{p}} V_a(\mathbf{r}'-\mathbf{n}) + W(\mathbf{r}') \right] \psi_a(\mathbf{r}'-\mathbf{p}) \right\}. \end{aligned} \quad (3.85)$$

Далі виділимо в сумі  $\sum_{\mathbf{n}} V_a(\mathbf{r}-\mathbf{n})$  доданок із  $\mathbf{n} = \mathbf{p}$ , і врахуємо, що

$$\begin{aligned} & \left[ -\frac{\hbar^2}{2m}\Delta' + V_a(\mathbf{r}'-\mathbf{p}) + \sum_{\mathbf{n}\neq\mathbf{p}} V_a(\mathbf{r}'-\mathbf{n}) + W(\mathbf{r}') \right] \psi_a(\mathbf{r}'-\mathbf{p}) = \\ & = E_a \psi_a(\mathbf{r}'-\mathbf{p}) + \left[ \sum_{\mathbf{n}\neq\mathbf{p}} V_a(\mathbf{r}'-\mathbf{n}) + W(\mathbf{r}') \right] \psi_a(\mathbf{r}'-\mathbf{p}). \end{aligned}$$

Зверніть увагу на те, що ліва частина у (3.85) є енергією електрона в кристалі, обчислена у першому порядку теорії наближень і помножена на нормувальний фактор  $ND$ , тобто її можна переписати у вигляді

$$E^{(1)} = D^{-1} \left\{ \sum_{\mathbf{p}} e^{i\mathbf{k}\cdot\mathbf{p}} \int d\mathbf{r}' \psi_a^*(\mathbf{r}') \left[ E_a + \sum_{\mathbf{n}\neq\mathbf{p}} V_a(\mathbf{r}'-\mathbf{n}) + W(\mathbf{r}') \right] \psi_a(\mathbf{r}'-\mathbf{p}) \right\}. \quad (3.86)$$

Як результат, звідси отримуємо

$$\begin{aligned} & E^{(1)} = E_a + D^{-1} \int d\mathbf{r}' \psi_a^*(\mathbf{r}') \left[ E_a + \sum_{\mathbf{n}\neq\mathbf{0}} V_a(\mathbf{r}'-\mathbf{n}) + W(\mathbf{r}') \right] \psi_a(\mathbf{r}') \\ & + D^{-1} \left\{ \sum_{\mathbf{p}\neq\mathbf{0}} e^{i\mathbf{k}\cdot\mathbf{p}} \int d\mathbf{r}' \psi_a^*(\mathbf{r}') \left[ E_a + \sum_{\mathbf{n}\neq\mathbf{p}} V_a(\mathbf{r}'-\mathbf{n}) + W(\mathbf{r}') \right] \psi_a(\mathbf{r}'-\mathbf{p}) \right\}. \end{aligned} \quad (3.87)$$

Введемо усереднену, самоузгоджену для атома потенціальну енергію електрона у полі всіх інших ядер

$$C = \int d\mathbf{r} (\psi_a(\mathbf{r}))^* \left[ \sum_{\mathbf{n}\neq\mathbf{0}} V_a(\mathbf{r}-\mathbf{n}) + W(\mathbf{r}) \right] \psi_a(\mathbf{r}) \quad (3.88)$$

та обмінну енергію, що зумовлена ненульовою ймовірністю знаходження електрона поблизу будь-якого довільного атома

$$A(\mathbf{p}) = \int d\mathbf{r} (\psi_a(\mathbf{r}))^* \left[ \sum_{\mathbf{n}\neq\mathbf{p}} V_a(\mathbf{r}-\mathbf{n}) + W(\mathbf{r}) \right] \psi_a(\mathbf{r}-\mathbf{p}). \quad (3.89)$$

Зрозуміло, що величина  $A(\mathbf{p})$  помітно відрізнятиметься від нуля лише за малих відстаней  $|\mathbf{p}|$  завдяки тому, що атомні хвильові функції швидко



(експоненційно) спадають при віддаленні від атома. Це означає, що в основному обмін відбувається між найближчими сусідами, а два довільних атоми обмінюються електронами через багаторазовий послідовний обмін між атомами-сусідами, тобто там, де локалізація електронів відбувається за рахунок "вільного" переміщення електронів від атома до атома завдяки ефекту обміну.

Таким чином, для енергії електрона в кристалі можна записати

$$E^1 = E_a + CD^{-1} + D^{-1} \sum_{\mathbf{p}} e^{i\mathbf{k}\cdot\mathbf{p}} A(\mathbf{p}). \quad (3.90)$$

Із формули видно, що в нульовому наближенні енергія електронів системи атомів дорівнює енергії електронів в ізольованому атомі. При зближенні атомів (коли взаємодія електронів з іншими ядрами та електронами інших атомів стає помітною) енергетичні рівні зсуваються на величину  $C$  і розщеплюються в смугу деякої ширини.

Проаналізуємо тепер вираз (3.90). Спочатку припустимо, що хвильова функція відповідає  $s$ -стану. Тоді обмінні інтеграли будуть ізотропними. Врахуємо обмін тільки між найближчими атомами. Оскільки всі вузли ґратки Браве є еквівалентними, то зрозуміло, що обмінний інтеграл матиме один і той самий вигляд для будь-якого вузла ґратки, тобто його можна вважати сталим. Тоді величину  $A$  можна винести за знак суми. У простій кубічній ґратці кожен атом оточений шістьма сусідами, що знаходяться на однакових відстанях від даного атома. У системі координат, вісі якої направлені вздовж ребер куба, для вузла  $p = 0$  сусідні вузли будуть розташовані на відстані  $+a$  та  $-a$ . Таким чином, сума в (3.90) легко обчислюється

$$\sum_i^6 e^{i\mathbf{k}\cdot\mathbf{p}} = e^{i\mathbf{k}_x a} + e^{-i\mathbf{k}_x a} + e^{i\mathbf{k}_y a} + e^{-i\mathbf{k}_y a} + e^{i\mathbf{k}_z a} + e^{-i\mathbf{k}_z a} = 2(\cos k_x a + \cos k_y a + \cos k_z a),$$

а енергія електрона у періодичному полі простої кубічної ґратки матиме вигляд

$$E^1 = E_a + C' + 2A'(\cos k_x a + \cos k_y a + \cos k_z a), \quad (3.91)$$

де  $C' = CD^{-1}$ ,  $A' = AD^{-1}$ . Тепер можна дійти деяких висновків, що безпосередньо випливають із (3.91):

- ♦ енергетичні рівні  $E_a$  ізольованого атома при утворенні кристалічної ґратки у результаті взаємодії між атомами зсуваються на величину  $C'$ ;
- ♦ атомний рівень у кристалічній ґратці розщеплюється у смугу (зону), усередині якої енергія електрона періодично залежить від компонентів хвильового вектора  $\mathbf{k}$ ;
- ♦ екстремальні значення (3.91), що мають місце при  $\cos k_j a = \pm 1$ , спостерігатимуться при  $k_j = (2\pi/a)n$ ,  $n = 0, 1, 2, \dots$  і дорівнюватимуть

$$E_+ = E_a + C' + 6A', \quad (3.92)$$

$$E_- = E_a + C' - 6A'. \quad (3.93)$$

Звідси видно, що ширина енергетичної зони

$$E_+ - E_- = 12A'; \quad (3.94)$$

- ♦ кожен енергетичний рівень ізольованого атома в кристалі розщеплюється у зону. Оскільки величина обмінного інтеграла залежить від перекривання хвильових функцій, то чим дужче перекриваються хвильові функції сусідніх атомів (тобто чим більшою є величина  $A$  – див. (3.89)), тим ширшою є енергетична зона;

- ♦ енергетичні зони можуть розділятися проміжками заборонених інтервалів енергії  $E_g$ , що називаються забороненими зонами;

- ♦ рівень  $E_a$  в ізольованому атомі може бути вироджений. При утворенні кристалічної ґратки виродження може повністю або частково зникати. Атомний рівень при цьому розщеплюється на кілька зон. Максимальне число їх дорівнює ступеню виродження. Для атома фактор виродження  $g = 2l + 1$  ( $l$  – азимутальне квантове число). Тоді, наприклад для  $p$ -стану, коли  $l = 1$ ,  $g = 3$ ;

- ♦ за дії на кристал факторів, що викликають зміни відстані між атомами (температура, тиск), змінюватимуться інтеграли перекривання, отже ширини енергетичних і заборонених зон.

Таким чином, розгляд руху електрона у полі періодичного потенціалу в моделі як квазівільного, так і квазізв'язаного електрона допомагає дійти важливого висновку:

**Незалежно від моделі рух електрона у періодичному потенціалі має спільні риси: електронні стани характеризуються наявністю смуг дозволених значень енергії, що розділені енергетичними смугами недозволених (заборонених) станів. Усередині кожної енергетичної смуги (енергетичної зони) залежність енергії від хвильового вектора є періодичною, із періодом, що визначається кристалічною ґраткою.**

### ■ 3.7. Типи міжатомних зв'язків у кристалах

Усі види зв'язку між атомами кристала зумовлені притяганням або відштовхуванням електричних зарядів. Тип чи сила зв'язку визначається електронною будовою атомів, що утворюють кристал. Слабкі ван-дер-ваальсові зв'язки виникають завдяки квантовим флуктуаціям електронної густини, що приводить до виникнення та взаємодії між електричними диполями на атомах. Такими слабкими зв'язками часто можна нехтувати за наявності сильніших іонних, ковалентних або металічних зв'язків.

Існування стабільних зв'язків між парою ізольованих атомів або між атомами у тривимірному кристалі передбачає, що повна енергія просторової конфігурації позитивних іонних остовів і зовнішніх електронів є меншою, ніж для будь-якої конфігурації (у т.ч. у випадку розведення атомів на нескінченну відстань). Дефіцит енергії такої конфігурації порівняно з енергією ізольованих атомів називається енергією зв'язку, величина якої складає від 0,1 еВ на атом для твердих тіл зі зв'язками типу Ван-дер-Ваальса до 7,0 еВ – на атом і ще більше – для ковалентних та іонних сполук і деяких металів. Таким чином усі тверді тіла можна розподілити за характером міжатомних сил. Згідно із такою класифікацією тверді тіла можна розподілити на п'ять категорій: молекулярні, іонні, ковалентні, металічні кристали, кристали із водневими зв'язками.

Зв'язок між атомами в молекулярних кристалах обумовлюється слабкими ван-дер-ваальсовими силами притягання. Класичним прикладом молекулярних кристалів є кристали інертних газів, які називаються молекулярними, оскільки міжатомні сили, що відповідають за хімічний зв'язок у них, належать до типу, який запропонував Ван-дер-Ваальс для пояснення поведінки молекул у реальному газі. Ван-дер-ваальсові сили притягання є єдиною силою взаємодії між атомами із заповненими електронними оболонками, коли перекривання електронних хвильових функцій відсутнє. Енергія зв'язку у цьому випадку є малою із типовим значенням  $\sim 0,1$  еВ на один атом. В іонних і ковалентних кристалах ця енергія складає величину  $\sim 10,0$  еВ на атом; у металах ця енергія лежить в інтервалі від 1,0 до 5,0 еВ на атом; у кристалах із водневим зв'язком (напр., в  $\text{H}_2\text{O}$ ) енергія зв'язку, віднесена до однієї молекули, складає величину  $\sim 0,5$  еВ. Як вже зазначалося, ван-дер-ваальсово притягання між атомами із заповненими електронними оболонками виникає через їхню взаємну поляризацію (рис. 3.12), поява якої пов'язана з тим, що в атома (незважаючи на відсутність середнього дипольного моменту) наявний миттєвий флуктуаційний дипольний момент, здібний індукувати дипольний момент в іншого атома, унаслідок чого два атоми, що знаходяться на відстані  $R$  один від одного, притягуються з енергією

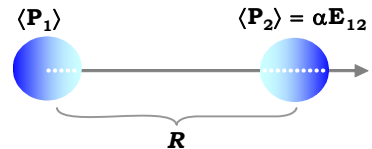


Рис. 3.12. Виникнення ван-дер-ваальсової взаємодії

$$\varphi_w = -\frac{C}{R^6}, \quad (3.95)$$

де  $C$  – стала, характерні значення якої складають, наприклад для Kr величину  $0,01 \cdot 10^{60}$  Ерг $\cdot$ см $^6$ . Наприклад, при  $R = 4 \text{ \AA}$  енергія  $\varphi_w \approx 0,013$  еВ, тобто  $\varphi_w/k_b \approx 100$  К, що за порядком величини збігається з температурою плавлення кристала криптонію. Виникнення ван-дер-ваальсової взаємодії можна краще зрозуміти за допомогою таких простих міркувань. Нехай на першому атомі спонтанно виник дипольний момент

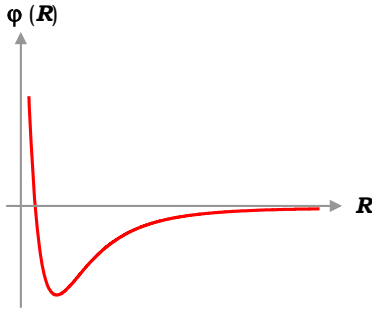
$\mathbf{P}_1$ , що індукує у точці знаходження другого атома електричне поле  $\mathbf{E}_{12} = \mathbf{P}_1/R^3$ , яке, у свою чергу, поляризує другий атом так, що дипольний момент на другому атомі дорівнюватиме  $\mathbf{P}_2 = \alpha \mathbf{E}_{12}$  ( $\alpha$  – поляризованість атома). Потенціальна енергія взаємодії індукованого дипольного моменту  $\mathbf{P}_2$  із полем  $\mathbf{E}_{12}$  визначатиме міжатомну взаємодію Ван-дер-Ваальса:  $\varphi_w = \mathbf{P}_2 \cdot \mathbf{E}_{12}$ , або в явному вигляді  $\varphi_w = -\alpha \mathbf{P}_1^2/R^6$  (тобто має вигляд (3.95)). Оскільки при зменшенні відстані між атомами  $R$  сила притягання зростає, вона могла б привести до необмеженого стиснення кристалу. Але у реальній ситуації за достатньо малих відстаней між атомами виникає зумовлена наявністю антизв'язуючих орбіталей (або кулонівським відштовхуванням електронних хмар, які перекриваються) сила відштовхування, яка описується потенціалом

$$\varphi_r = \frac{A}{R^n}, \quad (3.96)$$

де  $A$  – стала,  $n = 12$ . Таким чином, у молекулярних кристалах потенціальна енергія взаємодії, що викликає конденсацію речовини, описується потенціалом Леонарда–Джонса

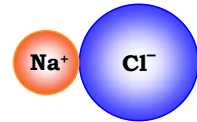
$$\varphi(R) = \varphi_0 \left[ \left( \frac{\sigma}{R} \right)^{12} - \left( \frac{\sigma}{R} \right)^6 \right], \quad (3.97)$$

типову поведінку якого подано на рис. 3.13. Мінімум потенціалу формується на відстані  $R_{\min}$ , що відповідає міжатомній відстані в кристалі.



**Рис. 3.13.** Потенціал міжатомної взаємодії типу Леннарда–Джонса в молекулярному кристалі

Характерним прикладом кристала з іонним зв'язком є кристали NaCl (рис. 3.14), при формуванні яких менш електровід'ємний атом натрію віддає свій валентний електрон атому хлору. Такі кристали називаються іонними саме тому, що утворені позитивними іонами натрію ( $\text{Na}^+$ ) і негативними іонами хлору ( $\text{Cl}^-$ ). Іонний зв'язок виникає завдяки електростатичній взаємодії системи катіонів ( $\text{Na}^+$ )



**Рис. 3.14.** Утворення іонного зв'язку в кристалах NaCl

та аніонів ( $\text{Cl}^-$ ), що по чергово займають позиції у вузлах кристалічної ґратки. Електронна конфігурація іонів за сутністю збігається з конфігурацією атомів інертного газу, і розподіл заряду поблизу кожного з іонів є сферично-симетричним. Відповідно іонні сполуки кристалізуються у щільно упаковані структури. Найтипівішими структурами іонних кристалів є прості кубічні ґратки. Іонні кристали характери-

зуються високим питомим опором та є ізоляторами. Через значне перекриття електронних оболонок іонів в іонних кристалах міжатомний зв'язок виникає у результаті двох протидіючих внесків. Перш за все, внесок дає потенціальна енергія кулонівської взаємодії

$$\varphi_C = \pm \frac{e^2}{R} \quad (3.98)$$

між двома іонами, що розташовані на відстані  $R$  один щодо одного. При цьому знак "+" відповідає взаємодії між іонами одного сорту (напр.,  $\text{Na}^+ - \text{Na}^+$ , або  $\text{Cl}^- - \text{Cl}^-$ ), а знак "-" належить іонам різного сорту (напр.,  $\text{Na}^+ - \text{Cl}^-$ ). Оскільки заряди в іонних кристалах розташовані почергово (позитивний і негативний) у просторі, то сумарна кулонівська взаємодія між сусідніми іонами відповідає притягання, що само собою викликало б необмежене стиснення ґратки. По-друге, у кристалі існує міжатомна взаємодія, обумовлена перекриттям електронних оболонок сусідніх атомів, яку за Борном–Майором можна записати у вигляді

$$\varphi_{B-M} = Be^{-R/b}, \quad (3.99)$$

де  $B$  та  $b$  є емпіричними константами. Таким чином повну потенціальну енергію міжатомної взаємодії в іонних кристалах можна записати у вигляді суми (3.98) і (3.99)

$$\varphi_i = -\frac{e^2}{R} + Be^{-R/b}. \quad (3.100)$$

Як видно із рис. 3.15, результуючий потенціал має чіткий мінімум. Він відповідає жорсткому відштовхуванню за малих значень міжатомної відстані  $R$  і притягання кулонівського типу на великих відстанях.

При визначенні ковалентного зв'язку часто важко зробити чітке розрізнення між іонним і ковалентним зв'язками. Дійсно, на користь умовності такого розрізнення свідчить той факт, що характерна енергія зв'язку в іонних і ковалентних кристалах має той самий порядок величини (10 еВ на атом), відповідно й температури кипіння та плавлення в одному класі кристалів є такими самими великими, як і в іншому.

Але існує чимало притаманних ковалентним кристалам рис, які відрізняють їх від іонних кристалів, наприклад ковалентні кристали характеризуються не тільки великою твердістю, але й крихкістю. Класичними

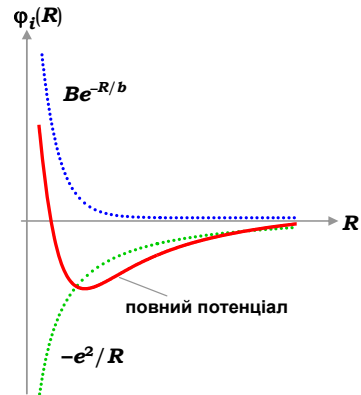


Рис. 3.15. Внески кулонівських сил і сил відштовхування до міжатомного потенціалу в іонних кристалах

прикладом кристалів із ковалентним зв'язком можуть служити напівпровідники IV групи (вуглець, кремній, германій). Ковалентні зв'язки переважають також у кристалах типу  $A^{III}B^V$ . Природу ковалентного зв'язку можна визначити таким чином: ковалентний зв'язок (його ще називають валентним або гомеополярним) є парним зв'язком, коли пара електронів одночасно належить двом атомам. У результаті перекривання хвильових функцій електронів в області простору між атомами утворюється висока щільність електричного заряду (рис. 3.16). Число ко-

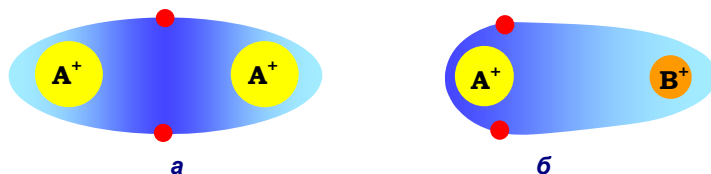


Рис. 3.16. Схема ковалентного зв'язку:

**а** – між двома однаковими атомами;

**б** – між двома різними атомами із визначеним ступенем іонності

валентних зв'язків, що здатний утворювати атом, є обмеженим і дорівнює числу електронів, яких не вистачає у зовнішній оболонці даного атома для утворення заповненої електронної конфігурації найближчого (у періодичній таблиці) інертного газу. Крім того, ковалентні зв'язки характеризуються чітко вираженою просторовою направленістю. Наприклад, атом вуглецю може мати чотири ковалентні зв'язки, які попарно утворюють між собою кут  $109,5^\circ$ , тобто структура ковалентних зв'язків вуглецю являє собою тетраедр. Характерне при цьому тетраедричне розташування атомів реалізується в алмазі та великій кількості органічних речовин. За ковалентного зв'язку відбувається "усуспільнювання" атомами парної кількості електронів: по два на кожний зв'язок. Кратний зв'язок утворюють  $2n$  електронів, де  $(n-1)$  – порядок зв'язку, який часто називають порядком  $\pi$ -зв'язку. У загальному випадку розподіл електронної густини між атомами, що утворюють ковалентний зв'язок, зсувається у напрямку від катіону до аніону. Про такий зв'язок кажуть як про частково іонний, тобто утворюється зв'язок типу  $(A^- - B^+)$  і частково ковалентний зв'язок –  $(A:B)$  (рис. 3.16 б). У рамках квантової теорії його часто описують суперпозицією хвильових функцій іонного та ковалентного станів.

Часто тверді тіла з ковалентним зв'язком можуть кристалізуватись у кількох структурних модифікаціях із майже однакою енергією зв'язку. Кожна із таких модифікацій є енергетично найвигіднішою у визначених умовах (напр., у різних інтервалах температур і тиску). Існування однієї і тієї самої речовини у різних структурних модифікаціях називають алотропією або поліморфізмом. Наведемо два приклади:

- ♦ кристали ZnS, які можуть кристалізуватись у кубічну (типу цинкової обманки) або гексагональну структуру (типу вюрциту). Число найближчих сусідів (у першій координаційній сфері) для обох модифікацій є однаковим, але число атомів у другій координаційній сфері є різним, що й обумовлює деяку різницю в енергії зв'язку між структурами;
- ♦ напівпровідник карбід кремнію може існувати у кількох кристалічних структурах – від простої кубічної до гексагональної, які незначно різняться за своїми електронними властивостями.

Для структур із *металічним зв'язком* характерними є відносно великі міжатомні відстані та чимале число найближчих сусідів. Наприклад, у літію міжатомна відстань складає 0,304 нм (у молекулярного кристалу хлору, що кристалізується у тетрагональну структуру, ця відстань є ще більшою – 0,434 нм). Число найближчих сусідів для літію складає 8, а для міді – 12. Кожен атом металу втягнутий в утворення великої кількості парних зв'язків таким чином, що вони не можуть бути локалізовані. Тому за рахунок резонансу зовнішніх електронів кожного атома утворюються делокалізовані електронні стани, які можна розглядати як наявність електронного газу, що разом з іонним остовом і реалізує металічний зв'язок. У багатьох металів в утворенні таких зв'язків беруть участь лише по одному електрону із зовнішньої оболонки. Наприклад, атом літію має тільки один електрон у незаповненій зовнішній оболонці, тому кристал літію можна уявити як систему правильно розташованих іонів  $\text{Li}^+$  (сферичної форми із радіусом 0,068 нм), що оточені електронним газом. Велика міжатомна відстань у металах, що є більшою за відстань у відповідних двоатомних молекулах (довжина ковалентного зв'язку в молекулі  $\text{Li}_2$  складає 0,267 нм), свідчить про малі значення енергії, що приходить на кожен зв'язок у металі. Але повна енергія зв'язку в металі є вищою, ніж в окремих молекулах за рахунок великої кількості зв'язків (хоча й слабкіших) у кристалі металу. Наприклад, енергія зв'язку зростає з 0,6 еВ у  $\text{Li}_2$  до 1,8 еВ – для металічного літію. Отже можна стверджувати, що для металічного зв'язку характерним є те, що хвильові функції електронів, які утворюють електронний газ, сильно перекриваються і фактично являють собою єдину хвильову функцію електронного газу, що задовольняє умовам руху у періодичному кристалічному потенціалі. Електрони, які беруть участь у формуванні зв'язку, мають різні енергії. При цьому середня енергія валентного електрона є меншою, ніж в ізольованому атомі, чим і зумовлюється наявність стабільної кристалічної структури.

Окремо серед міжатомних зв'язків, що забезпечують виникнення стабільних структур за конденсації речовини у тверде тіло, стоїть водневий зв'язок. Атом водню, який має лише один електрон, здатний утворювати ковалентний зв'язок лише з одним атомом. Але відомо, що атом водню може брати участь в утворенні додаткового електростатичного зв'язку з іншим атомом, що характеризується значною величиною електровід'ємності. Таким атомом може бути атом фтору, кисню або азоту. Цей додатковий *водневий зв'язок* характеризується енергією від 0,1 до 0,5 еВ і забезпечує з'єднання молекул  $\text{H}_2\text{O}$  в кристалах льоду, що мають структуру типу вюрцити, де відстань між атомами кисню двох сусідніх молекул дорівнює 0,276 нм, що більш ніж у два рази перевищує відстань зв'язку O–H у молекулі води (0,096 нм).

### ■ 3.8. Метали, діелектрики, напівпровідники. Елементарна класифікація

Тепер, після того, як ми ознайомились з методами класифікації кристалічних ґраток, обговорили фундаментальні наближення та моделі, що лежать в основі фізики твердого тіла, можемо, користуючись елементарними міркуваннями, визначити два принципово різні типи твердого тіла, в які відбувається конденсація атомів. Уявимо, що в процесі конденсації бере участь дуже велика кількість атомів. Зменшення відстані між атомами до розмірів сталої ґратки приведе до деякого зниження потенціальної енергії, і взаємодія між атомами в ґратці викличе розмиття енергетичних рівнів валентних електронів у майже суцільні смуги – енергетичні зони (див. рис. 1.27). Стани у зонах заповнюватимуться електронами відповідно до принципу Паулі (тобто на одному енергетичному рівні може бути лише один електрон, або два – із протилежними спінами). Таким чином, якщо в кристалічне тверде тіло конденсувались атоми з одним валентним електроном, то енергетична зона, в яку розмісться рівень валентних електронів, буде заповнена рівно наполовину – кожне дозволене значення енергії (а їх енергій буде рівно стільки, скільки атомів конденсувались у кристал) буде зайнято двома електронами із протилежними спінами. Як результат маємо випадок типового метала. Дійсно, основний стан є симетричним щодо заміни знака у будь-якого із хвильових векторів, тобто в основному стані електронний струм у такій системі є нульовим. Але для збудження електрона необхідна зовсім невелика енергія, оскільки всі рівні вище рівня з максимальною енергією (яка має назву



енергії Фермі) порожні, тобто інфінітезимально мале поле, прикладене до такого кристала, викликає рух електронів. Зрозуміло, що перехід електронів, які лежать поблизу енергії Фермі, у збуджений стан відбуватиметься завдяки будь-якому збуренню системи, коли перехід не є забороненим. Наприклад, таким збудженням може бути ненульова температура. Це означає, що метали характеризуються наявністю великої кількості збуджених електронів, що приходиться на одиницю об'єму. При цьому такі електрони у більшості випадків зі статистичного погляду поводитимуться як молекули майже ідеального газу.

Якщо ж атоми мали по два валентних електрони, то за їх конденсації у кристал енергетична зона, утворена розмиттям рівня валентних електронів, буде повністю заповнена, і для збудження будь-якого з них потрібно витратити деяку енергію, яка, щонайменше, дорівнює енергетичній відстані між заповненою і порожньою зонами та утворилась із першого збудженого рівня валентного електрона в ізолюваному атомі. Із міркувань симетрії ясно, що електричний струм у такій системі не має протікати, оскільки в повністю заповненій зоні кожному електронному стану, що характеризується імпульсом  $\mathbf{p}$ , відповідатиме стан, що характеризується імпульсом  $-\mathbf{p}$  і, таким чином, сумарний потік електронів дорівнюватиме нулю. Щоб викликати струм в такій системі треба прикласти електричне поле, яке збуджує деякі електрони в дозволені стани, що лежать вище забороненої зони та дають результуючий внесок у струм. Однак для цього необхідно надати електрону енергію, більшу за ширину забороненої зони. Слабке постійне поле таку енергію передати електронам не може. Зрозуміло, що у випадку достатньо великих значень енергії забороненої зони для збудження електронів у верхню порожню зону не достатнє нагріти кристал до температури порядку кімнатної. Дійсно, концентрація електронів, які за температури  $T$  спроможні подолати енергетичний зазор, що дорівнює енергії забороненої зони  $E_\Delta$ , буде експоненційно малою  $\sim \exp(-E_\Delta/kT)$ .

У результаті термічно збуджені електрони в таких системах характеризуватимуться дуже малими концентраціями і не впливатимуть помітно на властивості такого кристалу. Ці кристали, зокрема характеризуються дуже низькими значеннями електропровідності, і називаються діелектриками.

Припустимо тепер, що ширина забороненої зони не є великою. Тоді внаслідок теплових флуктуацій за ненульової температури незначна, але відмінна від нуля кількість електронів збуджуватиметься у верхню зону. Число таких електронів  $\sim \exp(-E_g/kT)$ . Збуджені електрони можуть брати участь у багатьох процесах, що визначають реакцію системи на зовнішнє збурення, зокрема переносити струм. Подібний матеріал

характеризуватиметься ненульовою провідністю, яка швидко зростає зі зростанням температури. Як буде видно з подальшого, такі кристали – напівпровідники – характеризуватимуться великою кількістю інших цікавих і корисних властивостей. Таким чином, можна стверджувати:

**Різниця між напівпровідниками та діелектриками є скоріше кількісною, ніж якісною, і всі характерні властивості напівпровідників формуються електронними станами поблизу країв енергетичних зон та енергетичними рівнями домішок, що вводяться у напівпровідниковий кристал.**

Залишаючись у рамках простої моделі конденсації атомів у кристал, яка розглядалась в першому розділі, можна сформулювати точніші правила передбачення властивостей твердого тіла:

**Тверде тіло, в якому на елементарну комірку приходиться один вільний електрон, завжди буде металом.  
Тверде тіло, в якому на елементарну комірку приходиться непарне число електронів, також завжди буде металом.**

Однак тут існують кілька винятків. Такі елементи, як As, Sb і Bi, в яких на валентних рівнях знаходяться по п'ять електронів, кристалізуються у структури із двома атомами на елементарну комірку, тобто у таких кристалах на елементарну комірку приходиться парне число електронів, і вони є неметалами. Дійсно, десять електронів у цих речовинах майже повністю заповнюють п'ять енергетичних зон, але остання з них заповнена не остаточно і, крім того, ще й перекривається із шостою зоною. Завдяки цьому за будь-якої температури (аж до абсолютного нуля) завжди існує деяка кількість електронів, що здатні переходити у збуджені стани, тобто переносити електричний струм. З іншого боку, ці речовини характеризуються незначною, порівняно із металами ( $\sim 10^{18}$ – $10^{20}$  см<sup>-3</sup>), концентрацією електронів, і зі зростанням температури у цих речовинах концентрація носіїв збільшується, тобто зростає їх електропровідність. Такі речовини, що за своїми властивостям знаходяться у проміжку між металами та напівпровідниками, називаються напівметалами.

**Тверде тіло, в якому на елементарну комірку приходиться парне число електронів, як правило, є діелектриком.**

Але може відбутися й так, що в деяких напрямках енергетичні зони можуть перекриватись. Тоді енергетично вигідним може бути не заповнення нижньої зони, а перехід частини електронів у верхню зону, і тверде тіло поводитиметься як напівметал.

### ■ 3.9. Задачі

1. Як вплине дія слабкого зовнішнього електричного поля на енергетичні зони електрона в кристалічній ґратці.

**Розв'язок.** Нехай зовнішнє поле направлене вздовж вісі  $OX$ , тобто  $\mathbf{E} = (E_x, 0, 0)$ . Вважатимемо, що зовнішня сила, яка діє на електрон, не може викликати переходи електрона із зони до зони, тобто виконується нерівність

$$Fa \ll E_g, \quad (3.101)$$

де  $a$  – стала ґратки).

Потенціальна енергія електрона визначається формулою

$$U(\mathbf{r}) = -e\mathbf{E} \cdot \mathbf{r} = -eEx, \quad (3.102)$$

і рівняння Шредингера набуває вигляду

$$\left[ -\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}) + U(\mathbf{r}) \right] \Psi = E\Psi. \quad (3.103)$$

Якщо потенціал слабо змінюється на відстанях порядку розміру сталої ґратки, то в області зміни хвильової функції атома його можна вважати сталим. Тоді (3.103) можна переписати як

$$\left[ -\frac{\hbar^2}{2m_0} \Delta + V(\mathbf{r}) \right] \Psi = (E - U)\Psi = \bar{E}\Psi, \quad (3.104)$$

де  $\bar{E}$  – енергія електрона за відсутності зовнішнього поля. Таким чином, можна записати енергію електрона в кристалі, на який діє зовнішнє поле

$$E = \bar{E} + U. \quad (3.105)$$

Враховуючи формулу для енергії електрона в кристалі (3.91), маємо

$$E^1 = E_a + C' + 2A'(\cos k_x a + \cos k_y a + \cos k_z a) - eEx. \quad (3.106)$$

Отже можна дійти такого висновку: якщо зобразити залежність енергії від координати  $x$ , то видно, що дія поля викличе нахил енергетичної діаграми. При цьому ширина зони та ширина забороненої зони зберігатимуться (рис. 3.17).

2. Поясніть, чому на рис. 1.27 (див. розд. 1) показано, що зони стають ширшими зі зростанням енергії.

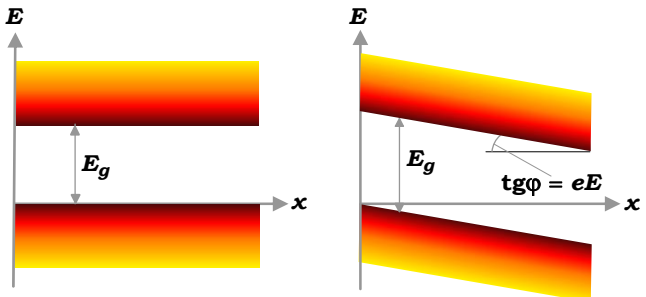


Рис. 3.17

3. Отримайте зв'язаний стан у моделі Кроніна–Пенні за умови, що  $b \rightarrow \infty$ .
4. Обчисліть енергію в методі квазізв'язаного електрона для тетрагонального кристала.

**Указівка.** Використайте (3.90), вважаючи, що міжатомні відстані вздовж різних напрямків є різними, обмінний інтеграл має різні сталі значення для попарної взаємодії даного електрона на даному вузлі із найближчими сусідами.

## Список літератури

1. *Киреев П.С.* Физика полупроводников. – М.: Высшая шк., 1975.
2. *Анималу А.* Квантовая теория кристаллических твердых тел. – М.: Мир, 1981.
3. *Давыдов А.С.* Квантовая механика. – М.: Наука, 1973.
4. *Ашкрофт Н., Мермин Н.* Физика твердого тела: В 2 т. – М.: Мир, 1979.
5. *Ming-Fu Li.* Modern Semiconductor Quantum Physics. – Singapore: World Scientific, 1995.
6. *Kasap S.O.* Principles of Electronic Materials and Devices. – N.Y.: McGraw-Hill, 2005.

## Розділ 4

# МЕТОДИ ОБЧИСЛЕННЯ ЗОННОЇ СТРУКТУРИ КРИСТАЛІЧНИХ ТВЕРДИХ ТІЛ

Розрахунки зонної структури напівпровідників (як зазвичай і кристалів) є важливою і досить складною задачею теорії твердого тіла. При цьому більшість застосовуваних для цього методів є комбінацією методів квазівільного та квазізв'язаного електрона. У цьому розділі дамо короткий огляд найбільш уживаних методів обчислення зонної структури твердих тіл. Але спочатку викладемо загальні міркування щодо обчислення зонної структури твердих тіл і напівпровідників, зокрема. Як згадувалось вище, енергетичний спектр електронів у кристалі із застосуванням одноелектронного та адіабатичного наближень визначається із рівняння Шредингера

$$\left[ -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (4.1)$$

де  $U(\mathbf{r})$  – періодичний потенціал кристала. Таким чином знаходження власних функцій і значень цього рівняння зводиться до визначення потенціалу  $U(\mathbf{r})$  і розв'язку рівняння з обумовленим потенціалом. Зрозуміло, що потенціал варто було б визначити із самоузгодженої схеми Хартрі–Фока, а далі розв'язувати рівняння. Але нездоланні обчислювальні труднощі не дозволяють послідовно виконати цю програму цілком і змушують використати наближені методи розрахунку. Наведемо загальну схему майже всіх методів розрахунку зонної структури. Будь-який розв'язок (4.1) розкладають у ряд за повним набором функцій, що задовольняють тим самим граничним умовам, що і рівняння (4.1). При цьому зазвичай вибирають деякий обмежений набір функцій  $\varphi_1, \varphi_2, \varphi_3, \dots, \varphi_l$  таким чином, щоб вони найкраще відповідали суті задачі. Тепер хвильова функція записується у вигляді

$$\psi(\mathbf{r}) = \sum_{i=1}^l c_i \varphi_i(\mathbf{r}). \quad (4.2)$$

Як найдоцільніше вибрати систему функцій  $\varphi_i(\mathbf{r})$ ? Розглянемо блоківську функцію

$$\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (4.3)$$

що має вигляд плоскої хвилі із довжиною, більшою за довжину сталої ґратки, модульованої множником  $u_{n\mathbf{k}}(\mathbf{r})$  із періодичністю ґратки. Оскільки поблизу атомних остовів хвильові функції поведуться подібно до атомних функцій, то через велику довжину плоскої хвилі  $e^{i\mathbf{k}\cdot\mathbf{r}}$  швидка

зміна функцій біля остова забезпечується тільки модулюючим множителем  $u_{n\mathbf{k}}(\mathbf{r})$ . Іншими словами, блохівські функції поводяться вдаліні від атомних остовів по суті як плоскі хвилі, а поблизу остовів – мають характер атомних функцій. Це наводить на думку, що як функції  $\varphi_i(\mathbf{r})$ , за якими відбувається розкладення хвильової функції електрона в кристалі, можна вибирати плоскі хвилі або атомні функції (згадаємо наближення майже вільних і квазізв'язаних електронів). Залежно від того, на яких функціях відбувається розкладення кристалічної хвильової функції, усі методи розрахунків зонної структури поділяються на два типи – методи комірок і варіаційні методи.

У першому методі як основна використовується періодичність власних функцій рівняння (4.1), завдяки якій можна обмежитись розглядом розв'язку рівняння Шредингера в одній елементарній комірці. Як комірки зазвичай використовують комірку Вігнера–Зейтца. Після розв'язку для однієї комірки Вігнера–Зейтца не складно, використовуючи граничні умови, що випливають із блохівського вигляду власних функцій, знайти хвильові функції для сусідніх комірок

$$\psi_{\mathbf{k}}(\mathbf{r}_b) = e^{i\mathbf{k}\cdot\mathbf{b}}\psi_{\mathbf{k}}(\mathbf{r}), \text{ та } \frac{\partial\psi_{\mathbf{k}}(\mathbf{r})}{\partial\mathbf{r}_b} = -e^{i\mathbf{k}\cdot\mathbf{b}}\frac{\partial\psi_{\mathbf{k}}(\mathbf{r})}{\partial\mathbf{r}_b}, \quad (4.4)$$

де  $\mathbf{r}$  та  $\mathbf{r}_b$  – радіус-вектори спряжених точок поверхні багатогранника Вігнера–Зейтца, що з'єднуються вектором трансляції  $\mathbf{b}$ . Похідні в (4.4) є нормальними до спряжених граней комірки Вігнера–Зейтца у точках  $\mathbf{r}$  та  $\mathbf{r}_a$ , а  $\mathbf{r}_b$  – зовнішні нормалі до поверхні багатогранника у цих точках. Розв'язок рівняння Шредингера всередині даної комірки знаходиться у вигляді суперпозиції самоузгоджених атомних хвильових функцій, тобто суми добутків сферичних гармонік на радіальні функції. Подальший розв'язок задачі зводиться до чисельного розв'язку рівняння Шредингера для радіальної функції із задоволенням відповідних граничних умов.

Варіаційні методи спрощують розв'язок диференційного рівняння Шредингера для варіаційної задачі про знаходження мінімуму деякого функціонала  $L$

$$\delta L = 0. \quad (4.5)$$

Ідея такого підходу полягає у наступному. Підставимо розкладання хвильової функції (4.2) до рівняння Шредингера (4.1) й подіємо на отримане рівняння зліва інтегральним оператором  $\int d\mathbf{r}\varphi_j^*(\mathbf{r})\dots$ . У результаті отримаємо систему лінійних алгебраїчних рівнянь щодо коефіцієнтів  $c_i$  розкладання (4.2)

$$\sum_{i=1}^l (\mathfrak{J}_{ji} - E\mathfrak{R}_{ji})c_i = 0. \quad (4.6)$$

Тут індекс  $j$  пробігає значення від 1 до  $l$  і введено такі позначення

$$\mathfrak{S}_{ji} = \int d\mathbf{r} \varphi_j^* H \varphi_i, \quad \mathfrak{R}_{ji} = \int d\mathbf{r} \varphi_j^* \varphi_i. \quad (4.7)$$

Зрозуміло, що у випадку вибору ортонормованого набору функцій  $\varphi_i(\mathbf{r})$  параметр

$$\mathfrak{R}_{ji} = \delta_{ji}. \quad (4.8)$$

Система (4.6) має нетривіальний розв'язок тільки у випадку, коли її детермінант дорівнює нулю

$$\det|\mathfrak{S}_{ji} - E\mathfrak{R}_{ji}| = 0. \quad (4.9)$$

Із секулярного рівняння (4.9) знайдемо власні значення  $E$ , тобто (4.9) дає можливість визначити зонну структуру системи, що вивчається. Незважко впевнитись, що (4.9) є результатом варіаційної задачі для функціонала

$$L\{\psi, \psi^*, E\} = \int d\mathbf{r} \psi^* (H - E)\psi \quad (4.10)$$

на класі пробних функцій  $\psi$ , які можна шукати у вигляді розкладання за деяким набором базисних функцій (4.2). Останні вибираються так, щоб вони були екстремалами функціонала (4.10), тобто варіація цього функціонала дорівнювала нулю. Таким чином, підставляючи розкладання (4.2) до (4.10), отримуємо, що

$$L = \sum_{i=1}^l \sum_{j=1}^l c_j^* c_i (\mathfrak{S}_{ji} - E\mathfrak{R}_{ji}). \quad (4.11)$$

Умова мінімуму функціонала  $L$  (4.5) еквівалентна рівнянням

$$\frac{\partial L}{\partial c_j^*} = 0, \quad \frac{\partial L}{\partial c_i} = 0. \quad (4.12)$$

Як видно, перше рівняння (4.12) приводить до системи (4.6), а друге – до комплексно спряженої системи (4.6)

$$\sum_{i=1}^l c_j^* (\mathfrak{S}_{ji} - E\mathfrak{R}_{ji}) = 0. \quad (4.13)$$

Розмірність секулярного рівняння визначається числом базисних функцій, за допомогою яких удається отримати задовільну точність розв'язку  $\psi$ . Отже основним чинником при розв'язку задачі на визначення зонної структури є вдалий вибір пробних функцій варіаційної задачі. Незважаючи на те, що обидва підходи мають загальний характер, пошуки розв'язків задачі для реальних напівпровідників у вигляді розкладання тільки за плоскими хвилями або атомними функціями не є цілком прийнятними. Наприклад, глибокі, цілком заповнені стани утворюють вузькі енергетичні зони та описуються хвильовими

функціями, що осцилюють поблизу ядер. Високі валентні та вільні стани електронів відповідають широким енергетичним зонам, а їхні хвильові функції "розмазані" по кристалу. Зрозуміло, що у першому випадку природніше використовувати підходи, які базуються на методі сильного зв'язку, що бере за основу атомні функції, а у другому – метод майже вільних електронів. На практиці ж зазвичай найбільший інтерес являють проміжні області енергетичного спектра – зони, побудовані із валентних електронів (валентна зона), і перша зона збуджених станів (зона провідності), що мають вирішальне значення для напівпровідників, зокрема для напівпровідникової електроніки. Для цих областей розроблено багато методів розрахунку зонної структури. З деякими з них ознайомимось нижче.

## ■ 4.1. Метод плоских хвиль

Завдяки періодичності потенціалу в кристалі потенціал і хвильову функцію можна розкласти у ряд Фур'є

$$U(\mathbf{r}) = \sum_{\mathbf{b}} c_{\mathbf{b}} e^{i\mathbf{b}\cdot\mathbf{r}}, \quad (4.14)$$

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{g}} a_{\mathbf{g}} e^{i(\mathbf{k}+\mathbf{g})\cdot\mathbf{r}}. \quad (4.15)$$

Із рівняння Шредингера отримаємо систему лінійних рівнянь для визначення коефіцієнтів розкладання  $a_{\mathbf{g}}$ :

$$\frac{\hbar^2(\mathbf{k}+\mathbf{g})}{2m} a_{\mathbf{g}} + \sum_{\mathbf{b}} c_{\mathbf{b}} a_{\mathbf{g}-\mathbf{b}} = E a_{\mathbf{g}}. \quad (4.16)$$

Для нетривіального розв'язку цієї системи її детермінант має дорівнювати нулю. У результаті отримаємо секулярне рівняння для визначення зв'язку між енергією та хвильовим вектором. Задля точності необхідно враховувати велику кількість плоских хвиль (або коефіцієнтів  $a_{\mathbf{g}}$ ), тобто під час практичних обчислень зонної структури розмірність такої системи виявляється дуже значною. Цей факт вимагає використання великих обчислювальних потужностей за невеликої гарантії прийнятної точності результатів. Тому такий метод найчастіше застосовують лише для оціночних розрахунків.



## ■ 4.2. Метод ортогоналізованих плоских хвиль

Оскільки поблизу ядра хвильова функція електрона суттєво відрізняється від плоскої хвилі, то для розкладення у ряд Фур'є необхідна наявність багатьох плоских хвиль зі значними хвильовими векторами. Із цієї причини метод плоских хвиль не є цілком прийнятним. Для його поліпшення можна поблизу ядра за хвильову функцію вибрати суперпозицію атомних хвильових функцій

$$\varphi_{\mathbf{k},j}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{n}} e^{i\mathbf{k}\cdot\mathbf{n}} \psi_j^a(\mathbf{r}-\mathbf{n}), \quad (4.17)$$

де  $j$  – атомний стан ( $s$ ,  $p$ ,  $d$  і т.д.),  $N$  – кількість атомів у кристалі, положення яких задається вектором  $\mathbf{n}$ . Тепер за основу вибираємо хвильову функцію у вигляді

$$\chi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{L^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} - \sum_j \mu_{\mathbf{k},j} \varphi_{\mathbf{k},j}(\mathbf{r}), \quad (4.18)$$

яка поблизу ядра поводитья в основному як атомна, а вдалині – як плоска хвиля. Використовуючи ці хвильові функції, можна далі діяти як і в попередньому методі, отримуючи секулярне рівняння, з якого і визначати закон дисперсії.

## ■ 4.3. Метод приєднаних плоских хвиль

Цей метод є вдосконаленням попереднього методу. Тут хвильова функція будується як комбінація плоских хвиль зовні атома та атомних хвильових функцій усередині атома

$$\Phi_{\mathbf{k}}(\mathbf{r}) = \theta(\mathbf{r}-\mathbf{r}_j) a_0 e^{i\mathbf{k}\cdot\mathbf{r}} + \sum_{lm} b_{lm} \theta(\mathbf{r}_j-\mathbf{r}) \psi_{lm}^a(\mathbf{r}), \quad (4.19)$$

де  $\theta(\mathbf{x})$  – функція Хевісайда:

$$\theta(\mathbf{x}) = \begin{cases} 1, & |\mathbf{x}| = 0 \\ 0, & |\mathbf{x}| \neq 0 \end{cases}. \quad (4.20)$$

При цьому  $j$ -й атом моделюється кулею із радіусом  $r_j$ . Цей метод, як і попередні, використовується в основному для оціночних обчислень.

## ■ 4.4. Метод псевдопотенціалу

Нехай  $\psi_{\mathbf{k}}^{(\alpha)}(\mathbf{r})$  – хвильова функція, що задовольняє рівнянню Шредингера для кристала, а  $\varphi_{\mathbf{k},j}(\mathbf{r})$  – сума, складена з атомних хвильових функцій так, щоб задовольнялась теорема Блоха (див. (4.17)). Важливо, що вимагається ортогональність цих функцій:

$$\int d\mathbf{r}(\psi_{\mathbf{k}}^{(\alpha)}(\mathbf{r}))^* \varphi_{\mathbf{k},j}(\mathbf{r}) = 0. \quad (4.21)$$

Будуємо таку функцію:

$$\Phi(\mathbf{r}) = \psi_{\mathbf{k}}^{(\alpha)}(\mathbf{r}) - \sum_j \alpha_{\mathbf{k},j} \varphi_{\mathbf{k},j}(\mathbf{r}). \quad (4.22)$$

Оскільки поблизу ядер  $\varphi_{\mathbf{k},j}(\mathbf{r})$  поводитья як одна з атомних функцій, що входять до блохівської суми, то функція  $\Phi(\mathbf{r})$  буде плавною всюди, у тому числі й поблизу ядер. Коефіцієнти  $\alpha_{\mathbf{k},j}$  при цьому можна знайти з умови ортогональності

$$\alpha_{\mathbf{k},j} = - \int d\mathbf{r}(\varphi_{\mathbf{k},j}(\mathbf{r}))^* \Phi(\mathbf{r}) = - \varphi_{\mathbf{k},j} \cdot \Phi. \quad (4.23)$$

При визначенні рівняння для функції  $\Phi(\mathbf{r})$  варто згадати, що хвильові функції, що її утворюють, мають задовольняти рівнянням  $H\psi_{\mathbf{k}}^{(\alpha)} = E(\mathbf{k})\psi_{\mathbf{k}}^{(\alpha)}$  та  $H\varphi_{\mathbf{k},j} = E_j\varphi_{\mathbf{k},j}$ . Тоді, використовуючи визначення (4.22), скомпонуємо із цих рівнянь таке

$$\left[ -\frac{\hbar^2}{2m} \Delta + U(\mathbf{r}) \right] \Phi(\mathbf{r}) + \sum_j \alpha_{\mathbf{k},j} E_j \varphi_{\mathbf{k},j} = E(\mathbf{k})\Phi(\mathbf{r}) + E(\mathbf{r}) \sum_j \alpha_{\mathbf{k},j} \varphi_{\mathbf{k},j}, \quad (4.24)$$

яке зручно переписати у вигляді

$$\left[ -\frac{\hbar^2}{2m} \Delta + U(\mathbf{r}) + \sum_j \frac{\alpha_{\mathbf{k},j} (E_j - E(\mathbf{r})) \varphi_{\mathbf{k},j}}{\Phi(\mathbf{r})} \right] \Phi(\mathbf{r}) = E(\mathbf{k})\Phi(\mathbf{r}). \quad (4.25)$$

Введемо псевдопотенціал

$$V_p = U(\mathbf{r}) + \sum_j \frac{\alpha_{\mathbf{k},j} (E_j - E(\mathbf{r})) \varphi_{\mathbf{k},j}}{\Phi(\mathbf{r})}. \quad (4.26)$$

Тоді (4.25) можна записати у стандартній формі

$$\left[ -\frac{\hbar^2}{2m} \Delta + V_p(\mathbf{r}) \right] \Phi(\mathbf{r}) = E(\mathbf{k})\Phi(\mathbf{r}). \quad (4.27)$$

Оскільки  $\Phi(\mathbf{r})$  є достатньо гладкою функцією, то її можна утворити з невеликого числа плоских хвиль, які забезпечать швидку збіжність проце-

дури для обчислення дисперсійної залежності  $E(\mathbf{k})$ . Дійсно, псевдопотенціал (4.26) складається із суми атомного  $U(\mathbf{r})$  та локального потенціалів

$$V_R(\mathbf{r}) = \sum_j \frac{\alpha_{\mathbf{k},j}(E_j - E(\mathbf{r}))\varphi_{\mathbf{k},j}}{\Phi(\mathbf{r})}, \quad (4.28)$$

тобто 
$$V_p = U(\mathbf{r}) + V_R(\mathbf{r}). \quad (4.29)$$

Локальний потенціал  $V_R(\mathbf{r})$  поводитья як потенціал відштовхування. Із (4.28) видно, що оскільки матричний елемент  $\alpha_{\mathbf{k},j}$  є негативним та енергетичні рівні  $E_j$  – меншими за рівні зовнішніх електронів  $E(\mathbf{r})$ , то  $V_R(\mathbf{r}) > 0$ . З іншого боку, оскільки  $U(\mathbf{r}) < 0$ , то атомний (кристалічний) потенціал є потенціалом притягання. Іншими словами, ці потенціали взаємно компенсують дію один одного. Така компенсація не є випадковою, а означає, що відповідним вибором хвильової функції  $\Phi(\mathbf{r})$  можна домогтися істотного зменшення псевдопотенціалу  $V_p$ . Звідси випливає, що псевдопотенціал є гладкішою функцією координат, ніж функція кристалічного потенціалу, оскільки поблизу атомних остовів кристалічний потенціал із глибокими провалами значною мірою компенсується потенціалом відштовхування. Нам відомо, що чим гладкішим є потенціал, тим гладкішою є хвильова функція, отже вона може бути представлена меншою кількістю членів ряду Фур'є.

Може здатися, що введення псевдопотенціалу не має великої користі, оскільки локальний потенціал залежить від функцій  $\Phi$  та енергії  $E(\mathbf{r})$ , які ще необхідно визначити. На практиці ці залежності є вельми слабкими, оскільки у (4.28), що визначає потенціал  $V_R(\mathbf{r})$ , функція  $\Phi$  входить під знаком інтеграла у чисельнику та знаменнику, завдяки чому залежність від  $\Phi$  значно слабшає. До того ж, оскільки  $E \gg E_j$ , то потенціал  $V_R(\mathbf{r})$  слабо залежить від енергії стану, що розглядається. Отже  $V_R(\mathbf{r})$  можна вважати просто функцією, яка в основному залежить від координат, і використовувати його як деякий емпіричний потенціал у рівнянні Шредингера. Фізично найуспішнішими є обчисленнями псевдопотенціалу, засновані на ідеї Хейне та Абаренкова. Сенс полягає у заміні істинного потенціалу модельним, що має вигляд потенціальної ями із плоским дном в області знаходження атомного остова,  $V_{H-A}(\mathbf{r}) = -A_l$  при  $|\mathbf{r}| \leq R_M$ , а за областю атомного остова потенціал вибирається у вигляді кулонівського:  $V_{H-A} = -ze^2/|\mathbf{r}|$  при  $|\mathbf{r}| > R_M$  ( $z$  – хімічна валентність). За вдалого вибору радіусу модельного потенціалу  $R_M$  глибина ями  $A_l$  для стану із заданим значенням моменту кількості руху  $l$  має бути підібрана так, щоб розв'язок модельного хвильового рівняння приводив до значень спектроскопічних термів для відповідного вільного іона. Метод псевдопотенціалу є одним із сучасних широко застосовуваних при обчисленні зонної структури напівпровідників і металів. Щодо останнього можна стверджувати: застосування методу псевдопотенціалу у

сучасній теорії твердого тіла привело до значного прогресу у розумінні властивостей металів і напівпровідників.

## ■ 4.5. **k-p-метод**

У теорії напівпровідників широко використовується метод, розроблений Шоклі, Кейном і Кіттелем. Для розуміння ідеї цього методу згадаємо, що для хвильової функції, яка задовольняє теоремі Блоха

$$\Psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\varphi_{\mathbf{k}}(\mathbf{r}), \quad (4.30)$$

рівняння Шредингера має такий вигляд

$$-\frac{\hbar^2}{2m}\Delta\varphi_{\mathbf{k}}(\mathbf{r}) + \left[ \frac{\hbar^2\mathbf{k}^2}{2m} - i\frac{\hbar^2}{m}\mathbf{k}\cdot\nabla + U(\mathbf{r}) \right] \varphi_{\mathbf{k}}(\mathbf{r}) = E(\mathbf{k})\varphi_{\mathbf{k}}(\mathbf{r}). \quad (4.31)$$

Оскільки оператор імпульсу  $\mathbf{p} = -\hbar\nabla$ , то (4.31) можна переписати як

$$\left[ -\frac{\hbar^2}{2m}\Delta + U(\mathbf{r}) + \frac{\hbar}{m}\mathbf{k}\mathbf{p} \right] \varphi_{\mathbf{k},n}(\mathbf{r}) = \left[ E_n(\mathbf{k}) - \frac{\hbar^2\mathbf{k}^2}{2m} \right] \varphi_{\mathbf{k},n}(\mathbf{r}). \quad (4.32)$$

Вважатимемо, що для деякого стану  $\mathbf{k}_0$  розв'язок рівняння Шредингера відомий. Тоді можна шукати розв'язок цього рівняння в області поблизу стану  $\mathbf{k}_0$  як суперпозицію хвильових функцій, відповідних стану  $\mathbf{k}_0$

$$\varphi_{\mathbf{k},n}(\mathbf{r}) + \sum_m c_{nm}(\mathbf{k} - \mathbf{k}_0)\varphi_{\mathbf{k}_0,n}(\mathbf{r}). \quad (4.33)$$

Оскільки хвильова функція  $\varphi_{\mathbf{k}_0,n}$  є розв'язком рівняння Шредингера

$$H_{\mathbf{k}_0}\varphi_{\mathbf{k}_0,n} = \left[ \frac{\mathbf{p}^2}{2m} + \frac{\hbar}{m}\mathbf{k}_0\mathbf{p} + \frac{\hbar^2\mathbf{k}^2}{2m} + U(\mathbf{r}) \right] \varphi_{\mathbf{k}_0,n} = E_n(\mathbf{k}_0)\varphi_{\mathbf{k}_0,n}, \quad (4.34)$$

то (4.34) можна записати у вигляді

$$\left[ H_{\mathbf{k}_0} + \frac{\hbar}{m}(\mathbf{k} - \mathbf{k}_0)\cdot\mathbf{p} + \frac{\hbar^2}{2m}(\mathbf{k}^2 - \mathbf{k}_0^2) \right] \varphi_{\mathbf{k},n} = E_n(\mathbf{k})\varphi_{\mathbf{k},n}.$$

Підставимо сюди хвильову функцію у вигляді (4.33), і подіємо на отримане рівняння оператором  $I_{\mathbf{k}_0} = \int d\mathbf{r}\varphi_{\mathbf{k}_0,n}^*$ ... У результаті отримаємо систему лінійних алгебраїчних рівнянь для визначення коефіцієнтів  $c_{mn}$

$$\sum_m \left\{ \left[ E_n(\mathbf{k}_0) + \frac{\hbar^2}{2m}(\mathbf{k}^2 - \mathbf{k}_0^2) \right] \delta_{mn} + \frac{\hbar}{m}(\mathbf{k} - \mathbf{k}_0)\cdot\mathbf{p}_{mn} \right\} \cdot c_{mn} = E_n(\mathbf{k})c_{mn}, \quad (4.35)$$

де введено позначення  $\mathbf{p}_{mn} = \int d\mathbf{r}\varphi_{\mathbf{k}_0,m}^*(\mathbf{r})\mathbf{p}\varphi_{\mathbf{k}_0,n}(\mathbf{r})$ . (4.36)

Якщо останній доданок у фігурних дужках розглядати як збурення, то у другому порядку теорії збурень отримаємо

$$E_n(\mathbf{k}) = E_n(\mathbf{k}_0) + \frac{\hbar}{m}(\mathbf{k} - \mathbf{k}_0)\mathbf{p}_{nn} + \frac{\hbar^2}{2m}(\mathbf{k}^2 - \mathbf{k}_0^2) + \frac{\hbar^2}{m^2} \sum_{n'} \frac{|(\mathbf{k} - \mathbf{k}_0) \cdot \mathbf{p}_{nn'}|^2}{E_n(\mathbf{k}_0) - E_{n'}(\mathbf{k}_0)}. \quad (4.37)$$

Якщо стан  $\mathbf{k}_0$  відповідає екстремуму енергії і  $\mathbf{p}_{nm} + \hbar\mathbf{k}_0 = 0$ , то закон дисперсії (4.37) відповідає сфероїдальним ізоенергетичним поверхням. У головних осях  $\mathbf{i}$  компоненти ефективних мас

$$\frac{1}{m_i^*} = \frac{1}{m} + \frac{2}{m^2} \sum_{n'} \frac{(\mathbf{i} \cdot \mathbf{p}_{nn'})^2}{E_n(\mathbf{k}_0) - E_{n'}(\mathbf{k}_0)} \quad (4.38)$$

та енергія мають стандартний вигляд із квадратичним законом дисперсії. Тобто  $\mathbf{k}\mathbf{p}$ -метод приводить до стандартного результату, що дає метод ефективної маси. За допомогою  $\mathbf{k}\mathbf{p}$ -методу зручно розглядати дисперсію електронів з урахуванням спин-орбітальної взаємодії, гамільтоніан якої має такий вигляд

$$H_{\text{SO}} = \frac{\hbar^2}{4m^2c^2} (\nabla U \times \mathbf{p}) \cdot \boldsymbol{\sigma}, \quad (4.39)$$

де  $\boldsymbol{\sigma}$  – матриці Паулі. Приєднання до гамільтоніана доданка, що описує спин-орбітальну взаємодію, викликає появу у квадратних дужках рівняння (4.35) доданків

$$H_{\text{SO}}^{(1)} = \frac{\hbar}{4m^2c^2} (\nabla U \times \mathbf{p}) \cdot \boldsymbol{\sigma} \quad (4.40)$$

та

$$H_{\text{SO}}^{(1)} = \frac{\hbar}{4m^2c^2} (\nabla U \times \mathbf{k}) \cdot \boldsymbol{\sigma}. \quad (4.41)$$

Ці оператори діють на модулюючу частину блохівської функції. У багатьох випадках їх можна розглядати як збурення. Таким чином загальний оператор збурення в  $\mathbf{k}\mathbf{p}$ -методі із врахуванням спин-орбітальної взаємодії матиме вигляд

$$H' = \frac{\hbar}{m} \left( \mathbf{k}\mathbf{p} + \frac{1}{4mc^2} (\nabla U \times (\mathbf{p} + \hbar\mathbf{k})) \boldsymbol{\sigma} \right). \quad (4.42)$$

Доданок  $H_{\text{SO}}^{(1)}$ , що не залежить від  $\mathbf{k}$ , аналогічний атомному параметру спин-орбітального розщеплення, а  $H_{\text{SO}}^{(2)}$ , пропорційний  $\mathbf{k}$ , – являє собою додаткову спин-орбітальну енергію, яка походить від особливостей руху електрона у періодичному полі кристала. Внесок від  $H_{\text{SO}}^{(2)}$  є набагато меншим за внесок до енергії електрона від гамільтоніана  $H_{\text{SO}}^{(1)}$ , оскільки квазіімпульс  $|\hbar\mathbf{k}|$  є дуже малим порівняно із середнім імпульсом еле-

ктрона  $|\mathbf{p}| = \langle \varphi_{\mathbf{k}}^* | \mathbf{p} | \varphi_{\mathbf{k}} \rangle \sim \hbar/a$  ( $a$  – стала ґратки) в області атомного остова, де спін-орбітальна взаємодія найістотніша. Це означає, що швидкість електрона на атомній орбіті є набагато більшою за швидкість хвильового пакета, який складається із хвильових векторів в околі  $\mathbf{k}_0$ . Найбільш плідним **кп**-метод виявився для розрахунків зонної структури кристалів із симетрією типу алмазу (напр., кремнію й германію) та цинкової обманки (напр., InSb і GaAs). Детальніше застосування **кп**-методу для обчислення закону дисперсії носіїв у напівпровіднику із врахуванням спін-орбітальної взаємодії буде розглянуто нижче.

## ■ 4.6. Метод ефективної маси

Напівпровідник в основному стані ( $T = 0^\circ \text{K}$ ) являє собою багатеелектронну систему із повністю заповненою валентною зоною, порожньою зоною провідності й неіонізованими домішками. Зрозуміло, що тільки завдяки збуренням (напр., ненульовій температурі) у зоні провідності можуть з'явитися заповнені стани, що заповнюватимуться, починаючи з тих, які характеризуються найменшими значеннями енергії. Іншими словами, починають заповнюватися стани у зоні провідності поблизу мінімальних значень енергії. Це означає, що найцікавішими є такі значення енергії електрона у напівпровіднику, що знаходяться поблизу мінімуму зони провідності. Раніше ми з'ясували, що поблизу екстремуму енергетичний спектр електрона в кристалі є принаймні квадратичною функцією квазіімпульсу. Цей факт є основою широко вживаного методу ефективної маси при визначенні енергетичного спектра електронів у напівпровіднику.

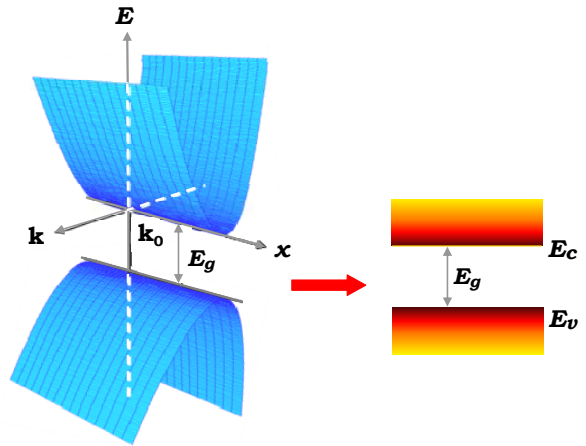


Рис. 4.1. Енергетичні стани поблизу дна зони провідності та стелі валентної зони

### 4.6.1. Розкладання енергії електрона поблизу екстремальних точок

Основна ідея методу ефективної маси полягає в наступному. Як відомо, хвильова функція електрона в ідеальному кристалі являє собою блохівську функцію  $\psi_{\mathbf{k}}(\mathbf{r})$ , що задовольняє рівнянню

$$\left[ -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) \right] \psi_{\mathbf{k}}(\mathbf{r}) = E(\mathbf{k}) \psi_{\mathbf{k}}(\mathbf{r}). \quad (4.43)$$

Поблизу екстремуму  $\mathbf{k}_0$  енергію електрона можна представити у вигляді квадратичної функції квазіімпульсу

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{\hbar^2}{2m^*} (\mathbf{k} - \mathbf{k}_0)^2. \quad (4.44)$$

Тепер уявимо, що ми вивчаємо рух вільної частинки із масою, яка збігається з ефективною масою електрона в кристалі  $m^*$  у сталому полі  $E(\mathbf{k}_0)$ . Динаміка такої частинки описується рівнянням Шредингера

$$H_e \Psi_0 = E_0 \Psi_0 \quad (4.45)$$

із гамільтоніаном

$$H_e = -\frac{\hbar^2}{2m^*} \Delta^2 + E(\mathbf{k}_0). \quad (4.46)$$

Неважно впевнитись, що хвильова функція  $\psi_0 = A e^{i(ax+by+cz)}$  є розв'язком рівняння (4.45) із гамільтоніаном (4.46). При цьому власні значення

$$E = E(\mathbf{k}_0) + \frac{\hbar^2}{2} \left( \frac{a^2}{m^*} + \frac{b^2}{m^*} + \frac{c^2}{m^*} \right). \quad (4.47)$$

Порівнюючи вирази для енергій (4.44) та (4.46), можна побачити, що енергії збігаються у випадку, коли

$$a = k_x - k_{0x}, \quad b = k_y - k_{0y}, \quad c = k_z - k_{0z}. \quad (4.48)$$

Таким чином, можна дійти дуже важливого висновку: гамільтоніан (4.46) має той самий спектр енергій, що й гамільтоніан поля ґратки в околі екстремуму  $\mathbf{k}_0$ , а його власні функції мають вигляд плоскої хвилі

$$\psi_0(\mathbf{r}) = A \times \exp\{i(\mathbf{k} - \mathbf{k}_0) \cdot \mathbf{r}\}, \quad (4.49)$$

тобто рух електрона у кристалі з енергіями, що знаходяться поблизу екстремального значення  $E(\mathbf{k}_0)$ , подібний до руху вільної частинки з ефективною масою  $m^*$  у сталому полі  $E(\mathbf{k}_0)$ . Таким чином:

**Метод ефективної маси неявно містить ідею квазічастинки – динаміку реальної частинки у складному кристалічному потенціалі можна описати деякою уявною вільною частинкою, що характеризується ефективною масою  $m^*$ .**

У випадку анізотропного кристала (4.44) перейде у рівняння

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{\hbar^2}{2} (m_{ij}^*)^{-1} (\mathbf{k}_i - \mathbf{k}_{0i})(\mathbf{k}_j - \mathbf{k}_{0j}), \quad (4.50)$$

а (4.47), записане в головних осях, – у рівняння

$$E = E(\mathbf{k}_0) + \frac{\hbar^2}{2} \left( \frac{(\mathbf{k}_x - \mathbf{k}_{0x})^2}{m_{xx}^*} + \frac{(\mathbf{k}_y - \mathbf{k}_{0y})^2}{m_{yy}^*} + \frac{(\mathbf{k}_z - \mathbf{k}_{0z})^2}{m_{zz}^*} \right). \quad (4.51)$$

Іншими словами, тепер квазічастинка характеризуватиметься різними значеннями мас за руху у різних напрямках. Отже, можна стверджувати, що для енергій поблизу мінімумів зони провідності електрон у кристалі можна вважати квазічастинкою, яка описується тензором ефективних мас.

#### 4.6.2. Динамічний аспект проблеми ефективної маси

Покажемо, що компоненти тензора  $(m_{ij}^*)^{-1}$  у (4.50) або параметр  $m^*$  у розкладанні (4.44) визначають динамічні властивості електронних збуджень у кристалі, та їх можна вважати масою деякої частинки. Розглянемо рух електрона, коли на кристал діє зовнішнє поле. Нехай це буде однорідне силове поле

$$V(\mathbf{r}) = -\mathbf{F}_{ext} \mathbf{r}. \quad (4.52)$$

Динаміка частинки у цьому випадку описується гамільтоніаном

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) + V(\mathbf{r}) = \hat{H}_0 + V(\mathbf{r}). \quad (4.53)$$

Визначимо оператор прискорення як похідну за часом від оператора швидкості

$$\hat{\mathbf{a}} = \frac{d\hat{\mathbf{v}}}{dt} = -\frac{i}{\hbar} (\hat{\mathbf{v}}\hat{H} - \hat{H}\hat{\mathbf{v}}). \quad (4.54)$$

Оскільки оператор швидкості комутує з гамільтоніаном  $\hat{H}_0$ , то

$$\hat{\mathbf{a}} = -\frac{i}{\hbar} (\hat{\mathbf{v}}V(\mathbf{r}) - V(\mathbf{r})\hat{\mathbf{v}}). \quad (4.55)$$

Представимо оператор швидкості через похідну від енергії за квазіімпульсом частинки  $\hat{\mathbf{v}} = dE/\hbar d\mathbf{k}$  і підставимо його до (4.55)

$$\hat{\mathbf{a}} = \frac{1}{i\hbar^2} \left[ \frac{dE}{d\mathbf{k}} V(\mathbf{r}) - V(\mathbf{r}) \frac{dE}{d\mathbf{k}} \right]. \quad (4.56)$$

Використаємо тепер явний вигляд оператора зовнішнього поля, попередньо записавши його в  $\mathbf{k}$ -просторі за допомогою представлення оператора координати в (4.52) через оператор імпульсу



$$\mathbf{r}(\mathbf{k}) = i\nabla_{\mathbf{k}} - i(\nabla_{\mathbf{k}} \ln \chi_{\mathbf{k}}^*), \quad (4.57)$$

де  $\chi_{\mathbf{k}}$  – блохівська хвильова функція (див. додаток С). Тоді оператор (4.52) можна записати у вигляді суми операторів диференціювання та множення

$$V(\mathbf{k}) = -i\mathbf{F}_{ext} \nabla_{\mathbf{k}} + i\mathbf{F}_{ext} \nabla_{\mathbf{k}} \ln \chi_{\mathbf{k}}^*. \quad (4.58)$$

Підставляючи цей вираз до (4.56) і зважаючи на те, що оператор швидкості комутує з другим доданком оператора (4.58), отримуємо

$$\hat{a}_i = \frac{1}{\hbar^2} F_j^{(ext)} \frac{d^2 E}{dk_j dk_i}. \quad (4.59)$$

При виведенні формули використано той факт, що  $i(\mathbf{F}_{ext} \cdot \nabla_{\mathbf{k}})(dE/d\mathbf{k})$  є добутком операторів диференціювання та множення, і його дія на деяку функцію дає такий результат

$$i(\mathbf{F}_{ext} \cdot \nabla_{\mathbf{k}}) \frac{dE}{d\mathbf{k}} \varphi(\mathbf{k}) = i \frac{dE}{d\mathbf{k}} (\mathbf{F}_{ext} \cdot \nabla_{\mathbf{k}} \varphi(\mathbf{k})) + i\mathbf{F}_{ext} \cdot \frac{d}{d\mathbf{k}} \frac{dE}{d\mathbf{k}} \varphi(\mathbf{k}). \quad (4.60)$$

Таким чином, комутатор

$$\begin{aligned} \frac{dE}{d\mathbf{k}} (-i\mathbf{F}_{ext} \nabla_{\mathbf{k}}) + (i\mathbf{F}_{ext} \nabla_{\mathbf{k}}) \frac{dE}{d\mathbf{k}} &= -\frac{dE}{d\mathbf{k}} (i\mathbf{F}_{ext} \nabla_{\mathbf{k}}) + \\ + \frac{dE}{d\mathbf{k}} (i\mathbf{F}_{ext} \nabla_{\mathbf{k}}) + i \left( \mathbf{F}_{ext} \cdot \frac{d}{d\mathbf{k}} \right) \frac{dE}{d\mathbf{k}} &= i\mathbf{F}_{ext}^{(ext)} \frac{d^2 E}{d\mathbf{k} d\mathbf{k}}. \end{aligned}$$

Взявши до уваги, що

$$\left( \mathbf{F}_{ext} \cdot \frac{d}{d\mathbf{k}} \right) \frac{dE}{d\mathbf{k}} = F_j^{(ext)} \cdot \frac{d}{dk_j} \frac{dE}{dk_i} \quad \text{та} \quad \frac{1}{\hbar^2} \frac{d^2 E}{dk_j dk_i} = (m_{ij}^*)^{-1}, \quad (4.61)$$

з'ясуємо таке: вектор прискорення за руху частинки в однорідному силловому полі зв'язаний із силою, що діє на частинку через тензор ефективної маси

$$a_i = F_j^{(ext)} (m_{ij}^*)^{-1}, \quad (4.62)$$

тобто прискорення надається електрону зовнішньою силою. При цьому динамічною характеристикою, що визначає реакцію електрона в кристалі на дію зовнішньої сили, є ефективна маса. Отже, незважаючи на те, що поле ґратки не прискорює електрон, воно діє на нього при формуванні його руху під дією зовнішніх полів. У випадку, коли ефективна маса є скаляром (або діагональним тензором) вектори  $\mathbf{a}$  та  $\mathbf{F}_{ext}$  є колінеарними. В інших випадках прискорення не колінеарне зовнішній силі. Розглянемо для простоти випадок руху електрона, що харак-

теризується скалярною ефективною масою в однорідному електричному полі ( $\mathbf{F}_{ext} = -e\mathbf{E}$ ,  $e$  – заряд електрона). Тоді, якщо зонний електрон знаходиться в околі мінімуму енергії, ефективна маса є величиною позитивною, і прискорення направлене проти поля, за напрямком сили. Якщо ж зонний електрон знаходиться в околі максимуму енергії, то ефективна маса частинки є негативною, і прискорення, яке отримує електрон, направлене по полю, тобто він рухається так, якби його заряд був позитивним. У випадках, коли частинки характеризуються негативною ефективною масою та позитивним зарядом, кажуть, що під дією зовнішнього поля рухаються *дірки*. Таким чином, при збудженні напівпровідника в стани з енергіями (які знаходяться поблизу дна зони провідності) квазічастинками, що описують такі збудження, існують зонні електрони. Якщо ж збудження відбуваються в стани з енергіями, що знаходяться поблизу стелі валентної зони, квазічастинки, які описують такі збудження, є дірками.

## ■ 4.7. Спін-орбітальна взаємодія

Як відомо, електрон характеризується "внутрішнім" моментом – спіном; в атомах взаємодія магнітного моменту спину електрона з його орбітальним рухом приводить до розщеплення та зсуву рівнів енергії атома. Зрозуміло, що спін-орбітальна взаємодія викликає аналогічні ефекти й у твердому тілі. Якщо електрон локалізований в околі певної просторової точки, то він знаходиться в електричному полі  $\mathbf{E}$ , що утримує його поблизу цієї точки. Але це поле діє також і на спін електрона. Строгий вивід гамільтоніана такої взаємодії проводиться в рамках релятивістської квантової теорії, але з точністю до безрозмірного фактора гамільтоніан можна знайти, виходячи із простих фізичних міркувань. Розглянемо електрон як класичну частинку та введемо дві системи координат – лабораторну, що пов'язана з центром локалізації електрона, і таку, що рухається разом з електроном зі швидкістю  $\mathbf{v}$ . Згідно із перетвореннями Лоренца для електромагнітного поля, електричне поле, що існує в лабораторній системі координат, обумовлює в рухомій системі магнітне поле

$$\mathbf{H}_L = \frac{1}{c} \mathbf{E} \times \mathbf{v}. \quad (4.63)$$

Якщо електрон рухається у самоузгодженому періодичному потенціалі кристала  $U(\mathbf{r})$ , то

$$\mathbf{E} = \frac{1}{e} \nabla U(\mathbf{r}). \quad (4.64)$$

Із теорії спину електрона відомо, що власний вектор оператора спину

$$\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\sigma}, \quad (4.65)$$

де  $\boldsymbol{\sigma}$  – дворядні матриці Паулі

$$\boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (4.66)$$

Видно, що проекція спіну на вісь  $OZ$  має значення  $s\hbar = \pm\hbar/2$ . З релятивістської теорії Дірака відомо, що магнітний момент обумовлений спіном електрона

$$\boldsymbol{\mu} = -\boldsymbol{\mu}_0 \boldsymbol{\sigma}, \quad (4.67)$$

де магнетон Бора

$$\boldsymbol{\mu}_0 = \frac{e\hbar}{2mc}. \quad (4.68)$$

Енергія дипольного моменту  $\boldsymbol{\mu}$  у магнітному полі  $\mathbf{H}$

$$W = -\boldsymbol{\mu} \cdot \mathbf{H}. \quad (4.69)$$

Використовуючи (4.63) та (4.64), маємо

$$\boldsymbol{\mu} \cdot \mathbf{H} = \boldsymbol{\mu}_0 \boldsymbol{\sigma} \cdot \left( \frac{1}{e} \nabla U(\mathbf{r}) \times \frac{\mathbf{v}}{c} \right). \quad (4.70)$$

Оскільки  $\mathbf{v} = \mathbf{p}/m = -(\i\hbar/m)\nabla$ , то (4.70) можна переписати

$$\boldsymbol{\mu} \cdot \mathbf{H} = -\frac{i\hbar^2}{2m^2c^2} \boldsymbol{\sigma} \cdot (\nabla U(\mathbf{r}) \times \nabla). \quad (4.71)$$

Більш послідовна релятивістська теорія дає вдвічі меншу величину енергії, ніж оцінка (4.71), і тому замість 2 у знаменнику (4.71) необхідно писати коефіцієнт 4. Таким чином гамільтоніан спин-орбітальної взаємодії, отриманий із послідовної релятивістської теорії,

$$H_{s-o} = -\frac{i\hbar^2}{4m^2c^2} \boldsymbol{\sigma} \cdot (\nabla U(\mathbf{r}) \times \nabla). \quad (4.72)$$

У результаті отримуємо, що гамільтоніан електрона у кристалі із врахуванням спин-орбітальної взаємодії має такий вигляд

$$H = H_0 - \frac{i\hbar^2}{4m^2c^2} \boldsymbol{\sigma} \cdot (\nabla U(\mathbf{r}) \times \nabla), \quad (4.73)$$

де

$$H_0 = -\frac{\hbar^2}{2m} \Delta + U(\mathbf{r}). \quad (4.74)$$

Тепер запишемо рівняння на власні значення

$$H(\boldsymbol{\sigma}, \mathbf{r})\Psi(s, \mathbf{r}) = E\Psi(s, \mathbf{r}), \quad (4.75)$$

де  $s = \pm 1/2$  – проекція спіну на вісь OZ. Хвильову функцію шукаємо у вигляді

$$\Psi(s, \mathbf{r}) = \Psi_1(\mathbf{r})v_1(s) + \Psi_2(\mathbf{r})v_2(s), \quad (4.76)$$

де  $v_1(s)$  і  $v_2(s)$  – спінові функції Паулі (спінори), які дорівнюють амплітудам ймовірності того, що спин направлений уздовж або проти вісі OZ

$$v_1(1/2) = 1, v_1(-1/2) = 0, v_2(1/2) = 0, v_2(-1/2) = 1, \quad (4.77)$$

тобто кожен стан тепер буде двічі вироджений за спіном. Гамільтоніан (4.73), діючи на блоховські функції  $\psi_i(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}}u(\mathbf{r})$ , дає

$$\left\{ \frac{\hat{\mathbf{p}}}{2m} + V(\mathbf{r}) + \frac{\hbar}{\mathbf{m}} \mathbf{k} \cdot \hat{\mathbf{p}} + \frac{\hbar}{4m^2c^2} \boldsymbol{\sigma} \cdot (\nabla V \times \hat{\mathbf{p}}) \right\} u(\mathbf{r}) = \left( E_{\mathbf{k}} - \frac{\hbar^2 \mathbf{k}^2}{2m} \right) u(\mathbf{r}). \quad (4.78)$$

Якщо тепер застосовувати  $\mathbf{k}\text{-p}$ -метод, то отримаємо секулярне рівняння типу

$$\left| \frac{\hbar^2}{m^2} \sum_{n,s} \frac{\langle l, j | \mathbf{k} \cdot \mathbf{p} | n, s \rangle \langle n, s | \mathbf{k} \cdot \mathbf{p} | l, j' \rangle}{E_1^0 - E_n^0} - E^{(2)} \delta_{jj'} \right| = 0 \quad (4.79)$$

для визначення доданків до енергії завдяки  $\hbar(\mathbf{k} \cdot \mathbf{p})/m$  доданку в гамільтоніані. Це секулярне рівняння має такий загальний вигляд

$$\left| \mathfrak{R}_{jj'} - E^{(2)} \delta_{jj'} \right| = 0. \quad (4.80)$$

Із врахуванням спіну електрона стану у центрі валентної зони відповідатимуть уже шість хвильових функцій. Якщо ці функції вибрати у вигляді

$$\begin{aligned} Y_{3/2}^{3/2} &= -\frac{1}{\sqrt{2}}(x+iy)v_1, & Y_{3/2}^{1/2} &= -\frac{1}{\sqrt{6}}[(x+iy)v_2 - 2zv_1], \\ Y_{3/2}^{-1/2} &= \frac{1}{\sqrt{6}}[(x-iy)v_1 + 2zv_2], & Y_{3/2}^{-3/2} &= \frac{1}{\sqrt{2}}(x-iy)v_2, \\ Y_{1/2}^{1/2} &= -\frac{1}{\sqrt{3}}[(x+iy)v_2 + zv_1], & Y_{1/2}^{-1/2} &= \frac{1}{\sqrt{3}}[(x-iy)v_1 - zv_2], \end{aligned} \quad (4.81)$$

то вони діагоналізують гамільтоніан спин-орбітальної взаємодії. У результаті секулярне рівняння розпадеться на два із визначниками четвертого та другого порядків

$$\begin{vmatrix} F - \varepsilon & \varepsilon & J & 0 \\ H^* & G - \varepsilon & 0 & J \\ J^* & 0 & G - \varepsilon & -H \\ 0 & J^* & -H^* & F - \varepsilon \end{vmatrix} = 0, \quad \begin{vmatrix} Ak^2 - \Delta - \varepsilon & 0 \\ 0 & Ak^2 - \Delta - \varepsilon \end{vmatrix} = 0, \quad (4.82)$$

$$\text{де } F = Ak^2 + \frac{B}{2}(k^2 - 3k_z^2); \quad G = Ak^2 - \frac{B}{2}(k^2 - 3k_z^2); \quad H = -Dk_z(k_x - ik_y);$$

$$J = -\frac{\sqrt{3}}{2}(B(k_x^2 - k_y^2) + iDk_x k_y); \quad A, B, D - \text{константи.}$$

Розв'язок (4.82) дає дворазово вироджені значення енергії

$$\varepsilon_{1,2} = Ak^2 \pm \sqrt{B^2 k^4 + C^2(k_x^2 k_y^2 + k_x^2 k_z^2 + k_y^2 k_z^2)}, \quad C^2 = D^2 - 3B^2 \quad (4.83)$$

$$i \quad \varepsilon_3 = -\Delta + Ak^2, \quad (4.84)$$

де  $\Delta$  – спін-орбітальне розщеплення рівнів у точці  $\mathbf{k} = 0$ .

Для кремнію  $A = 4,1 \pm 0,2$ ;  $B = 1,6 \pm 0,2$ ;  $C = 3,3 \pm 0,4$ . Ізоенергетичні поверхні являють собою гофровані сфери. Якщо усереднити енергію (4.83) за кутами, то отримуємо для збуджень у валентній зоні

$$\varepsilon_{1,2} = E(0) - \frac{\hbar^2 k^2}{2m} \left[ A \pm \sqrt{B^2 + \frac{C^2}{5}} \right], \quad (4.85)$$

тобто гофр замінюється деякою "середньою" сферичною поверхнею з ефективною масою

$$m_p = \frac{m}{A \pm \sqrt{B^2 + C^2/5}}, \quad (4.86)$$

що свідчить про існування двох типів дірок:

♦ **важких з ефективною масою**

$$m_p^h = \frac{m}{A - \sqrt{B^2 + C^2/5}} \quad (\text{для кремнію } m_p^h = 0,52m); \quad (4.87)$$

♦ **легких з ефективною масою**

$$m_p^l = \frac{m}{A + \sqrt{B^2 + C^2/5}} \quad (\text{для кремнію } m_p^l = 0,16m). \quad (4.88)$$

Для третьої гілки (4.84) ефективна маса є скалярною величиною  $m_p^c = m/A$ ,  $m_p^c = 0,24m$ . Треба зазначити, що завдяки спін-орбітальній взаємодії виродження у центрі зони Бриллюена знімається – одна з гілок ( $\varepsilon_3$ ) зсувається на величину  $\Delta$ . Однак експериментально дірки із середнім значенням ефективної маси не спостерігаються, оскільки відповідна енергія зсунута на величину  $\Delta$  щодо енергій легких і важких дірок. Мінімальна відстань між дном зони провідності та стелею валентної зони відповідає значенням енергії у різних точках зони Бриллюена. Саме ця мінімальна відстань і називається шириною забороненої зони, що визначає плин усіх процесів, які пов'язані із термічним збудженням.

Разом із методом псевдопотенціалу **kp**-метод з урахуванням спин-орбітальної взаємодії є найпоширенішим при обчисленні законів дисперсії елементарних збуджень у напівпровідниках.

## ■ 4.8. Деякі приклади розрахунків зонної структури напівпровідників

Можна стверджувати, що завдяки аналізу численних експериментальних даних і результатів розрахунків зонну структуру найпоширеніших напівпровідників, таких як Ge, Si та GaAs твердо встановлено. Розглянемо особливості їх зонної структури. Почнемо з германію та кремнію.

Електрони кремнію та германію розподілені за станами таким чином:  $\text{Si}^{(14)} \Rightarrow 1s^2 2s^2 2p^6 3s^2 3p^2$  та  $\text{Ge}^{(32)} \Rightarrow 1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^2$ , тобто зовнішня оболонка в них не заповнена. У ній є по два  $p$ -електрони з антипаралельними спінами. Оскільки зона провідності й валентна зона кремнію та германію вміщують  $p$ -стан, виродження для якого в кристалі знімається, то кожна із зон являє собою накладення трьох зон – має три гілки. У кремнії та германії у центрі зони Бриллюена (точка  $\Gamma$ ) з'єднуються дві валентні зони, а при врахуванні спин-орбітальної взаємодії – ще третя, зсунута на величину  $\Delta$  при  $P = 0$  зона. Для різних кристалографічних напрямків залежність енергії від імпульсу має різний вигляд. Як видно з рис. 4.2 та 4.3, одна із гілок зони провідності у цих матеріалів лежить нижче від інших, і саме абсолютний мінімум енергії визначає дно зони провідності. Зазвичай мінімуми енергії називають долинами.

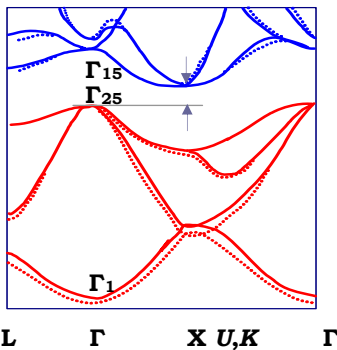


Рис. 4.2. Зонна структура кремнію

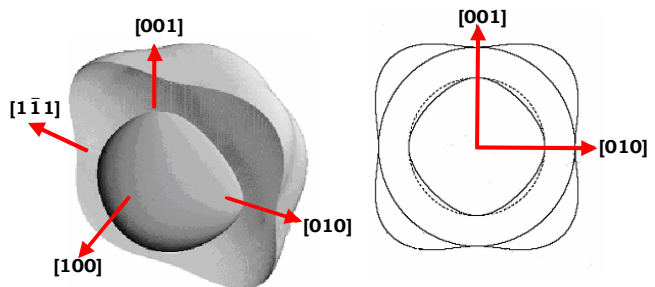


Рис. 4.3. Ізоенергетичні поверхні валентної зони кремнію

Абсолютний мінімум зони провідності у кремнії розташований у напрямку осей [100] (напрямок  $\Gamma \rightarrow X$  на діаграмі рис. 4.2), тобто в кремнії є шість еквівалентних мінімумів, і на першу зону Бриллюена приходить шість еліпсоїдів ізоенергетичних поверхонь (рис. 4.4). Останні поблизу абсолютних мінімумів у кремнії являють собою еліпсоїди обертання щодо великої напіввісі, яка збігається з кристалографічним напрямком [100]. Залежність енергії від квазіімпульсу поблизу мінімуму можна представити у вигляді

$$E(\mathbf{P}) = E(\mathbf{P}_0) + (1/2m_1^*)[(p_x - p_x^0)^2 + (p_y - p_y^0)^2] + (1/2m_3^*)(p_z - p_z^0)^2, \quad (4.89)$$

де  $m_1^* = 0,19m_e$  і  $m_3^* = 0,98m_e$ . Велике значення відношення  $m_3^*/m_1^* = 5,16$  характеризує досить значну анізотропію ізоенергетичних поверхонь (рис. 4.4 а).

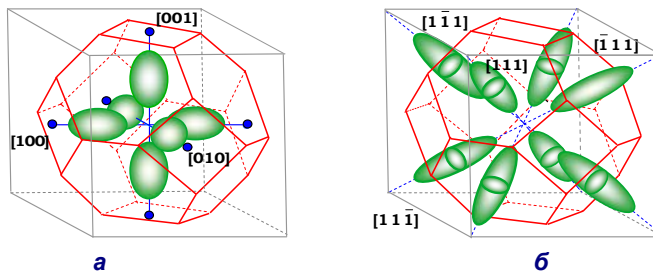


Рис. 4.4. Ізоенергетичні поверхні: а – кремнію, б – германію

Для валентної зони точка максимуму енергії знаходиться у центрі зони Бриллюена. При цьому у точці  $\mathbf{P} = 0$  усі три зони зникаються, тому енергія виявляється виродженою. Як зазначалося вище, за рахунок спин-орбітальної взаємодії виродження частково знімається, і одна із зон зсувається догори у цій точці на величину  $\Delta = 0,035$  еВ. Зв'язок між енергією та компонентами квазіімпульсу для валентної зони поблизу центра зони Бриллюена задається співвідношенням

$$E_{1,2}(\mathbf{P}) = \mathbf{E}(0) - \frac{1}{2m_e} \left[ AP^2 \pm \sqrt{B^2P^4 + C^2(P_x^2P_y^2 + P_y^2P_z^2 + P_z^2P_x^2)} \right], \quad (4.90)$$

де безрозмірні константи  $A = 4,1 \pm 0,2$ ;  $B = 1,6 \pm 0,2$ ;  $C = 3,3 \pm 0,5$ . Ізоенергетичні поверхні являють собою специфічні гофровані поверхні, які можна уявити як деформовані (стиснуті із шести боків) сфери (рис. 4.3). Якщо усереднити вираз для енергії (4.90) за кутами, то можна записати

$$E_{1,2}(\mathbf{P}) = \mathbf{E}(0) - \frac{P^2}{2m_e} \left[ A \pm \sqrt{B^2 + C^2/5} \right]. \quad (4.91)$$

Такою операцією гофрована сфера замінюється деякою усередненою сферою. При цьому тепер збудження кристала можна представити як ізотропний рух частинок з ефективною масою

$$m_p^* = \frac{m_e}{A \pm \sqrt{B^2 + C^2/5}}. \quad (4.92)$$

Це наближення приводить до того, що збудження у валентній зоні представляються двома типами дірок: легкими з ефективною масою  $m_l^* = 0,16m_e$  (відповідає знаку + у (4.92)) і важкими з ефективною масою  $m_h^* = 0,52m_e$ . Урахування спін-орбітальної взаємодії дає третій тип збуджень – дірок з ефективною масою  $m_m^* = 0,24m_e$ . Але експериментально дірки із середнім значенням ефективної маси не спостерігаються, оскільки відповідний закон дисперсії зсунутий догори на величину 0,035 eВ порівняно з енергією легких і важких дірок. Абсолютний мінімум зони провідності у германію розташований у напрямку осей [111] (рис. 4.4 б), тому існує вісім еквівалентних мінімумів енергії. Ізоенергетичні поверхні поблизу таких мінімумів являють собою еліпсоїди обертання щодо великої напіввісі, яка збігається з напрямком [111], і на першу зону Бриллюена приходиться половина кожного еліпсоїда. Іншими словами, у германію є чотири повних еліпсоїди енергії. Вираз для енергії має вигляд, аналогічний виразу для кремнію (див. (4.81)) з ефективними масами  $m_1^* = 0,082m_e$  і  $m_3^* = 1,58m_e$ . Відношення  $m_3^*/m_1^* = 19,3m_e$  майже вчетверо перевищує таке значення для кремнію, що свідчить про більшу анізотропію ізоенергетичних поверхонь у германії. Валентна зона у германії має структуру, аналогічну структурі у кремнії (див. (4.91)), де безрозмірні коефіцієнти мають такі значення:  $A = 13,0 \pm 0,2$ ;  $B = 8,9 \pm 0,1$  і  $C = 10,3 \pm 0,2$ . Ефективні маси легких дірок –  $m_l^* = 0,04m_e$ , а важких –  $m_h^* = 0,34m_e$ . Третя гілка завдяки спін-орбітальній взаємодії опускається порівняно зі смутою важких і легких дірок на величину  $\Delta = 0,28$  eВ (рис. 4.5). Так само, як і в кремнії, експериментально ці дірки не виявляються. Звертає на себе увагу, що екстремуми валентної зони в германії і кремнії не збігаються (за хвильовим вектором) з екстремумами зони провідності, тобто максимуми валентної зони, що знаходяться всередині зони Бриллюена ( $\mathbf{k} = 0$ ), не збігаються з мінімумами зони провідності, розташовані поблизу грані зони Бриллюена. Такі напівпровідники називаються *непрямозонними*.

Арсенід галію належить до напівпровідників типу  $A^{III}B^V$  і має кристалічну структуру типу цинкової обманки, яка відрізняється від структури типу алмазу (до якої належать напівпровідники Ge та Si) тільки тим, що побудована із різних атомів. У цій ґратці атоми сорту А та В



чергуються. Унаслідок цього кристалічне поле в напівпровідниках типу  $A^{III}B^V$  не має центра інверсії, тобто  $U(\mathbf{r}) \neq U(-\mathbf{r})$ . Ізоенергетичні поверхні як валентної, так і зони провідності у цих структурах є сферами. Це є свідченням того, що ефективні маси електронів і дірок в арсеніді галію є скалярними величинами. Валентна зона так само, як і в Ge та Si, складається із трьох енергетичних смуг, одна з яких завдяки спін-орбітальній взаємодії зміщена згори щодо двох інших на величину  $\Delta \approx 0,2$  еВ (рис. 4.6). В арсеніді галію абсолютний мінімум зони провідності (як і максимум валентної зони) лежить у центрі зони Брилюєна при  $\mathbf{k} = 0$ , отже цей напівпровідник належить до класу прямозонних. Ширина забороненої зони складає 1,43 еВ, ефективна маса електронів –  $m_e^* = 0,067m_e$ , а ефективна маса дірок –  $m_p^* = 0,62m_e$ . Укажемо також на наявність локального мінімуму у зоні провідності, що розташований поблизу точки X у напрямку  $[100]$ , характеризується важкими електронами з ефективними масами  $m_{eh}^* = 1,2m_e$  і зсунутий догори щодо енергії основного мінімуму на величину  $\Delta E_g = 0,36$  еВ. Ця особливість зонної структури арсеніду галію є вирішальною при формуванні дуже важливого як з фізичного, так і з прикладного погляду ефекту Ганна, який детально розглянемо нижче.

Нарешті зазначимо, що сучасні експериментальні методи з циклотронного резонансу та поглинання електромагнітного поля в ультрафіолетовому діапазоні дають цілком задовільні результати щодо визначення зонної структури напівпровідників. Незалежно від цього зонні структури напівпровідників обчислювались на сучасному комп'ютерному обладнанні за допомогою ефективних чисельних методів, що базуються на фізичних методах обчислення зонної структури.

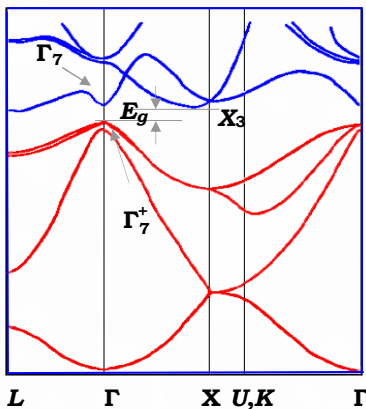


Рис. 4.5. Зонна структура германію

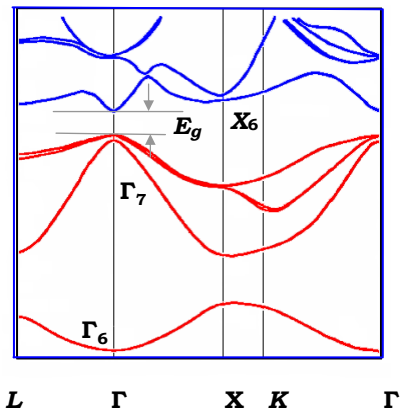


Рис. 4.6. Зонна структура арсеніду галію

Обчислення дали із задовільною точністю такі самі результати для більшості напівпровідників. Таким чином, зонна структура найбільш активно використовуваних напівпровідників у сучасній мікроелектроніці є добре встановленим фактом, що перевірений на практиці.

## ■ 4.9. Задачі

1. Поясніть, чому метод ефективної маси майже завжди є прийнятним для обчислення закону дисперсії напівпровідника і часто не дає коректних результатів для металу.

2. Нехай в одновимірному кристалі із параметром ґратки  $a$  закон дисперсії електрона має вигляд

$$E(k) = \frac{\hbar^2}{a^2 m} (7/8 - \cos ka + (1/8 \cos(2ka))). \quad (4.93)$$

Знайдіть ефективну масу електрона поблизу дна та стелі зони, попередньо зобразивши її графічно.

3. Розглянемо енергетичні зони (рис. 4.7). Нехай  $L$  – центр грані зони Бриллюена, де  $\mathbf{k} = \mathbf{g}/2$ . Тоді можна вважати, що електронні стани описуються хвильовими функціями:

♦ поблизу стелі нижньої зони –  $\psi^{(-)} = \sqrt{\frac{2}{\Omega}} \cos\left(\frac{gz}{2}\right)$ ,

♦ поблизу дна верхньої зони –  $\psi^{(+)} = \sqrt{\frac{2}{\Omega}} \sin\left(\frac{gz}{2}\right)$ .

Використовуючи так звані атомні одиниці (в яких відстань вимірюється у боровських радіусах –  $R_B = 0,529 \text{ \AA}$ , маса – в одиницях електронної маси, а енергія – в атомних одиницях енергії 1 ат.од. = 27,2 еВ), де вважається, що  $\hbar = m = e = 1$ , запишемо енергію, відповідну хвильовій функції  $\Psi^{(-)}$  як  $g^2/8 - W/2$ , а енергію, відповідну стану  $\Psi^{(+)}$ , як  $g^2/8 + W/2$ .

За допомогою  $\mathbf{k}\mathbf{p}$ -методу побудуйте всі компоненти тензору ефективної маси, враховуючи тільки взаємодію між двома зонами.

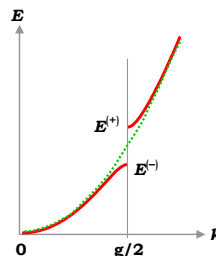


Рис. 4.7

## Список літератури

1. Цидильковский И.М. Электроны и дырки в полупроводниках. – М.: Наука, 1972.
2. Бонч-Бруевич В.Л., Калашников С.Г. Физика полупроводников. – М.: Наука, 1990.
3. Ю Питер, Кардона М. Основы физики полупроводников. – М.: Физматлит, 2002.
4. Korvink J., Greiner A. Semiconductors for Micro- and Nanotechnology: An Introduction for Engineers, Wiley-VCH Verlag GmbH, Weinheim, 2002.
5. Джонс Г. Теория зон Бриллюэна и электронные состояния в кристаллах. – М.: Мир, 1968.

# Розділ 5

## ЕЛЕКТРОН У МАГНІТНОМУ ПОЛІ

Сучасна фізика напівпровідників неможлива без розгляду впливу магнітного поля на електронні властивості напівпровідника. Вивчення взаємодії носіїв заряду в напівпровідниках із магнітним полем, починаючи з класичних гальваноманітних явищ – ефекту Холла та магніторезистивного ефекту й закінчуючи квантовим ефектом Холла та ефектами спин-залежної рекомбінації і перенесення, є важливим аспектом пізнання напівпровідникових матеріалів як з погляду фундаментальної науки, так і для їхнього практичного застосування. Розглянемо деякі проблеми взаємодії електрона в напівпровіднику із зовнішнім магнітним полем.

### ■ 5.1. Магнітний момент і спин електрона

Як відомо з основ квантової механіки, кожен електрон, що входить до складу атома, може певний час перебувати в одному з кількох стаціонарних станів, не випромінюючи електромагнітних хвиль. Кожному стаціонарному стану електрона в атомі відповідає певне значення енергії – енергетичний рівень. Якщо енергетичні рівні занумеровані у порядку зростання енергії, номер рівня  $n$  називають головним квантовим числом. Крім того, квантовий стан системи визначається орбітальним моментом, який є характеристикою просторової симетрії електронних станів. Спрощено його можна інтерпретувати як таку фізичну величину, що визначає форму електронних оболонок атома. Якою б не була ця форма, обертання атома у просторі на кут  $2\pi$  не може змінити його фізичний стан. Це міркування дозволяє дійти висновку, що орбітальний момент може набувати лише дискретних значень, кількість яких обмежена головним квантовим числом

$$|\mathbf{L}| = \hbar\sqrt{l(l+1)}, \quad l=0, 1, 2, \dots, n-1. \quad (5.1)$$

Число  $l$  називають орбітальним квантовим числом. Значення  $l=0$  відповідає електронній оболонці сферичної форми, ненульові значення орбітального квантового числа – оболонкам складнішої форми.

Завдяки наявності орбітального моменту електрон має магнітний момент

$$\mu_l = \mu_B \sqrt{l(l+1)}, \quad (5.2)$$

де  $\mu_B = |e|\hbar(2mc)^{-1}$  – магнетон Бора. Таким чином, величина орбітального моменту є добутком сталої Планка на ціле число, а гіромагнітне відношення  $\gamma_l = -\mu_l/L$  дорівнює  $e(2mc)^{-1}$ . З іншого боку, дослідження

надтонкої структури оптичних спектрів атомів показують, що електрон має характеризуватись також спіновим механічним моментом  $\mathbf{S}$ , причому проекція спіну на виділений напрямок у просторі (вісь квантування –  $OZ$ ) може набувати лише двох значень  $S_z = \pm\hbar/2$ . Оскільки величина спіну не є кратною до сталої Планка, його не можна пов'язати з механічним рухом електрона, коли йдеться про обертання електрона у просторі або про форму електрона. Крім того, спінове "гіромагнітне відношення"

$$\gamma = -\frac{\mu_s}{s\hbar} = \frac{e}{mc} \quad (5.3)$$

удвічі відрізняється від орбітального. Цей факт також свідчить про принципову відмінність між орбітальним і спіновим моментами, тобто спін є специфічною квантовою величиною, і на відміну від орбітального квантового числа спінове квантове число може набувати як цілих, так і напівцілих значень. При цьому частинки, що характеризуються цілим значенням спіну (включаючи нуль), і частинки, що характеризуються напівцільними значеннями спіну (напр., електрон, спін якого дорівнює  $1/2$  – спін вимірюють в одиницях  $\hbar$ ) підлягають різним типам статистики.

## ■ 5.2. Електронні статистичні системи

Фізичні системи, що складаються з великої кількості електронів, називають статистичними електронними системами, властивості яких визначаються принципом заборони Паулі, згідно з яким дві частинки, що характеризуються напівцілим спіном, не можуть перебувати в одному квантовому стані. Якщо статистична система складається з електронів, що входять до складу атомів, принцип Паулі зумовлює певну залежність між квантовими числами електронів одного й того самого атома. Квантові стани електронів, які належать до різних атомів, відрізняються за значеннями просторової координати, тому заборона на квантові числа (головне, орбітальне, спінове) не виникає. У найбільш цікавому для нас випадку (коли система складається з великої кількості колективізованих частинок, напр., електронів у зоні провідності або дірок у валентній зоні) виникає інша ситуація. Частинки, що складають таку систему, відрізнити за значеннями координати неможливо (якщо вони не є локалізованими на домішках), тому вони мають відрізнитись за своїм енергетичним станом. Отже, кількість енергетичних рівнів статистичної системи за порядком величини має дорівнювати кількості електронів  $N$ . У найпростішому випадку (коли

$l = 0$  та  $H = 0$ ) на кожному рівні енергії може знаходитися не більше двох електронів, причому спіни електронів, що знаходяться на одному рівні, відрізняються знаком проекції на вісь квантування. Іншими словами можна сказати, що енергетичні рівні електронів (ферміонів) є двократно виродженими за напрямком спіну. Енергетичний спектр такої електронної системи має складатись із великої кількості ( $N/2$ ) щільно розташованих рівнів  $\varepsilon_i$ .

Статистичну електронну систему, що перебуває у стані теплової рівноваги з термостатом (роль якого в твердому тілі зазвичай виконує ґратка), називають термодинамічною. За нульової температури електрони, що входять до термодинамічної системи, мали б перебувати на найнижчому енергетичному рівні, якби цьому не заважав принцип Паулі. З огляду на цей принцип пари електронів із протилежними спінами займають енергетичні рівні з енергіями від нуля до граничного значення

$$\varepsilon_F = \frac{\hbar^2}{2m^*} \left( 3\pi^2 \frac{N}{V} \right)^{2/3}, \quad (5.4)$$

що називається *енергією Фермі*. Тут  $V$  – об'єм системи,  $m^*$  – ефективна маса колективізованого електрона. Таким чином, за нульової температури кількість електронів на  $i$ -му рівні визначається як

$$n_i = \begin{cases} 2, & \varepsilon_i < \varepsilon_F \\ 0, & \varepsilon_i > \varepsilon_F \end{cases}. \quad (5.4)$$

Оскільки в дійсності температура не є нульовою, тепла енергія електронів обумовлює ненульову ймовірність заповнення рівнів з енергією  $\varepsilon_i > \varepsilon_F$ , а ймовірність заповнення рівнів з енергією  $\varepsilon_i < \varepsilon_F$  стає меншою від одиниці та визначається розподілом Фермі–Дірака

$$f(\varepsilon_i) = \left[ 1 + e^{\frac{\varepsilon_i - \varepsilon_F}{kT}} \right]^{-1}. \quad (5.5)$$

Ураховуючи виродження енергетичних рівнів за напрямком спіну для середньої кількості електронів  $n_i$  на  $i$ -му рівні, маємо

$$n_i = 2f(\varepsilon_i). \quad (5.6)$$

За порядком величини кількість рівнів в енергетичних зонах напівпровідника дорівнює кількості електронів, тому їхня щільність є надзвичайно високою. Зважаючи на це, дискретні значення  $\varepsilon_i$  зазвичай замінюють на неперервну змінну  $E$ , яку називають енергією електронів.

Розглянемо тепер спіновий магнетизм статистичної системи. Сумарний орбітальний і спіновий моменти атома  $\mathbf{L}_a$  та  $\mathbf{S}_a$  складаються із моментів електронів, що входять до його складу. Якщо  $\mathbf{L}_a = \mathbf{S}_a = 0$ , то за

відсутності зовнішнього магнітного поля магнітний момент атома дорівнює нулю. Зовнішнє магнітне поле  $\mathbf{H}$  трохи деформує електронну оболонку атома, у результаті чого виникає ненульовий магнітний момент, спрямований проти поля, тобто момент, що зумовлює діаманітний ефект. У межах уявлень класичної фізики цей ефект можна пояснити за допомогою теореми Лармора, згідно з якою ввімкнення магнітного поля впливає на електродинамічні властивості фізичної системи так само, як її прецесія – навколо напрямку поля з кутовою частотою  $\Omega = |e|H/(2mc)$ . Зважаючи на малу величину діаманітного ефекту, атом із нульовим механічним (орбітальним і спіновим) моментом називають діаманітним або немагнітним.

У разі, коли атом має ненульовий механічний момент, його магнітний момент не дорівнює нулю навіть за відсутності зовнішнього поля. Такий атом називають парамагнітним. Фізичні тіла, до складу яких входять парамагнітні атоми, називають парамагнетиками. Оскільки в нульовому магнітному полі просторова орієнтація магнітних моментів атомів є хаотичною, середня (за координатами та часом) намагніченість парамагнетика дорівнює нулю. Увімкнення зовнішнього магнітного поля викликає парамагнітний ефект, який полягає у виникненні намагніченості, спрямованої за полем. Парамагнітний ефект значно перевищує за величиною діаманітний, тому намагнічуватися проти поля можуть лише ті системи, до складу яких не входять парамагнітні атоми. До того ж величина намагніченості в діаманетиках завжди буде дуже малою. У парамагнетиках діаманітний ефект хоч і присутній, але непомітний на фоні парамагнітного ефекту.

Парамагнетизм, обумовлений електронними спінами, локалізованими поблизу атомів (парамагнетизм Ланжевена), можна охарактеризувати спіноювою поляризацією

$$P = \frac{N_+ - N_-}{N_+ + N_-}, \quad (5.7)$$

де  $N_+$  і  $N_-$  – кількість спінів, спрямованих уздовж і проти магнітного поля, відповідно. Для невироджених систем, згідно із формулою Больцмана,

$$N_{\pm} \propto \exp(\mu_B H / kT). \quad (5.7)$$

У переважній більшості випадків  $\mu_B H \ll kT$  (ця нерівність перестає справджуватися лише за зростання поля до величини  $\sim 10^5$  Е та зменшення температури до величин  $\sim 10$  К). У такому разі

$$P \approx \frac{\mu_B H}{kT}, \quad (5.9)$$

а намагніченість системи можна оцінити за формулою

$$M \propto \mu_B \frac{N}{V} P \propto \frac{N}{V} \frac{\mu_B^2 H}{kT}, \quad (5.10)$$

яка вказує, що намагніченість зменшується зі зростанням температури.

Парамагнетизм колективізованих електронів (парамагнетизм Паулі) описують спіноюю поляризацією

$$P_i = \frac{n_i^+ - n_i^-}{n_i^+ + n_i^-}, \quad (5.11)$$

де  $n_i^+$  і  $n_i^-$  дорівнюють середній кількості електронів зі спінами вздовж і проти поля на  $i$ -му енергетичному рівні. Підставляючи до (5.5) значення енергії  $\varepsilon_i(H) = \varepsilon_i(0) \pm \mu_B H$  із (5.11) у випадку, коли  $\mu_B H \ll kT$ , отримаємо

$$P_i = \left(1 - \frac{n_i}{2}\right) \frac{\mu_B H}{kT}. \quad (5.12)$$

Ця формула показує, що відчутний внесок до поляризації дають лише електрони, розташовані на енергетичних рівнях, ймовірність заповнення яких істотно різниться від одиниці ( $n_i \neq 2$ ). За нульового значення температури такі рівні відсутні, а за ненульових – вони розташовані в енергетичному інтервалі завширшки  $kT$  навколо енергії Фермі. За значень  $kT \ll \varepsilon_F$  кількість електронів  $\Delta N$  на напівзаповнених рівнях лінійно зростає з температурою, тобто  $\Delta N = \text{const } T$ . Завдяки цьому намагніченість

$$M \propto \mu_B \Delta N \frac{\mu_B^2 H}{kT} \sim \frac{\mu_B^3 H}{k}. \quad (5.13)$$

Як уже зазначалось, за полів  $H \leq 10^5 \text{ E}$  і температур, вищих за 10 К, спінова поляризація та намагніченість систем є незначними. Однак існує велика кількість речовин, які намагнічуються до значень  $M \sim \mu_B N/V$  у набагато слабкіших полях (зазвичай менших або порядку  $10^2 \text{ E}$ ) та іноді зберігають намагніченість і після вимкнення зовнішнього магнітного поля. Такі речовини називаються феромагнетиками. Усі феромагнетики переходять у парамагнітний стан, коли температура перевищує певне критичне значення  $T_C$ , що називається температурою Кюрі.

### ■ 5.3. Рівні Ландау

Для з'ясування явищ, що спостерігаються у сильному магнітному полі, розглянемо гамільтоніан частинки, де магнітне поле задається вектор-потенціалом  $\mathbf{A}(\mathbf{r})$ . У цьому випадку, як відомо з класичної механіки, узагальнений імпульс запишеться у вигляді

$$\mathbf{P} = \mathbf{p} + e\mathbf{A}. \quad (5.14)$$

З іншого боку, оскільки кінетична енергія

$$T = \mathbf{p}^2/2m, \quad (5.15)$$

то її можна записати через узагальнений імпульс

$$T = \frac{(\mathbf{P} - e\mathbf{A})^2}{2m}. \quad (5.16)$$

Згідно із принципом відповідності ми маємо замінити всі величини на їхні оператори. Таким чином частина гамільтоніана, що відповідає кінетичній енергії, запишеться як

$$T = \frac{(i\hbar\nabla + e\mathbf{A})^2}{2m}. \quad (5.17)$$

Далі знайдемо явний вигляд оператора  $(i\nabla + \mathbf{A})^2$

$$(i\nabla + \mathbf{A})^2 = -\Delta + i\nabla\mathbf{A} + i\mathbf{A}\nabla + \mathbf{A}^2. \quad (5.18)$$

Оскільки

$$\nabla\mathbf{A} = (\nabla\mathbf{A}) + \mathbf{A}\nabla, \quad (5.19)$$

і ми можемо вибрати таке калібрування, де  $\text{div}\mathbf{A} = 0$ , то

$$(i\nabla + \mathbf{A})^2 = -\nabla^2 + 2i\mathbf{A}\nabla + \mathbf{A}^2. \quad (5.20)$$

Підставимо цей вираз до (5.17), попередньо домноживши  $\Delta$  на  $\hbar^2$ ,  $\nabla$  на  $\hbar$  і  $\mathbf{A}$  – на  $e$ . Маємо

$$T = -\frac{\hbar^2}{2m}\Delta + \frac{e^2\mathbf{A}^2}{2m} + \frac{i\hbar e}{m}(\mathbf{A}\nabla). \quad (5.21)$$

Таким чином, розглядаючи рух вільного електрона у магнітному полі як гамільтоніан, можна використовувати оператор (5.21). Припустимо, що магнітне поле, яке діє на частинку, є однорідним і направлено вздовж вісі OZ системи координат:  $\mathbf{B} = (0, 0, B)$ . У такому випадку

$$\begin{aligned} \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} &= 0, \\ \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} &= 0, \\ \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} &= B. \end{aligned} \quad (5.22)$$

Ці рівняння задовольняються, якщо

$$A_y = A_z = 0; A_x = -yB. \quad (5.23)$$

До того ж, подібний вибір вектор-потенціалу задовольняє умові вибраного калібрування. Рівняння Шредингера тепер матиме вигляд



$$-\frac{\hbar^2}{2m}\Delta\psi - \frac{i\hbar eB}{m}y\frac{\partial\psi}{\partial x} + \frac{e^2B^2y^2}{2m}\psi = E\psi \quad (5.24)$$

та є диференційним рівнянням зі змінними, що розділяються, якщо шукати хвильову функцію у вигляді

$$\psi(x, y, z) = e^{ik_x x} e^{ik_z z} \varphi(y). \quad (5.25)$$

Підставимо (5.25) до (5.24) та отримаємо

$$-\frac{\hbar^2}{2m}\left(-k_x^2 - k_z^2 + \frac{d^2}{dy^2}\right)\varphi(y) + \frac{\hbar eBk_x}{m}y\varphi(y) + \frac{e^2B^2y^2}{2m}\varphi(y) = E\varphi(y). \quad (5.26)$$

Це рівняння можна істотно спростити, якщо припустити, що  $y = y_0 + y'$ . Тоді останні два доданки в лівій частині (5.26) можна записати як

$$\begin{aligned} & \frac{\hbar e k_x B}{m} y' + \frac{\hbar e k_x B}{m} y_0 + \frac{e^2 B^2}{2m} y'^2 + \frac{e^2 B^2}{2m} y_0^2 + \frac{e^2 B^2}{m} y' y_0 = \\ & = \frac{e^2 B^2}{2m} y_0^2 + \frac{e^2 B^2}{2m} y'^2 + \frac{\hbar e k_x B}{m} y_0 + \frac{eB}{m} (\hbar k_x + eB y_0) y'. \end{aligned} \quad (5.27)$$

Вибираючи  $y_0 = -\hbar k_x / eB$ , занулюємо доданок з  $y'$ . Тоді (5.27) запишеться як

$$\frac{e^2 B^2}{2m} y'^2 + \frac{\hbar^2 k_x^2}{2m} - \frac{\hbar^2 k_x^2}{m}. \quad (5.28)$$

Враховуючи ці обставини, маємо

$$-\frac{\hbar^2}{2m}\left(-k_x^2 - k_z^2 + \frac{d^2}{dy^2}\right)\varphi(y') + \left[\frac{e^2 B^2 y'^2}{2m} - \frac{\hbar^2 k_x^2}{2m}\right]\varphi(y') = E\varphi(y'). \quad (5.29)$$

Це рівняння можна переписати у вигляді

$$-\frac{\hbar^2}{2m}\frac{d^2}{dy^2}\varphi(y') + \frac{e^2 B^2 y'^2}{2m}\varphi(y') = \left(E - \frac{\hbar^2 k_z^2}{2m}\right)\varphi(y'). \quad (5.30)$$

Введемо позначення

$$\omega_0 = eB/m, \quad (5.31)$$

$$E' = E - \frac{\hbar^2 k_z^2}{2m}, \quad (5.32)$$

$$\text{перепишемо (5.30) як } -\frac{\hbar^2}{2m}\frac{d^2}{dy^2}\varphi(y') + \frac{m\omega_0^2 y'^2}{2}\varphi(y') = E'\varphi(y'). \quad (5.33)$$

Отримане рівняння є нічим іншим, як рівнянням гармонічного осцилятора з масою  $m$  і частотою  $\omega_0$ . Енергією осцилятора є  $E' = \hbar\omega_0(n + 1/2)$ . Звідси, враховуючи (5.32), маємо

$$E = \frac{\hbar^2 k_z^2}{2m} + \hbar\omega_0 \left( n + \frac{1}{2} \right). \quad (5.34)$$

Доходимо висновку: на рух уздовж поля (уздовж вісі OZ) магнітне поле не впливає. Енергія, пов'язана з рухом уздовж цього напрямку, не квантована, а рух у площині, перпендикулярній полю, є рухом за колом із частотою  $\omega_0$ , що називається циклотронною. Енергія, пов'язана з цим рухом, квантована. Спектр частинки в цьому випадку можна зобразити у вигляді набору квадратичних парабол, які зсунуті вздовж осі енергії на величину  $\hbar\omega_0$  (рис. 5.1). Якщо припустити, що

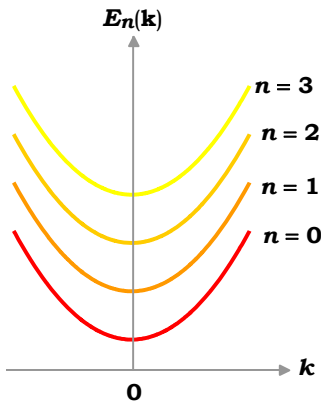


Рис. 5.1. Спектр електрона в магнітному полі

другий доданок у (5.34) пов'язаний із рухом уздовж кругової орбіти, то енергією такого руху є  $E_R = m\omega_0^2 R^2/2$ , а радіус орбіти залежить від поля та є квантованим

$$R = \sqrt{\frac{2\hbar}{m\omega_0} \left( n + \frac{1}{2} \right)} = \sqrt{\frac{2\hbar}{eB} \left( n + \frac{1}{2} \right)}. \quad (5.35)$$

Очевидно, що для руху електрона в кристалі ми отримаємо такі самі результати із заміною маси електрона на ефективну масу. Тоді

$$E(k_z, n, \omega_C) = \frac{\hbar^2 k_z^2}{2m^*} + \hbar\omega_C \left( n + \frac{1}{2} \right) \quad (5.36)$$

із циклотронною частотою  $\omega_C = eB/2m^*$ . (5.37)

Дамо тепер оцінку величини магнітного поля, за якого можливе спостереження ефектів квантування. Ясно, що ефекти, пов'язані з квантуванням енергії, спостерігатимуться лише у випадку, якщо відстань між рівнями перевищує їх ширину, що є наслідком термічного розширення. Іншими словами, для можливості спостереження ефектів квантування в магнітному полі необхідно виконання умови

$$\hbar\omega_C > kT. \quad (5.38)$$

Наприклад, для температури рідкого гелію звідси отримуємо

$$B > 4,2 \cdot 10^4 (m/m^*) [\text{Гс}], \quad (5.39)$$

тобто для спостереження квантування енергії зонного електрона в магнітному полі за температури порядку гелієвої потрібні поля в десятки кілогаусів.

## ■ 5.4. Задачі

1. Для якого з напівпровідників – кремнію чи арсеніду галію – за температури рідкого гелію потрібні менші магнітні поля для спостереження ефектів квантування Ландау.

2. Оцініть величину циклотронної частоти для кремнію та арсеніду галію в магнітному полі величиною 4,2 кГц за температури 4,2 К.

## Список літератури

1. *Киреев П.С.* Физика полупроводников. – М.: Высшая шк., 1975.
2. *Третьяк О.В., Львов В.А., Барабанов О.В.* Фізичні основи спінової електроніки. – К.: ВПЦ "Київський університет", 2002.
3. *Ридли Б.* Квантовые процессы в полупроводниках. – М.: Мир, 1986.
4. *Ансельм А.И.* Введение в теорию полупроводников. – М.: Наука, 1978.



# Розділ 6

## КОЛИВАННЯ АТОМІВ КРИСТАЛІЧНОЇ ҐРАТКИ

Розглядаючи рух електрона в періодичному полі кристалічної ґратки, при введенні поняття квазіімпульсу ми наголошували, що зонний електрон, рухаючись у такому полі, не розсіюється (див. коментар до формули (2.58)), і вся взаємодія електрона із ґраткою зводиться до формування закону дисперсії зонного електрона. У більшості випадків у напівпровідниках ця взаємодія зводиться до введення ефективної маси. Зрозуміло, що ідеальних ґраток у природі не існує, оскільки завжди мають місце різного роду дефекти кристалічної структури, а головне, за ненульових температур атоми в ґратці коливаються, тим самим спотворюючи кристалічний потенціал. Іншими словами, електрони тепер рухаються не в ідеальному періодичному потенціалі. На потенціал накладається коливальний рух атомів ґратки, тому у багатьох випадках розсіювання електронів на такого роду неідеальностях приводить до формування специфічних електронних властивостей напівпровідника. З'ясуємо, яким чином можуть бути описані коливання ґратки, а також ефекти, що можуть стати наслідками взаємодії електронів із коливаннями ґратки.

### ■ 6.1. Одновимірний лінійний ланцюжок атомів

Для вивчення характерних рис коливання атомів у кристалічній ґратці розглянемо просту одновимірну модель – лінійний ланцюжок однакових атомів масою  $m$ , що розташовані один від одного на відстані  $a$ . Нехай  $n$ -й атом змістився на відстань  $u_n$  (рис. 6.1). Якщо це зміщення незначне, то силу взаємодії можна розглядати як пружну, тобто пропорційну зміні відстані між атомами. Для простоти врахуємо лише взаємодію між найближчими сусідами. Тоді сила, що діє на атом  $n$ , дорівнюватиме сумі сил, які діють з боку атомів  $n + 1$  та  $n - 1$ :

$$F_n = -\beta(u_n - u_{n+1}) - \beta(u_n - u_{n-1}) = \beta(u_{n+1} + u_{n-1} - 2u_n), \quad (6.1)$$

де  $\beta$  – коефіцієнти квазіпружної сили. Рівнянням руху  $n$ -го атома (рівнянням Ньютона) є

$$m \frac{d^2 u_n}{dt^2} = \beta(u_{n+1} + u_{n-1} - 2u_n). \quad (6.2)$$

Шукаємо розв'язок цього рівняння у вигляді

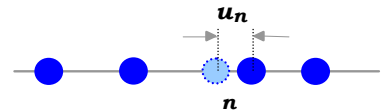


Рис. 6.1. Одновимірний ланцюжок атомів

$$u_n = Ae^{i(qan - \omega t)}, \quad (6.3)$$

де величину  $q$  можна розглядати як хвильовий вектор коливань, що відбуваються із частотою  $\omega$ . Підставимо цей вираз до (6.2)

$$-m\omega^2 = \beta(e^{iqa} + e^{-iqa} - 2) = \beta(2\cos qa - 2) = -2\beta(1 - \cos 2(qa/2))$$

$$\text{або} \quad m\omega^2 = 4\beta\sin^2(qa/2). \quad (6.4)$$

Таким чином із (6.4) отримуємо залежність частоти коливань в одновимірному ланцюжку атомів від хвильового числа (закон дисперсії або дисперсійне співвідношення)

$$\omega = \pm \omega_m \sin(qa/2), \quad (6.5)$$

$$\text{де} \quad \omega_m = 2\sqrt{\beta/m}. \quad (6.6)$$

Оскільки швидкість розповсюдження пружної хвилі вздовж ланцюжка атомів пов'язана із частотою та довжиною хвилі через співвідношення

$$\omega = 2\pi \frac{v}{\lambda}, \quad (6.7)$$

то із (6.5) знайдемо зв'язок швидкості розповсюдження вздовж ланцюжка атомів пружної хвилі із довжиною  $\lambda$

$$v = \frac{\lambda}{\pi} \sqrt{\frac{\beta}{m}} \sin(a\pi/\lambda). \quad (6.8)$$

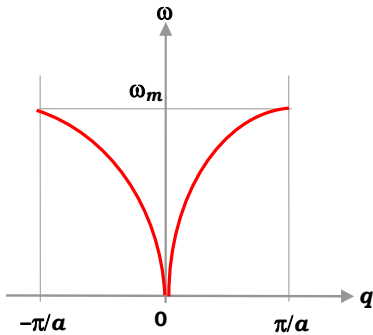
Для визначення меж зміни хвильового числа порівняємо зміщення з рівноважного положення для двох значень хвильового числа –  $q$  та  $q' = q + (2\pi/a)g$  із цілим невід'ємним числом  $g$

$$u_{n,g} = Ae^{i(qan + 2\pi gn - \omega t)} = Ae^{i(qan - \omega t)} e^{2\pi gn} = u_n \quad (6.9)$$

і впевнимися, що хвиля із хвильовим числом  $q$  збігається із хвилею із хвильовим числом, що зсунуте на величину, кратну  $2\pi/a$ . Це означає, що коливання із хвильовими числами  $q$  та  $q' = q + (2\pi/a)g$  характеризують ті самі коливання. Тому можемо характеризувати всі коливання хвильовими числами з інтервалу

$$-\frac{\pi}{a} \leq q \leq \frac{\pi}{a}. \quad (6.10)$$

**Область, де знаходяться всі можливі значення хвильових векторів, збігається із зоною Бриллюена для хвильового вектора  $\mathbf{k}$  для електронів у випадку одновимірного кристала.**



**Рис. 6.2.** Закон дисперсії коливань одновимірного кристала з однакових атомів

Такі коливання, коли при  $q \rightarrow 0$  частота також прямує до нуля, називаються *акустичними*. Закон дисперсії коливань одновимірного ланцюжка з однакових атомів подано на рис. 6.2. Нехай атомний ланцюжок налічує  $N$  атомів. Застосуємо до цих коливань граничні умови Борна–Кармана

$$u_{n \pm N} = u_n. \quad (6.11)$$

Враховуючи розв'язок (6.3), звідси маємо

$$Ae^{i(qa(n \pm N) - \omega t)} = Ae^{i(qan - \omega t)} e^{\pm iqaN}, \quad e^{\pm iqaN} = 1 \quad (6.12)$$

$$\text{або} \quad qaN = 2\pi \cdot g \quad (6.13)$$

$$\text{Звідси знаходимо} \quad q = \frac{2\pi}{aN} g. \quad (6.14)$$

Підставляючи цей вираз до (6.10), маємо

$$-\frac{N}{2} \leq q \leq \frac{N}{2}. \quad (6.15)$$

Таким чином, для атомного ланцюжка, який складається з  $N$  однакових атомів, хвильовий вектор змінюється дискретно в інтервалі значень  $-\pi/a \leq q \leq \pi/a$  і може набувати  $N$  значень – стільки, скільки елементарних комірок міститься у ланцюжку. Для великих значень  $N$  можна вважати, що хвильовий вектор змінюється квазінеперервно.

## ■ 6.2. Довгодіючі сили та метод оберненої ґратки

Раніше ми розглянули коливання одновимірного атомного ланцюжка у випадку, коли між атомами існують тільки короткодіючі сили, й лише взаємодія між найближчими сусідами є суттєвою. У реальних кристалах (крім короткодіючих сил, що зумовлені перекриванням хвильових функцій), існують також довгодіючі (напр., кулонівські). Розглянемо, як відбувається коливання атомів у такому кристалі на прикладі одновимірного ланцюжка, що утворює іонний кристал (ланцюжок із позитивними та негативними зарядами, що чергуються). Потенціальна енергія в такій системі записується у вигляді

$$\varphi = \frac{1}{2} \sum_{i,j} W(x_j - x_i), \quad (6.16)$$

де  $W(x_i - x_j)$  – потенціальна енергія парної взаємодії двох іонів, що розташовані в точках  $x_i$  та  $x_j$  лінійного ланцюжка, відповідно.

Рівняння руху ланцюжка іонів маси  $m$  матиме такий вигляд

$$m \frac{d^2 u_j}{dt^2} = - \sum_i \left( \frac{\partial^2 W}{\partial x_j \partial x_i} \right)_0 u_i, \quad (6.17)$$

де  $u_j$  – зміщення  $j$ -ї частинки з її рівноважного стану. Оскільки енергія  $W$  має описувати рівноважну кристалічну структуру, у стані рівноваги похідні  $\partial W / \partial x_j$  мають дорівнювати нулю для будь-яких  $j$ . Але другі похідні

$$\left( \frac{\partial^2 W}{\partial x_j \partial x_i} \right)_0 \neq 0 \quad (6.18)$$

визначають міжатомні силові сталі. З іншого боку, за зміщення всіх атомів на одну й ту саму величину результуюча сила, що діє на будь-який атом, обертається нулем, тобто має виконуватись умова

$$\sum_i \left( \frac{\partial^2 W}{\partial x_j \partial x_i} \right)_0 = 0. \quad (6.19)$$

Крім того, трансляційна інваріантність кристала вимагає, щоб похідні (6.18) були функціями від  $|x_j - x_i|$ . Загальний розв'язок рівняння (6.17) знов шукаємо у вигляді

$$u_j = A e^{i(qx_j^0 - \omega t)}, \quad (6.20)$$

де  $x_j^0 = ja$  – рівноважна координата  $j$ -го атома. Із (6.17) тепер маємо

$$m\omega^2 = \sum_i \left( \frac{\partial^2 W}{\partial x_j \partial x_i} \right)_0 e^{iq(x_i^0 - x_j^0)}. \quad (6.21)$$

Позначимо праву частину (6.21) через  $D(q)$ , тобто

$$D(q) = \sum_i \left( \frac{\partial^2 W}{\partial x_j \partial x_i} \right)_0 e^{iq(x_i^0 - x_j^0)}.$$

Беручи до уваги рівняння (6.19) (тобто додаючи до (6.21) нуль), запишемо

$$D(q) = - \sum_i \left( \frac{\partial^2 W}{\partial x_j \partial x_i} \right)_0 \left[ 1 - e^{iq(x_i^0 - x_j^0)} \right]. \quad (6.22)$$

Для короткодійчих сил (у наближенні взаємодії найближчих сусідів) можна вважати, що

$$\left( \frac{\partial^2 W}{\partial x_j \partial x_i} \right)_0 = \begin{cases} -f, & |i - j| = 1 \\ 0, & |i - j| \neq 1 \end{cases}, \quad (6.23)$$



тоді із (6.22) маємо результат, який уже отримували раніше

$$D(q) = f(2 - e^{iqa} - e^{-iqa}) = 2f(1 - \cos qa). \quad (6.24)$$

Для системи із довгодіючим парним потенціалом розкладемо функцію  $W(x_j - x_i)$  у ряд Фур'є

$$W(x_j - x_i) = \sum_{q'} W(q') e^{iq'(x_j - x_i)}. \quad (6.25)$$

Тоді силові сталі отримуємо простим диференціюванням

$$\left( \frac{\partial^2 W}{\partial x_j \partial x_i} \right)_0 = \sum_{q'} q'^2 W(q') e^{iq'(x_j^0 - x_i^0)}, \quad (6.26)$$

і (6.22) набуде вигляду

$$D(q) = - \sum_i \sum_{q'} q'^2 W(q') e^{iq'(x_j^0 - x_i^0)} \left[ 1 - e^{iq(x_i^0 - x_j^0)} \right]. \quad (6.27)$$

Перегрупуємо це рівняння

$$D(q) = \sum_q \left\{ \sum_i q'^2 W(q') e^{i(q'-q)(x_j^0 - x_i^0)} - \sum_i W(q') e^{iq'(x_j^0 - x_i^0)} \right\} \quad (6.28)$$

і врауємо тотожність (див. задачу 3 розд. 1)

$$\sum_n e^{ipx_n^0} = N \delta_{p,g}, \quad (6.29)$$

де  $g$  – вектор оберненої ґратки. У результаті маємо

$$\omega^2 = \frac{N}{m} \sum_g (q+g)^2 W(q+g) - \frac{N}{m} \sum_g g^2 W(q). \quad (6.30)$$

Для прикладу розглянемо випадок кулонівського екранованого потенціалу

$$W(g) = \frac{4\pi \cdot (Ze)^2}{\Omega(g^2 + \lambda^2)}, \quad (6.31)$$

де  $\lambda$  – параметр екранування,  $\Omega = Na$  – об'єм кристала. Оскільки за малих  $\lambda$  у (6.30) внесок дає тільки доданок із нульовим вектором оберненої ґратки в першій сумі, то маємо наближено

$$\omega^2 \approx \frac{4\pi \cdot (Ze)^2 N}{m\Omega} \frac{q^2}{q^2 + \lambda^2} + \frac{\omega_p^2}{1 + \lambda^2/q^2}. \quad (6.32)$$

Зауважимо, що за нехтування екрануванням ми отримаємо не коректний результат бездисперсійних коливань ґратки, де частота не наближається до нуля за зменшення хвильового вектора, тобто коли-

вання не будуть акустичними. За  $\lambda \neq 0$  і малих  $q^2$  із (6.32) отримаємо наближено закон дисперсії акустичних коливань у твердому тілі

$$\omega^2 \approx c^2 q^2, \quad c = \omega_p / \lambda. \quad (6.33)$$

Отже врахування ефектів довгодії не викликає принципової відмінності коливань одновимірного ланцюжка однакових атомів, що зумовлені близькодючими силами.

### ■ 6.3. Коливання двоатомного лінійного ланцюжка

Розглянемо коливання одновимірного кристала, в елементарній комірці якого є два атоми з масами  $m_1$  та  $m_2$ . Атоми першого типу знаходяться у вузлах ґратки  $n'$ , а другого – у вузлах  $n''$ . Їх зміщеннями зі стану рівноваги в  $n$ -й комірці є  $u'_n$  та  $u''_n$ , відповідно (рис. 6.3). Припустимо, що коефіцієнти квазіпружної сили між сусідніми атомами однакові й дорівнюють  $\beta$ . Тоді рівняннями руху для атомів обох типів є

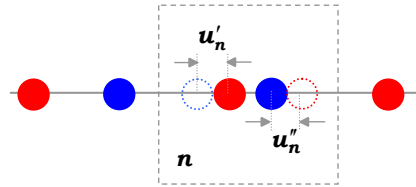


Рис. 6.3. Коливання двоатомного лінійного ланцюжка. Штрихованими кульками позначено місця рівноважної позиції атомів  $n$ -ї комірки

$$m_1 \ddot{u}'_n = \beta(u''_n + u''_{n-1} - 2u'_n), \quad (6.34)$$

$$m_2 \ddot{u}''_n = \beta(u'_n + u'_{n+1} - 2u''_n).$$

Шукаємо розв'язок цієї системи рівнянь у вигляді

$$\begin{aligned} u'_n &= A_1 e^{i(qan - \omega t)}, \\ u''_n &= A_2 e^{i(qan - \omega t)}. \end{aligned} \quad (6.35)$$

Підставимо (6.35) до системи (6.34).

$$\begin{aligned} -\omega^2 m_1 A_1 &= \beta \cdot A_2 (1 + e^{-iqa}) - 2\beta \cdot A_1, \\ -\omega^2 m_2 A_2 &= \beta \cdot A_1 (1 + e^{iqa}) - 2\beta \cdot A_2. \end{aligned} \quad (6.36)$$

Перепишемо цю систему лінійних алгебраїчних рівнянь щодо  $A_n$  у стандартному вигляді

$$\begin{aligned} (2\beta - \omega^2 m_1) A_1 - \beta (1 + e^{-iqa}) A_2 &= 0, \\ -\beta \cdot (1 + e^{iqa}) A_1 + (2\beta - \omega^2 m_2) A_2 &= 0. \end{aligned} \quad (6.37)$$

Для існування ненульових розв'язків цієї системи необхідною умовою є рівність нулю детермінанта

$$\begin{vmatrix} 2\beta - \omega^2 m_1 & -\beta(1 + e^{-iqa}) \\ -\beta(1 + e^{iqa}) & 2\beta - \omega^2 m_2 \end{vmatrix} = 0. \quad (6.38)$$

$$\text{Звідси маємо } (2\beta - \omega^2 m_1)(2\beta - \omega^2 m_2) - \beta^2(1 + e^{-iqa})(1 + e^{qa}) = 0 \quad (6.39)$$

$$\text{або } \omega^4 - 2\beta \frac{m_1 + m_2}{m_1 m_2} \omega^2 + \frac{\beta^2}{m_1 m_2} [4 - (1 + e^{-iqa})(1 + e^{qa})] = 0. \quad (6.40)$$

Зважаючи на те, що вираз у квадратних дужках дорівнює  $4\sin^2(qa/2)$  і вводячи позначення

$$\omega_0^2 = 2\beta \frac{m_1 + m_2}{m_1 m_2} \quad (6.41)$$

$$\text{та } \gamma^2 = 4 \frac{m_1 m_2}{m_1 + m_2}, \quad (6.42)$$

знайдемо розв'язки біквдратного рівняння (6.40)

$$\omega_+^2 = \frac{\omega_0^2}{2} \left[ 1 + \sqrt{1 - \gamma^2 \sin^2(qa/2)} \right], \quad (6.43)$$

$$\omega_-^2 = \frac{\omega_0^2}{2} \left[ 1 - \sqrt{1 - \gamma^2 \sin^2(qa/2)} \right]. \quad (6.44)$$

Отримано нетривіальний результат: кожному значенню  $q$  відповідає два значення частоти, тобто в одновимірному кристалі, у якого в елементарній комірці знаходяться два атоми, існують дві гілки коливань.

Звернемо увагу, що як й в одноатомному ланцюжку розв'язок (6.40) містить корінь  $\omega_-$ , який поблизу точки  $q = 0$  стає пропорційним  $q$  та прямує до нуля. Це коливання називається *акустичним*, оскільки воно є аналогічним довгохвильовому коливанню ланцюжка, що розглядається як пружний континуум. При цьому групова швидкість такого коливання дорівнює швидкості звука у середовищі.

Гілка  $\omega_+$  відповідає оптичному коливанню. Дійсно, при  $q = 0$  підґратки легких і важких атомів рухаються без деформацій у протилежних напрямках. Інакше кажучи, двоатомна молекула в кожній елементарній комірці коливається так, начебто вона незалежна від сусідів. Якщо атоми різного типу мають електричний заряд протилежного знаку (як буває у випадку іонних кристалів), то виникає оптично активний осцилюючий дипольний момент, тобто оптична гілка коливань у кристалах з іонним зв'язком може збуджуватись світлом відповідної частоти.

У граничних випадках найдовших ( $q = 0$ ) і найкоротших ( $q = \pi/a$ ) хвиль із (6.43) і (6.44) випливає, що

$$\omega_-(0) = 0, \quad \omega_-(\pi/a) = \frac{\omega_0}{\sqrt{2}} \sqrt{1 - \sqrt{1 - \gamma^2}}, \quad (6.45)$$

$$\omega_+(0) = \omega_0, \quad \omega_+(\pi/a) = \frac{\omega_0}{\sqrt{2}} \sqrt{1 + \sqrt{1 - \gamma^2}}. \quad (6.46)$$

Звідси можна визначити ієрархію граничних частот

$$\omega_+(0) > \omega_+(\pi/2) > \omega_-(\pi/2) > \omega_-(0). \quad (6.47)$$

Використовуючи визначення (6.35) і (6.37) для випадку  $\omega_-(0) = 0$ , маємо

$$\begin{pmatrix} u'_n \\ u''_n \end{pmatrix}_- = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}_- = 1, \quad (6.48)$$

тобто всі атоми мають одне й те саме зміщення. Це означає, що в нескінченно довгій хвилі акустичної гілки коливань атоми рухаються синхронно – комірочки зміщуються згруповано, як одне ціле.

Нескінченно довгій хвилі оптичних коливань відповідає частота  $\omega_0$ . Із (6.35) та (6.37) отримуємо, що

$$\begin{pmatrix} u'_n \\ u''_n \end{pmatrix}_+ = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}_+ = \frac{2\beta}{2\beta - m_1\omega_0^2} = \frac{1}{1 - (m_1 + m_2)/m_2} = -\frac{m_2}{m_1}. \quad (6.49)$$

Ці співвідношення означають, що для нескінченно довгої хвилі оптичної моди коливань атоми у комірці зміщуються у протилежних напрямках – коливаються у протифазі. При цьому центр тяжіння кожної пари атомів лишається нерухомим:  $m_1 u'_n + m_2 u''_n = 0$ . Закон дисперсії коливань одновимірної ґратки із двома атомами на елементарну комірочку показано на рис. 6.4. Треба звернути увагу на наявність енергетичної щілини між акустичною та оптичною гілками коливань.

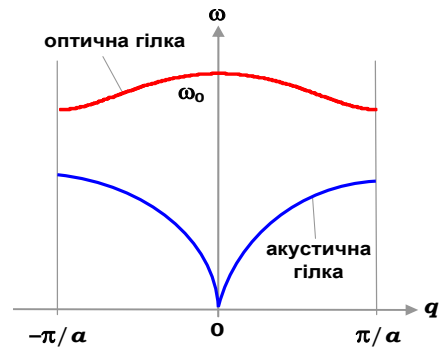


Рис. 6.4. Закон дисперсії коливань одновимірної ґратки із двома атомами в комірці

## ■ 6.4. Коливання атомів кристала. Фонони

Зміщення атомів кристалічної тривимірної ґратки можна описати як суперпозицію нормальних коливань ґратки. Кожне нормальне коливання являє собою хвилю, напрямком розповсюдження якої визначається

хвильовим вектором  $\mathbf{q}$ . Останній має стільки значень, скільки елементарних комірок існує у кристалі. Значення  $\mathbf{q}$  змінюються у межах

$$-\pi \leq q_i a_i \leq \pi, \quad (6.50)$$

де  $a_i$  – стала ґратки ( $i = x, y, z$ ). Кристал містить  $N = N_x N_y N_z$  елементарних комірок, у кожній з яких може бути  $s$  атомів. Кожний атом визначається трьома ступенями свободи, тобто кристал характеризуватиметься  $3sN$  ступенями свободи. Це означає, що в кристалі існуватиме  $3sN$  хвиль. У тривимірному кристалі можуть існувати два типи пружних коливань:

*повздовжні* – хвилі стиснень і розріджень та *поперечні* – хвилі деформацій зсуву. В останніх зсув атомів ґратки відбувається в площині, перпендикулярній до напрямку розповсюдження хвилі, і може мати два головних напрямки поляризації. Таким чином, у кристалі  $s$ -типу коливання будуть повздовжніми, а  $2s$ -типу – поперечними. Кожному значенню хвильового вектора  $\mathbf{q}$  відповідає  $3s$  коливань, що відрізняються частотою та поляризацією. Три гілки, для яких у довгохвильовому наближенні групова та фазова швидкості дорівнюють одна одній  $v_{zp} = v_{\phi} = v_{ze}$  ( $v_{ze}$  – швидкість звуку у кристалі), будуть акустичними. Інші ( $3s-3$ ) гілки, для яких при  $q \rightarrow 0$ ,  $v_{ph} \rightarrow \infty$ , а  $v_{zp} \rightarrow 0$ , будуть оптичними.

Кожне із  $3sN$  нормальних коливань можна уявити як гармонічний осцилятор (тобто повна енергія коливань кристала дорівнює сумі енергій коливань  $3sN$  гармонічних осциляторів, що не взаємодіють між собою). Якщо  $q_j$  осцилятору, що коливається із частотою  $\omega_j(\mathbf{q})$ , приписати енергію  $E_{qj} = \hbar \omega_j(\mathbf{q})(n_{qj} + 1/2)$ , то повна енергія коливань кристала

$$E = \sum_{qj} \hbar \omega_j(\mathbf{q}) \left( n_{qj} + \frac{1}{2} \right), \quad n_{qj} = 0, 1, 2, \dots \quad (6.51)$$

де  $\mathbf{q}$  – хвильовий вектор, що набуває  $N$  дозволених значень. Число  $n_{qj}$  має сенс квантового числа осцилятора, яке в свою чергу може набувати цілих значень, тобто мінімальна зміна енергії може відбутися при  $\Delta n_{qj} = \pm 1$  і складати  $\hbar \omega_j(\mathbf{q})$ . Мінімальна порція енергії, яку може поглинути ( $\Delta n_{qj} = 1$ ) або випромінювати ( $\Delta n_{qj} = -1$ ) ґратка, називається *фононом*. Його можна розглядати як квазічастинку зі всіма притаманними частинкам властивостями: вони можуть взаємодіяти один з

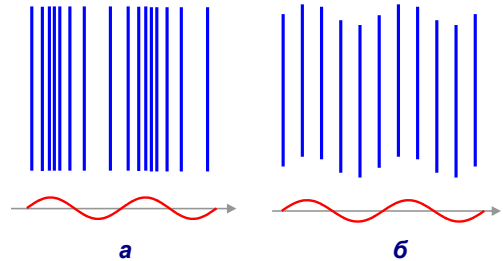


Рис. 6.5. Схема виникнення коливань ґратки кристала:

**а** – повздовжніх, **б** – поперечних.

Лініями показано атомні площини. Видно, що повздовжні коливання є хвилями стиснень і розтягнень, а поперечні – хвилями деформацій зсуву

одним (розсіювання фонона на фонони), з електронами (розсіювання електрона на фонони) та дірками. За невеликого ангармонізму їх можна розглядати як фононний газ: при  $\Delta n_{qj} = 1$  маємо процес породження фонона, а при  $\Delta n_{qj} = -1$  – процес зникнення. Важливим аспектом введення квазічастинок власних коливань ґратки – фононів – є той факт, що такі частинки є бозонами.

## ■ 6.5. Статистика фононів

Середнє число фононів з енергією  $\hbar\omega_j(\mathbf{q})$ , що існують у даному нормальному коливанні за температури  $T$ , виражається через середню енергію осцилятора  $\langle E_{qj} \rangle$  формулою

$$\langle n_{qj} \rangle = \frac{\langle E_{qj} \rangle}{\hbar\omega_j(\mathbf{q})}. \quad (6.52)$$

Статистична фізика стверджує, що ймовірність того, що осцилятор у тепловій рівновазі за температури  $T$  знаходиться в стані  $s$  з енергією  $E_s$ , складає

$$W_s = C e^{-E_s/kT}, \quad (6.53)$$

де нормуючий коефіцієнт визначається з умови нормування ймовірності на одиницю

$$\sum_s W_s = 1, \quad C = \left( \sum_s e^{-E_s/kT} \right)^{-1}. \quad (6.54)$$

Тоді середня (рівноважна) енергія дорівнює сумі енергій станів, помножених на відповідну ймовірність знаходження фонона у даному стані

$$\langle E \rangle = \sum_n E_{s(n)} W_{s(n)}. \quad (6.55)$$

Підставимо вирази для ймовірності знаходження фонона в  $s$ -му стані

$$\langle E \rangle = \frac{\sum_{n=0}^{\infty} E_s e^{-E_s/kT}}{\sum_{n=0}^{\infty} e^{-E_s/kT}} = - \frac{d}{d(1/kT)} \ln \left( \sum_{n=0}^{\infty} e^{-E_s/kT} \right). \quad (6.56)$$

Враховуючи, що енергією гармонічного осцилятора, який характеризується квантовим числом  $n$ , є  $E_s = \hbar\omega(n + 1/2)$ , обчислимо статистичну суму як суму геометричної прогресії

$$\sum_{n=0}^{\infty} e^{-E_s/kT} = \sum_{n=0}^{\infty} e^{-\hbar\omega(n+1/2)/kT} = \sum_{n=0}^{\infty} e^{-\hbar\omega/2kT} (e^{-\hbar\omega/kT})^n = \frac{e^{-\hbar\omega/2kT}}{1 - e^{-\hbar\omega/kT}}. \quad (6.57)$$

Підставимо цю формулу до (6.56)

$$\langle E \rangle = -\frac{d}{d(1/kT)} \ln \frac{e^{-(\hbar\omega/2)/kT}}{1 - e^{-\hbar\omega/kT}} = \frac{\hbar\omega}{2} + \frac{d}{d(1/kT)} \ln(1 - e^{-\hbar\omega/kT}) = \frac{\hbar\omega}{2} + \frac{\hbar\omega e^{-\hbar\omega/kT}}{1 - e^{-\hbar\omega/kT}}.$$

Звідси маємо 
$$\langle E \rangle = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}. \quad (6.58)$$

Перший доданок у формулі не залежить від температури й називається нульовою енергією осцилятора. Оскільки нас у подальшому цікавлять процеси взаємодії електронів із фононами, то енергію нульових коливань можна не враховувати – у процесах розсіювання на електронах беруть участь фонони, збуджені вище рівня нульових коливань. Таким чином середня енергія фононів кристалічної ґратки за температури  $T$  задається формулою

$$\langle E \rangle = \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}. \quad (6.59)$$

Тоді середня щільність фононів, що характеризуються середньою енергією  $\langle E \rangle$ ,

$$\langle n \rangle = \frac{\langle E \rangle}{\hbar\omega} = \frac{1}{e^{\hbar\omega/kT} - 1}. \quad (6.60)$$

Ми впевнились, що фонони підлягають статистиці Бозе–Ейнштейна. Формула (6.60) означає, що  $\langle n \rangle$  є рівноважним середнім числом фононів

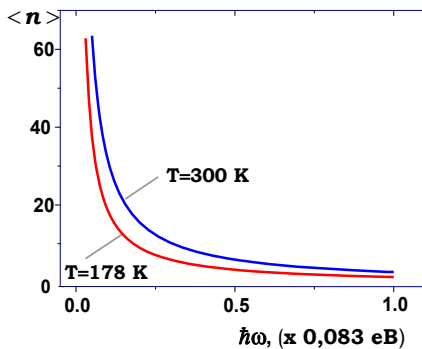


Рис. 6.6. Функція розподілу фононів

з енергією  $\hbar\omega$  в одному квантовому стані – комірці фазового простору з об'ємом  $\hbar^3$ . Поведінку функції  $\langle n \rangle$  подано на рис. 6.6. Середнє число фононів із малими енергіями є дуже великим і швидко зменшується зі зростанням енергії. За високих температур  $kT \gg \hbar\omega$  число фононів в одній комірці фазового простору складає:

$$\langle n \rangle \approx kT/\hbar\omega. \quad (6.61)$$

При цьому середня енергія осцилятора

$$\langle E \rangle = \langle n \rangle \hbar\omega = kT, \quad (6.62)$$

тобто за високих температур фонони у твердому тілі підлягають класичній статистиці.

За низьких температур  $kT < \hbar\omega$  середнє число фононів

$$\langle n \rangle = e^{-\hbar\omega/kT}, \quad (6.63)$$

а середня енергія коливань кристалічної ґратки

$$\langle E \rangle = \hbar\omega \cdot e^{-\hbar\omega/kT}. \quad (6.64)$$

Оскільки фонони є бозонами, то за взаємодії із зовнішніми полями або електронами їхня кількість може змінюватись. Зауважимо, що у твердому тілі можливі два типи коливань (акустичні та оптичні). Завдяки тому, що розкладення цих коливань за нормальними модами еквівалентне введенню фононів, існують також два типи фононів – оптичні та акустичні. За малих значень імпульсу  $q$  енергія акустичних фононів є значно меншою від енергії оптичних фононів. Це дуже важливий факт, адже за майже пружного розсіювання електронів на коливаннях ґратки зазвичай відбувається мала передача імпульсу. Це означає, що в основному електрони розсіюються на акустичних фононах. За низьких температур ( $kT \ll \hbar\omega_0$ ) оптичні фонони не збуджуються – у кристалі існують тільки акустичні фонони.

## ■ 6.6. Електрон-фононна взаємодія

Очевидно, що коливання ґратки мають впливати на поведінку електронів у напівпровіднику. Наприклад, повздовжні коливання іонів викликають неоднорідний розподіл зарядів, який залежить від часу. Заряди, відповідним чином екрановані, створюють потенціал, залежність якого від координат має той самий вигляд, що й залежність від координат амплітуди коливань ґратки. Цей потенціал, зрозуміло, входить до повного гамільтоніану електронів і визначає взаємодію між коливаннями ґратки та електронами.

Простіше за все розглянути електрон-фононну взаємодію в іонних кристалах, де позитивні та негативні іони в повздовжніх модах за великих довжин хвиль зміщуються у протилежних напрямках. Для спрощення позначень припустимо, що амплітуди коливань позитивних і негативних іонів однакові. Це припущення є виправданим, оскільки до остаточної відповіді входить лише різниця векторів цих амплітуд. Нехай зміщення позитивно та негативно заряджених іонів визначаються як

$$\delta \mathbf{r}_i^+ = \frac{\mathbf{u}}{\sqrt{N}} e^{i(\mathbf{q} \cdot \mathbf{r}_i^+ - \omega t)} \quad (6.65)$$

$$i \quad \delta \mathbf{r}_i^- = \frac{\mathbf{u}}{\sqrt{N}} e^{i(\mathbf{q} \cdot \mathbf{r}_i^- - \omega t)}. \quad (6.66)$$



Для повздовжньої моди вектори  $\mathbf{u}$  та  $\mathbf{q}$  паралельні. За великих довжин хвиль різницею координат іонів, що входять до тієї самої примітивної комірки, можна знехтувати та записати дипольний момент у вигляді  $(2Ze/\sqrt{N})\mathbf{u}\exp[i(\mathbf{q}\cdot\mathbf{r}_i - \omega t)]$ , де  $Ze$  – заряд кожного іона. Таким чином, за великих довжин хвиль поляризація (величина дипольного моменту в одиниці об'єму) визначається формулою

$$\mathbf{P} = \frac{2Ze}{\Omega_c} \frac{\mathbf{u}}{\sqrt{N}} e^{i(\mathbf{q}\cdot\mathbf{r} - \omega t)}, \quad (6.67)$$

де  $\Omega_c$  – об'єм комірки. Густина локального заряду дорівнює дивергенції поляризації зі знаком мінус. Якщо відомо густину заряду, то за допомогою рівняння Пуассона можна обчислити електростатичну потенціальну енергію. У результаті для електронного потенціалу отримуємо

$$V(\mathbf{r}) = i \frac{8\pi Ze^2}{q^2 \Omega_c \sqrt{N}} \mathbf{q} \cdot \mathbf{u} e^{i(\mathbf{q}\cdot\mathbf{r} - \omega t)}. \quad (6.68)$$

Цей потенціал описує електрон-фононну взаємодію. Зазначимо, що він прямує до нескінченності при  $q \rightarrow 0$ . У цьому випадку відбувається так званий сильний зв'язок, що викликає сильні поляронні ефекти. Цієї особливості не виникає для поперечних оптичних коливань, коли накопичування заряду не відбувається.

У неполярних напівпровідниках, наприклад германії та кремнії, ефектів поляризації немає, і для встановлення вигляду електрон-фононної взаємодії потрібно зважати на те, що основну роль тут відіграють довгохвильові моди коливань. Вільні електрони та дірки переважно зосереджуються в невеликих областях простору хвильових векторів, що розташовані поблизу краю зони, і розсіювання електронів переважно супроводжується випромінюванням довгохвильових фонових. Взаємодію із повздовжніми модами зазвичай описують за допомогою потенціалу деформації. Як відомо із детальних розрахунків зонної структури, якщо кристал розтягувати так, що відносна зміна його об'єму визначиться об'ємним розширенням  $\Delta$ , то дно зони провідності підніметься на величину  $\delta V_C$ , а стеля валентної зони незначно (на величину  $\delta V_V$ ) зсунеться у бік зони провідності. Ці енергетичні зсуви залежать від об'ємного розширення  $\Delta$  лінійно

$$\delta V_C = D_C \Delta, \quad \delta V_V = D_V \Delta. \quad (6.69)$$

Іншими словами, енергія електрона, що знаходиться на дні зони провідності, зсувається на величину, пропорційну об'ємному розширенню  $\Delta$ . Припустимо тепер, що завдяки взаємодії між електроном і ґраткою просторово-неоднорідне розширення визначає ефективний потенціал,

який діє на електрон і залежить від його координати. Це означає, що рівняння (6.69) повинні мати вигляд

$$\delta V_C(\mathbf{r}) = D_C \Delta(\mathbf{r}), \quad \delta V_V(\mathbf{r}) = D_V \Delta(\mathbf{r}). \quad (6.70)$$

Зрозуміло, що ефективні потенціали, які діють на електрони, розташовані не в екстремумах зони, дещо відрізняються від потенціалів, які діють на електрони точно в екстремальних точках. Але оскільки ми маємо справу з напівпровідниками, де основна кількість частинок, що визначають електронні властивості, знаходиться поблизу екстремумів, то можемо використовувати значення констант, обчислених в екстремумах зон. Величина констант потенціалу деформації має порядок ширини забороненої зони. Для полярних кристалів об'ємне розширення дорівнює величині  $iq(u/\sqrt{N})\exp[i(qr - \omega t)]$ , а згідно із (6.68) роль константи  $D$  відіграє величина  $(8\pi Ze^2)(\Omega_C q^2)$ . У випадку неполярних кристалів довгохвильова особливість не виникає.

Формулу (6.70) можна узагальнити, записавши деформацію у вигляді

$$\varepsilon_{ij} = \frac{1}{2} \left( \frac{\partial \delta_i}{\partial x_j} + \frac{\partial \delta_j}{\partial x_i} \right), \quad (6.71)$$

де  $\partial \delta_i / \partial x_j$  – зміна зміщення вздовж  $i$ -ї вісі за зміни відстані вздовж  $j$ -ї вісі. Тоді потенціал деформації набуде вигляду

$$\delta V(\mathbf{k}, \mathbf{r}) = \sum_{i,j} D_{ij}(\mathbf{k}) \varepsilon_{ij}(\mathbf{r}). \quad (6.72)$$

Розглянемо найпростіший випадок взаємодії із повздовжніми акустичними коливаннями. Оскільки акустичні хвилі характеризуються малими частотами, то можна точно обчислити їхній внесок до розсіювання електронів. Для цього зафіксуємо в уяві положення атомів у деформованому кристалі та обчислимо розсіювання електронів, вважаючи ці викривлення дефектом кристалу. Процес розсіювання електрона під час його переходу зі стану з хвильовим вектором  $\mathbf{k}$  до стану  $\mathbf{k}'$  визначається Фур'є-компонентою потенціалу деформації, що відповідає хвильовому вектору  $\mathbf{q} = \mathbf{k}' - \mathbf{k}$ . Величина Фур'є-компоненти потенціалу, у свою чергу, залежить лише від коливань ґратки із хвильовими векторами  $\pm \mathbf{q}$ , якщо зміщення ґратки записати у вигляді

$$\frac{\mathbf{u}}{\sqrt{N}} e^{i(\mathbf{q}\mathbf{r} - \omega t)} + \frac{\mathbf{u}^*}{\sqrt{N}} e^{-i(\mathbf{q}\mathbf{r} - \omega t)}.$$

Частота таких переходів пропорційна квадрату потенціалу деформації, тобто квадрату об'ємного розширення. Енергія відповідної моди також

пропорційна квадрату об'ємного розширення, за високої температури вона має бути пропорційною  $kT$ .

Процеси взаємодії електронів із коливаннями ґратки зручно розглядати як процеси взаємного розсіювання електронів і фононів. У такому випадку квантування коливань ґратки зводить процеси взаємодії електронів із коливаннями ґратки до процесів розсіювання частинок зі всіма наслідками, що витікають звідси. Наприклад, при розгляді процесів такого розсіювання, як послідовність елементарних актів електрон-фононої взаємодії, зручно використовувати діаграмні методи, основна ідея яких полягає в тому, що формальне введення графічних позначень приводить до істотного полегшення визначення амплітуд складних процесів розсіювання як суми елементарних. Позначимо через  $\mathbf{k}$  електрон, що характеризується енергією  $E(\mathbf{k})$  та імпульсом  $\mathbf{k}$ , а через  $\mathbf{q}$  – фонон, що характеризується енергією  $\hbar\omega(\mathbf{q})$  та імпульсом  $\mathbf{q}$ . Процеси розсіювання електрона на фононі із випромінюванням і поглинанням фонона позначатимуться діаграмами (рис. 6.7). У таких актах елементарного розсіювання виконується закон збереження імпульсу та енергії, тобто у діаграмі (а) маємо

$$\mathbf{k} = \mathbf{k}' + \mathbf{q} \quad \text{та} \quad E(\mathbf{k}) = E(\mathbf{k}' + \mathbf{q}) + \hbar\omega(\mathbf{q}).$$

Поданий процес (рис. 6.7 б) характеризується наступними законом збереження імпульсу та енергії:

$$\mathbf{k} + \mathbf{q} = \mathbf{k}' \quad \text{та} \quad E(\mathbf{k}) + \hbar\omega(\mathbf{q}) = E(\mathbf{k} + \mathbf{q}).$$

Оцінимо максимальну енергію, яку може мати фонон у процесі розсіювання електрона із випромінюванням фонона (рис. 6.7 а). Таку саму енергію може передати ґратка електрону за розсіювання електрона із поглинанням фонона (рис. 6.7 б). Зазначимо, що за розсіювання електрон може отримати максимальний імпульс, якщо за цього процесу абсолютна величина імпульсу збережеться, а напрямок – зміниться на протилежний (аналог лобового зіткнення із частинкою великої маси). Тоді  $\mathbf{k} - \mathbf{k}' = 2\mathbf{k}$ , тобто максимальне значення імпульсу фонона за розсіювання електрона на фононі  $|\mathbf{q}| = 2|\mathbf{k}|$ . Оскільки закон дисперсії акустичного фонона для невеликих значень імпульсу  $\omega = c_s q$  ( $c_s$  – швидкість звука в кристалі), то максимальна енергія фонона, який бере участь у процесі розсіювання електрона,  $E_{ph}^{(\max)} = 2\hbar|\mathbf{k}|c_s$ . Таким чином, відносна зміна енергії електрона за розсіювання на коливаннях ґратки не може перевищувати величину

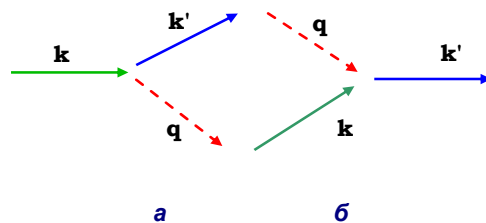


Рис. 6.7. Діаграми електрон-фононої взаємодії:

а – розсіювання з еманациєю фонона;  
б – розсіювання із поглинанням фонона

$$\frac{\delta E}{E} = \frac{E_{ph}^{(\max)}}{E(\mathbf{k})}. \quad (6.73)$$

Оскільки  $\hbar k = m^* V$  та енергія електрона  $E = m^* V^2/2$ , то із (6.69) маємо

$$\frac{\delta E}{E} = 4 \frac{c_v}{V}. \quad (6.74)$$

Отже відносна зміна енергії електрона за розсіювання на акустичних коливаннях ґратки складає за порядком величини відношення швидкості звуку в кристалі до швидкості електрона. Швидкість звуку в кристалі  $c_s \sim 10^5$  см·с<sup>-1</sup>, а швидкість електронів за кімнатної температури  $V \sim 10^7$  см·с<sup>-1</sup>, що дає

$$\frac{\delta E}{E} \sim 10^{-2}. \quad (6.75)$$

Таким чином, процеси розсіювання електронів на коливаннях ґратки за кімнатної температури можна з великою точністю вважати пружними. За низьких температур, коли швидкість електронів у валентній зоні є малою, відносна зміна енергії електрона за розсіювання на коливаннях ґратки матиме порядок величини енергії фонона.

Крім подібних процесів розсіювання можуть відбуватися переходи електронів між виродженими мінімумами у зоні провідності. Такі процеси називаються міждолинним розсіюванням.

## ■ 6.7. Задачі

1. В яких напівпровідниках існують, крім акустичних, оптичні гілки коливань ґратки – Si, Ge, GaAs, InSb?

2. Розгляньте лінійний ланцюжок атомів із виходом за межі врахування взаємодії тільки між атомами – найближчими сусідами. Вважайте, що пружну силу, яка діє на  $n$ -й атом, можна записати у вигляді

$$F_n = \sum_p \beta_p (u_{n+p} + u_{n-p} - 2u_n), \quad (6.76)$$

де  $u_n = u(na)$ ,  $a$  – стала ґратки (порівн. із (6.1)). Покажіть, що закон дисперсії коливань ґратки тепер має вигляд

$$\omega = (2/\sqrt{m}) \sqrt{\sum_{p>0} \beta_p \sin^2(pka/2)}, \quad (6.77)$$

а у довгохвильовому наближенні за умов збіжності ряду  $\sum_{p>0} p^2 \beta_p$  цей закон

переходить у такий

$$\omega = \frac{a}{\sqrt{m}} |k| \sqrt{\sum_{p>0} p^2 \beta_p}. \quad (6.78)$$

**3.** Знайдіть закон дисперсії довгохвильових коливань одновимірного одноатомного ланцюжка атомів у наближенні із попередньої задачі, і за припущення, що  $\beta_p = A p^{-\gamma}$ . Зокрема, розгляньте випадки  $1 < \gamma < 3$ ,  $\gamma = 3$  та  $\gamma = 1$ .

**Указівка.** У цьому випадку ряд під коренем (6.78) розходиться, що не дає можливості скористатися розкладенням синуса, але завдяки малим значенням хвильового числа  $k$  можна перейти від підсумовування до інтегрування.

### Список літератури

1. *Маделунг О.* Теория твердого тела. – М.: Наука, 1980.
2. *Смит Р.* Полупроводники. – М.: Мир, 1982.
3. *Рейсленд Дж.* Физика фононов. – М.: Мир, 1975.
4. *Борн М., Хуанг-Кунь.* Динамическая теория кристаллических решеток. – М.: ИЛ, 1958.
5. *Ансельм А.И.* Введение в теорию полупроводников. – М.: Наука, 1978.
6. *Фейнман Р.* Статистическая механика. – М.: Мир, 1978.



## Розділ 7

# ДОМІШКИ ТА ДЕФЕКТИ В НАПІВПРОВІДНИКАХ

У попередніх розділах ми розглянули основні ідеї зонної теорії ідеальних кристалів, усі вузли ґратки (що безперервно простягаються в усіх напрямках) яких заповнені тільки атомами певної речовини. Однак існуючі технології не надають можливості отримання ідеального кристала. Усі кристали тією чи іншою мірою характеризуються різноманітними недосконаlostями, що викривляють строгу періодичність ідеальної ґратки. Крім того, домішки та дефекти у напівпровідниках відіграють вирішальну, а часто й фатальну роль у формуванні їхніх властивостей. Наприклад, один з авторів (О.В.Т.) у 80-х роках був учасником такого повчального випадку. Під час державного приймання новозбудованого заводу для автоматизованого виробництва МОП інтегральних схем як вихідний матеріал використовувались кремнієві пластини виробництва однієї із західноєвропейських фірм. Після проходження всього технологічного ланцюжка, що включав кілька десятків операцій, вихід стандартної продукції складав близько 70 %, що відповідало всім тогочасним нормам. За спроби використати як вихідний матеріал кремнієві пластини вітчизняного виробництва, що за паспортом (тип провідності, концентрація домішок  $10^{15} \text{ см}^{-3}$ , концентрація дислокацій) точно відповідали іноземному аналогу, відсоток стандартної продукції ледь досяг кількох одиниць. Вивчення причин такої ситуації показало, що вітчизняні кремнієві пластини характеризувались концентрацією неконтрольованих домішок, на порядок вищою ( $\sim 10^{12} - 10^{13} \text{ см}^{-3}$ ) за європейський аналог ( $\sim 10^{10} - 10^{11} \text{ см}^{-3}$ ). Технологічний цикл виробництва мікросхем містив процес відпалювання за температур 700 °С. Виявилось, що за таких температур неконтрольовані дефекти призводили до значного зростання концентрації дислокацій та інших дефектів (т. зв. глибоких центрів). І якщо зростання концентрації таких дефектів в іноземному зразку не викликало фатальних наслідків, то на порядок вища концентрація неконтрольованих дефектів у вітчизняному – призводила до повної непридатності запропонованої технології для виготовлення інтегральних схем на базі вітчизняного кремнію. На щастя з'ясувалось, що зниження температури відпалювання на кілька десятків градусів, що припускалось технологічними умовами, не викликало значного зростання концентрації неконтрольованих дефектів. У результаті виробництво мікросхем на вітчизняній базі виявилось успішним. Цей випадок свідчить про суттєву роль домішок і дефектів у фізиці напівпровідників та є одним із прикладів, що спонукають зважати на наявність недосконалостей напівпровідникових кристалів.

Таким чином, питання про роль дефектів і домішок у твердому тілі, зокрема у напівпровідниках, виявилось основним у процесі впровадження напівпровідникових технологій. Фактично відповіді на це питання були присвячені основні зусилля науковців, які протягом десяти років розвивали науку й технологію напівпровідників. Більш того, можна стверджувати, що завдяки вмінню керувати властивостями домішок і дефектів у напівпровідниках і виникла сучасна напівпровідникова електроніка.

Зрозуміло, що існування домішок вплине на формування характеристик кристала, отже приведе до відмінності його властивостей від властивостей ідеального кристала. Перед аналізом впливу дефектів на електронні властивості напівпровідника коротко розглянемо класифікацію дефектів кристалічної ґратки.

В умовах термодинамічної рівноваги навіть в ідеальному кристалі існують відхилення від точного розташування атомів у вузлах ґратки Браве. Такі відхилення називаються дефектами кристалічної ґратки. Рівноважні дефекти – це такі дефекти кристала, що є елементарними збудженнями основного стану кристала. До цього типу дефектів належать фонони – колективні коливання атомів кристалічної ґратки. У реальних кристалах, крім рівноважних дефектів ґратки, існують нерівноважні дефекти, викликані переважним чином неідеальними умовами росту та існування кристала. Часто нерівноважні дефекти кристалічної ґратки стабілізуються електричними, магнітними або пружними полями, що виникають у кристалі під час його росту, фазових перетворень або зовнішніх дій, наприклад радіаційного опромінювання. Таким чином з'являється можливість суттєвого зменшення густини нерівноважних дефектів за рахунок удосконалення методів вирощування та обробки кристалів.

Можливість існування дефектів із великою енергією утворення в умовах термодинамічної рівноваги за температури  $T$  можна пояснити тим, що утворення дефектів викликає збільшення ентропії кристала у цілому щодо ідеального кристала. Нехай кристал має  $N$  однакових атомів, і в ньому утворюються  $n$  вакансій. Видалити із кристала, що містить  $N$  атомів,  $n$  атомів, і тим самим утворити  $n$  вакансій можна

$$C_N^n = \frac{N!}{n!(N-n)!} \quad (7.1)$$

способами. Тоді, згідно із формулою Больцмана, відповідне збільшення ентропії кристала виражатиметься величиною

$$\Delta S = k \ln \frac{N!}{n!(N-n)!} . \quad (7.2)$$



Якщо енергією утворення одного дефекту є  $E_d$ , то утворення  $n$  дефектів за температури  $T$  змінить вільну енергію кристала на величину

$$\Delta F = nE_d - T\Delta S. \quad (7.3)$$

Оскільки числа  $N$  та  $n$  є дуже великими для оцінки логарифма, від факторіала можна використовувати формулу Стирлінга

$$\ln N! = M \ln N. \quad (7.4)$$

Підставляючи вираз для ентропії до правої частини (7.3) і використовуючи формулу Стирлінга, отримуємо для вільної енергії

$$\Delta F = nE_d - kT\{M \ln N - n \ln n - (N - n) \ln (N - n)\}. \quad (7.5)$$

Параметр  $n$ , за якого система буде стабільною, можна знайти як розв'язок рівняння  $\partial \Delta F / \partial n = 0$ , тобто

$$\ln \frac{n}{N - n} = -\frac{E_d}{kT}. \quad (7.6)$$

Звідси, вважаючи, що  $N \gg n$ , отримуємо число рівноважних дефектів з енергіями утворення  $E_d$ , які можуть існувати у напівпровіднику

$$n = N e^{-E_d/kT}. \quad (7.7)$$

Аналогічно можна розглянути іншу ситуацію, коли дефекти мають складнішу природу. Наприклад, дефект, що складається з атома у міжвузлів'ї та вакансії (т. зв. дефект Френкеля – див. далі), може бути утворений, якщо в кристалі існує деяка кількість вакансій (нехай їхнє число є  $N_1$ ) та атомів у міжвузлів'ях ( $N_2$ ). Якщо енергією утворення такого дефекту є  $E_{Fr}$ , то рівноважна кількість таких дефектів у кристалі

$$n_{Fr} = \sqrt{N_1 N_2} e^{-E_{Fr}/kT}. \quad (7.8)$$

У кристалах може існувати багато типів дефектів, і всі вони можуть відігравати значну роль у формуванні властивостей напівпровідника, тому їхня класифікація є важливим завданням технології. Завдяки малій рухливості та великому часу життя дефекти кристалічної ґратки можуть бути класифіковані за геометричними ознаками.

## ■ 7.1. Типи дефектів у напівпровідниках

У реальних кристалах регулярне розташування атомів у кристалічній ґратці зберігається у досить обмежених областях. На кожній ділянці кристала з лінійними розмірами  $\sim 0,1\text{--}1,0$  мкм регулярне розташу-

вання атомів у вузлах переривається точковим або протяжним дефектом: точкові дефекти зазвичай пов'язуються з окремими атомами в локальних областях кристала; протяжні – можна поділити на *лінійні* (одновимірні) та *двовимірні* (поверхневі). У цьому сенсі й поверхню, якою закінчується кристал, можна розглядати як двовимірний дефект. До двовимірних дефектів також можна зарахувати межі розподілу різних фаз, межі блоків, зерен, доменів тощо. До тривимірних дефектів належать пори, включення тощо.

Точкові дефекти додатково класифікують таким чином:

- ♦ **вакансія** – дефект, викликаний відсутністю атома у вузлі;
- ♦ **атом у міжвузлів'ї** – дефект, обумовлений атомом (власної речовини або домішки), що займає місце між вузлами ґратки;
- ♦ **атом заміщення** – дефект, за якого атом домішки заміщує атом власної речовини у вузлі;
- ♦ **антисайт** – особливий тип дефекту заміщення, за якого власний атом А у сполучі займає місце іншого власного атома В;
- ♦ **дефект Френкеля**, який розглядають як точковий, що являє собою комплекс із вакансії та міжвузлів'я, тобто власний атом, зсунутий із вузла в напрямку найближчого міжвузлів'я.

Вакансії та антисайти є власними дефектами, що не пов'язані з наявністю сторонніх атомів у кристалі. Їхню концентрацію не можна визначити методами хімічного аналізу або мас-спектрометрії. Вони не є електрично активними і не викликають зміни концентрації носіїв. Але контролювання таких дефектів є дуже важливим, оскільки вони можуть визначати поведінку кристалів під час різноманітних технологічних операцій. Тому під час вирощування кристалів прагнуть домогтися, щоб концентрація таких дефектів не перебільшувала величину  $\sim 10^{10}\text{--}10^{12}\text{ см}^{-3}$ .

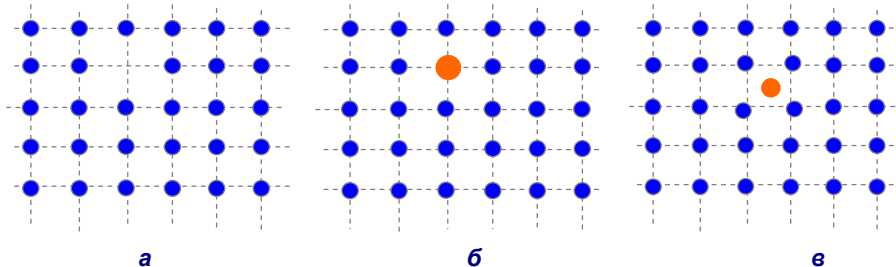


Рис. 7.1. Точкові дефекти: а – вакансія; б – атом заміщення; в – атом у міжвузлів'ї

Схематично точкові дефекти подано на рис. 7.1 та 7.2. Схеми не показують процес релаксації ґратки в околі дефекту, хоча зрозуміло, що в невеликому околі дефекту ґратка зазнає деяких деформацій, і недосконалої насправді займають великі ділянки ґратки. Характерним

представником неточкових дефектів є одновимірні дефекти – дислокації, і двовимірні – площини. На рис. 7.3 схематично подано такі дефекти. Тут показано плоский зріз кристала із дефектом. Дислокація є лінією краю напівплощини, що вставлена в кристал (приклад т. зв. крайової дислокації).

Двійникуванням називають утворення в монокристалі областей зі змінною орієнтацією кристалічної ґратки, що відбувається шляхом дзеркального відбиття вихідного кристала в площині двійникування. На рис. 7.3 б ця площина розташована нормально до поверхні рисунка та проходить через лінію АВ. Лінійні та плоскі дефекти майже завжди погіршують роботу електронних приладів, тому при виготовленні напівпровідникових пластин для подальшого застосування в електронній промисловості намагаються відбирати пластини з мінімальною кількістю лінійних і плоских дефектів.

Домішки є дефектами, що спеціально вводяться (або з'являються неконтрольовано) до напівпровідникових матеріалів з метою впливу на їхні електронні характеристики. Таким чином виникає проблема визначення енергетичних станів домішок у напівпровідниках.

## ■ 7.2. Енергетичні стани домішкових напівпровідників

Відомо, що в ідеальному кристалі утворюються зони заборонених і дозволених значень енергії. У зоні заборонених значень енергії не може існувати жоден дозволений стан електронів. Реальні ж кристали відрізняються від ідеальних наявністю дефектів кристалічної ґратки, зокрема домішок. При цьому енергії електронних станів, що локалі-

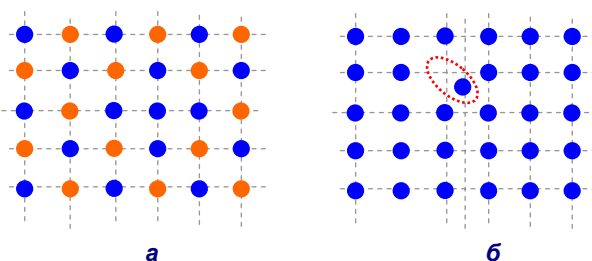


Рис. 7.2. Точкові дефекти:  
а – антисайт, б – Френкеля

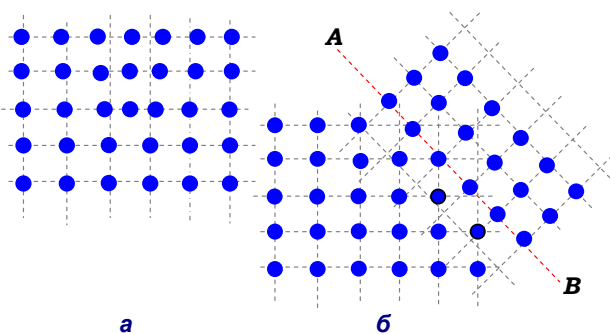


Рис. 7.3. Неточкові дефекти:  
а – дислокація; б – площина двійникування

зовані на домішкових рівнях, можуть лежати всередині забороненої зони та мати однакову величину для домішок одного типу. Дійсно, наявність домішок означає, що, крім кристалічного потенціалу, на електрон у кристалі діє потенціал, обумовлений сильним локальним збуренням поля домішки. Розглянемо загальну задачу про електрон у кристалі, на який накладено сильне локальне збурення. Нехай поле цього збурення локалізоване в області із центром у точці  $\mathbf{R}_i - W(\mathbf{r} - \mathbf{R}_i)$ . Рух електрона в такому випадку описується рівнянням Шредингера

$$\left( -\frac{\hbar^2}{2m} \Delta + U(\mathbf{r}) + W(\mathbf{r} - \mathbf{R}_i) \right) \psi(\mathbf{r}) = E \psi(\mathbf{r}). \quad (7.9)$$

Припустимо, що розв'язок цього рівняння для ідеального кристала, де  $W(\mathbf{r} - \mathbf{R}_i) = 0$ , відомий. Розкладемо невідому хвильову функцію в (7.9) за блохівськими функціями ідеального кристала

$$\psi(\mathbf{r}) = \sum_{\mathbf{k}} c_{\mathbf{k}} \psi_{\mathbf{k}}(\mathbf{r}). \quad (7.10)$$

Підставляючи це розкладання до (7.9), отримуємо

$$\sum_{\mathbf{k}} c_{\mathbf{k}} (E_0(\mathbf{k}) + W(\mathbf{r} - \mathbf{R}_i)) \psi_{\mathbf{k}}(\mathbf{r}) = E \sum_{\mathbf{k}} c_{\mathbf{k}} \psi_{\mathbf{k}}(\mathbf{r}). \quad (7.11)$$

Подіємо на це рівняння оператором  $I_{\mathbf{k}'}$  =  $\int d\mathbf{r} (\psi_{\mathbf{k}'}(\mathbf{r}))^* \dots$ . Із врахуванням ортонормованості хвильових функцій ідеального кристала  $\psi_{\mathbf{k}}(\mathbf{r})$  отримаємо рівняння для визначення коефіцієнтів розкладання (7.10)

$$[E - E_0(\mathbf{k}')] c_{\mathbf{k}'} - \sum_{\mathbf{k}} W_{\mathbf{k}'\mathbf{k}} c_{\mathbf{k}} = 0. \quad (7.12)$$

Оскільки збурення є малим усюди, крім малої області  $V_R$  в околі точки  $\mathbf{R}_i$ , то інтегрування при обчисленні матричних елементів  $W_{\mathbf{k}\mathbf{k}}$  можна проводити за малим об'ємом із центром у точці  $\mathbf{R}_i$ :

$$W_{\mathbf{k}\mathbf{k}} \equiv (\psi_{\mathbf{k}'}(\mathbf{R}_i))^* \psi_{\mathbf{k}}(\mathbf{R}_i) \int_{V_R} d\mathbf{r} W(\mathbf{r} - \mathbf{R}_i) = (\psi_{\mathbf{k}'}(\mathbf{R}_i))^* \psi_{\mathbf{k}}(\mathbf{R}_i) W_0 V_R, \quad (7.13)$$

де

$$W_0 = \frac{1}{V_R} \int_{V_R} d\mathbf{r} W(\mathbf{r} - \mathbf{R}_i) \quad (7.14)$$

– середнє значення енергії за об'ємом  $V_R$ . Суттєвим є те, що ця величина не залежить від координати точки  $\mathbf{R}_i$  локалізації дефекту. Підставимо (7.13) до (7.12)

$$[E - E_0(\mathbf{k}')] c_{\mathbf{k}'} - (\psi_{\mathbf{k}'}(\mathbf{R}))^* W_0 V_R \sum_{\mathbf{k}} c_{\mathbf{k}} \psi_{\mathbf{k}}(\mathbf{R}_i) = 0. \quad (7.15)$$

Враховуючи, що сума у другому доданку є хвильовою функцією електрона в неідеальному кристалі, обчислена в точці  $\mathbf{R}_i$ , отримаємо

$$c_{\mathbf{k}'} = \frac{(\psi_{\mathbf{k}'}(\mathbf{R}_i))^* W_0 V_R \psi(\mathbf{R}_i)}{E - E_0(\mathbf{k}')} . \quad (7.16)$$

Підставимо цей розв'язок до (7.15) і скоротимо отриманий вираз на  $(\psi_{\mathbf{k}'}(\mathbf{R}_i))^* W_0 V_R \psi(\mathbf{R}_i)$ . Маємо

$$1 - \sum_{\mathbf{k}} \frac{(\psi_{\mathbf{k}}(\mathbf{R}_i))^* \psi_{\mathbf{k}}(\mathbf{R}_i) W_0 V_R}{E - E(\mathbf{k})} = 0 \quad (7.17)$$

або

$$\frac{1}{W_0 V_R} = \sum_{\mathbf{k}} \frac{A_{\mathbf{k}}}{E - E(\mathbf{k})} , \quad (7.18)$$

де

$$A_{\mathbf{k}} = (\psi_{\mathbf{k}}(\mathbf{R}_i))^* \psi_{\mathbf{k}}(\mathbf{R}_i) .$$

Проаналізуємо цю формулу, яка є алгебраїчним рівнянням  $N$ -го ступеня щодо  $E$ , оскільки хвильовий вектор набуває  $N$  значень, де  $N$  – кількість елементарних комірок у кристалі. Введемо допоміжну функцію

$$\varphi(E) = \sum_{n=1}^N \frac{A_n}{E - E_n} , \quad (7.19)$$

де  $A_n = A_{\mathbf{k}} = |\psi_{\mathbf{k}}(\mathbf{R}_i)|^2$ ,  $E_n = E(\mathbf{k})$ , і позначимо  $B = 1/W_0 V_R$ . Тоді всі значення енергії електрона неідеального кристала можна знайти як розв'язки рівняння

$$\varphi(E) = B. \quad (7.20)$$

$B \rightarrow \infty$  при  $W_0 V_R \rightarrow 0$  і, оскільки  $A_n$  є константою, що не залежить від  $W_0 V_R$ , знаменник хоча б одного із доданків суми в (7.19) має бути нескінченно малою величиною. Таким чином за нескінченно малого значення збурення розв'язок "збуреної" задачі переходить у розв'язок "незбуреної". Функцію  $\varphi(E)$  можна представити як частку двох поліномів

$$P_1(E) = \sum_{j=1}^N A_j \prod_{n \neq j}^{N-1} (E - E_n) , \quad (7.21)$$

$$P_2(E) = \prod_{n=1}^N (E - E_n) \quad (7.21)$$

або

$$\varphi(E) = \frac{P_1(E)}{P_2(E)} . \quad (7.23)$$

Тут  $P_1(E)$  є поліномом  $N-1$  ступеня, а  $P_2(E)$  – поліномом  $N$  ступеня. Поведінку функції  $\varphi(E)$  подано на рис. 7.4. Вона має  $N-1$  вузлів, де  $\varphi(E) = 0$  і  $N$  точок, де  $\varphi(E) = \infty$ . Зрозуміло, що розв'язками (7.20) будуть усі точки пе-

ретину прямої  $\varphi = B$  і функції  $\varphi(E)$ . Таким чином, за відсутності дефекту ( $B \rightarrow \infty$ ) розв'язок (7.20) дає  $N$  особливих точок, що відповідають  $N$  значенням енергії у зоні, ширина якої –  $E_B$ . Ці розв'язки подано на рисунку чорними точками.

Наявність дефекту приводить до того, що від зони, значення  $N - 1$  рівнів в якій майже не відрізняються від вихідних  $E_n = E(\mathbf{k})$  ідеального кристала (оскільки  $N$  дуже велике), відщеплюється один рівень. Розв'язки (7.20) для неідеального напівпровідника показано зеленими точками. При цьому, якщо збурення  $W_0 > 0$ , то рівень відщеплюється від максимального значення енергії у зоні. Так формується, наприклад рівень акцепторної домішки в напівпровіднику. Якщо ж збурення  $W_0 < 0$  ( $B < 0$ ), то рівень відщеплюється від мінімального рівня енергії у зоні. За такою схемою, формується, наприклад рівень донорної домішки у напівпровіднику. Треба зауважити, що величина енергії рівня дефекту  $E_{loc}$  залежить тільки від типу дефекту, оскільки за великої кількості дефектів (7.20) буде записане для кожного із дефектів, що локалізовані у точка  $\mathbf{R}_i$  і характеризуватиметься потенціалом  $W(\mathbf{r} - \mathbf{R}_i)$ . Іншими словами:

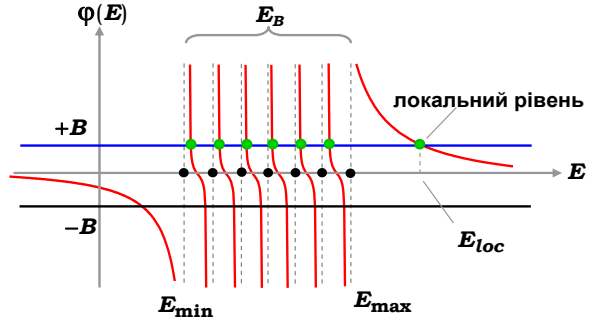


Рис. 7.4. Графічний розв'язок рівняння (7.20) для визначення енергії рівня на дефекті

**При накладанні на кристал локального збурення достатньої сили один із підрівнів енергії незбуреної системи відщеплюється від сукупності рівнів (зони) вниз від мінімальної енергії або зсувається вище від максимальної енергії дозволеної зони. Інші  $N-1$  рівнів майже не змінюють свого положення. При  $W_0 V_R > 0$  догори піднімається рівень максимального значення енергії, а при  $W_0 V_R < 0$  згори опускається рівень мінімальної енергії. Таким чином у забороненій зоні з'являється дозволений енергетичний стан.**

З'ясуємо тепер, чи локалізований такий стан у просторі? Нехай енергія такого стану  $E_{loc} > 0$ . Підставляючи значення для коефіцієнтів розкладення хвильової функції (7.16) до (7.10), маємо хвильову функцію такого стану

$$\psi_L(\mathbf{r}) = \psi_{loc}(\mathbf{R}) W_0 V_R \sum_{\mathbf{k}} \frac{(\psi_{\mathbf{k}}(\mathbf{R}))^*}{E_{loc} - E(\mathbf{k})} \psi_{\mathbf{k}}(\mathbf{r}). \quad (7.24)$$

Оскільки при  $W_0 V_R \rightarrow \infty$ ,  $E_{loc} / W_0 V_R \sim 1$ , то  $E_{loc} \gg E(\mathbf{k})$  і

$$\psi_{loc}(\mathbf{r}) \approx \psi_{loc}(\mathbf{R}) \sum_{\mathbf{k}} (\psi_{\mathbf{k}}(\mathbf{R}))^* \psi_{\mathbf{k}}(\mathbf{r}) = \psi_{loc}(\mathbf{R}) \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{R}} (\varphi_{\mathbf{k}}(\mathbf{R}))^* \psi_{\mathbf{k}}(\mathbf{r}). \quad (7.25)$$

Для оцінки суми у правій частині (7.25) зверху замінимо функцію  $(\varphi_{\mathbf{k}}(\mathbf{R}))^*$  деякою її максимальною величиною, незалежною від хвильового числа  $-\varphi^*(\mathbf{R})$ . Тоді

$$\Psi_{loc}(\mathbf{r}) < \Psi_{loc}(\mathbf{R})\varphi^*(\mathbf{R})\sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{R}}\Psi_{\mathbf{k}}(\mathbf{r}). \quad (7.26)$$

Праву частину цієї нерівності зручно переписати через так звану функцію Ваньє

$$\Phi(\mathbf{r}-\mathbf{R}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{R}}\Psi_{\mathbf{k}}(\mathbf{r}), \quad (7.27)$$

тобто

$$\Psi_{loc}(\mathbf{r}) < \Psi_{loc}(\mathbf{R})\varphi^*(\mathbf{R})\sqrt{N}\Phi(\mathbf{r}-\mathbf{R}). \quad (7.28)$$

Функція Ваньє відмінна від нуля лише в невеликій околиці точки  $\mathbf{R}$ . Таким чином, хвильова функція (7.25) пропорційна функції Ваньє. Електрон з енергією  $E_{loc}$  локалізований в деякому об'ємі в області збурення, й за обмеженої величини  $W_0 V_R$  чим меншою є енергія збурення, тим більшою буде область локалізації.

Таким чином, ми показали, що за невеликої концентрації дефектів одного типу у забороненій зоні напівпровідника виникає один дозволений енергетичний стан, локалізований у просторі поблизу дефектів, причому область локалізації є тим меншою, чим більшою є енергія збурення, яку привносить домішка до кристалу. Донорний рівень відщеплюється від дна зони провідності та розташовується неподалік від  $E_C$  з енергією  $E_d$ . Акцепторний рівень відщеплюється від стелі валентної зони  $E_V$  з енергією  $E_a$ . Зрозуміло, що енергії  $E_i^{(d)} = E_C - E_d$  та  $E_i^{(a)} = E_a - E_V$  будуть енергіями іонізації донорів та акцепторів, відповідно.

### ■ 7.3. Елементарна теорія дефектів

У сучасній мікроелектроніці найбільшу роль відіграють точкові дефекти: легуючі домішки та рекомбінаційні центри. Розглянемо рух електрона в слабкому полі потенціалу (мілкої) домішки. Необхідно врахувати, що: по-перше, у кристалі електрон характеризується ефективною масою; по-друге, його взаємодія з ядром буде ослаблена внаслідок екрануючих властивостей кристалічної ґратки. Формально цей факт виражається наявністю діелектричної проникності, що відрізняється від одиниці (напр., для кремнію  $\epsilon = 12$ ). Вважатимемо, що домішка є донором, тобто

$$U(r) = -\frac{Ze^2}{\epsilon r} \quad (7.29)$$

За початок відліку енергії зручно приймати дно зони провідності  $E_C$ . Рівняння Шредингера матиме вигляд

$$-\frac{\hbar^2}{2m^*} \nabla^2 \Psi - \frac{Ze^2}{\epsilon r} \Psi = E \Psi. \quad (7.30)$$

Розв'язок цього рівняння знаходимо за схемою, аналогічною для атома водню. Як результат отримуємо, що власні значення енергії електрона

$$E_n = E_C - \frac{m^* Z^2 e^4}{\hbar^2 \epsilon^2} \frac{1}{n^2}. \quad (7.31)$$

Для аналізу цієї формули зручно ввести енергію іонізації атома водню

$$E_I^{(H)} = \frac{me^4}{\hbar^2}. \quad (7.32)$$

Тоді рівні енергії електрона донорної домішки виражатимуться через енергію іонізації атома водню

$$E_n = E_C - \frac{1}{n^2} \frac{Z^2}{\epsilon^2} \left( \frac{m^*}{m} \right) E_I^{(H)}. \quad (7.33)$$

Таким чином, рівень енергії основного стану донорної домішки лежить у забороненій зоні нижче дна зони провідності на величину енергії іонізації донорної домішки

$$E_I^{(D)} = \frac{Z^2}{\epsilon^2} \left( \frac{m^*}{m} \right) E_I^{(H)}. \quad (7.34)$$

Як видно з цієї формули, енергія іонізації одноразово іонізованої донорної домішки в  $\epsilon^2$  разів менша за енергію іонізації атома водню. З іншого боку видно, що енергія дворазово іонізованої домішки лежить нижче рівня енергії одноразово іонізованої домішки. Оцінка величини енергії іонізації п'ятивалентних атомів домішок у германії, де  $m^* = 0,25 m$ , а діелектрична стала  $\epsilon = 16$ , дає величину  $\sim 10^{-2}$  еВ. Радіус орбіти електрона домішки у воднеподібній моделі виражається формулою

$$a = a_0 \left( \frac{m}{m^*} \right) n^2 \epsilon, \quad (7.35)$$

де 
$$a_0 = \frac{\hbar^2}{e^2 m^*} \quad (7.36)$$

є радіусом першої борівської орбіти атома водню ( $a_0 = 0,053$  нм). Згідно із (7.35) радіус першої електронної орбіти одноразово іонізованої домішки  $a = 3,4$  нм. Оскільки стала ґратки германію  $b = 0,56$  нм, то легко побачити, що локалізований на домішці електрон "розмазаний"



по досить великій області кристалу. Це означає, що рух електрона можна розглядати як рух вільної частинки із масою  $m^*$  у суцільному середовищі з діелектричною константою  $\epsilon$ , тобто припущення, зроблені при виведенні (7.13), є обґрунтованими.

Аналогічно можна розглянути випадок, коли домішка є акцепторною, тобто у домішкового атома не вистачає електрона для заповнення зв'язку атома напівпровідника, що розташований поруч із домішковим атомом. Розв'язок задачі про рух електрона в кулоновському полі позитивно зарядженого іона дозволяє записати енергію рівнів акцепторної домішки в напівпровіднику

$$E_p = E_v + \frac{Z^2}{\epsilon^2 n^2} \left( \frac{m^*}{m} \right) E_I^{(H)}. \quad (7.37)$$

Отже рівень акцепторної домішки лежить у забороненій зоні вище стелі валентної зони.

У літературі часто зустрічається термін **глибокий центр**, якому, однак, важко дати точне визначення. Якщо за глибокі центри вважати домішки, енергетичні рівні яких лежать поблизу середини забороненої зони, то виникає проблема, пов'язана з існуванням дефектів, властивості яких схожі на властивості глибоких центрів, енергії яких розташовані неподалік від країв зони провідності або валентної зони. Для розуміння причини утворення таких рівнів можна припустити, що в деяких атомах домішок електрони слабко взаємодіють з атомами основної речовини, орбіта електрона атома домішки має малий радіус, а потенціал є екрануючим

$$V(r) = -\frac{Ae^{-ar}}{r}. \quad (7.38)$$

Екрануючий потенціал на малих відстанях є великим, а за збільшення відстані – швидко зменшується. Із квантової механіки відомо, що чим вужчим є потенціальна яма, тим більшою є відстань між енергетичними рівнями і тим глибше лежать рівні енергії електрона в потенціальній ямі. Оскільки в полі потенціалу (7.38) за не дуже малих радіусів ( $R = 1/a$ ) можуть існувати зв'язані стани електронів, то такий атом може захватувати більше одного електрона, тому в деяких випадках домішкові рівні можуть бути багатозарядними. Крім того, на дефектах типу **глибокого центру** можуть захоплюватись як електрони, так і дірки, тобто існувати домішки, які в одних умовах віддають електрон і відіграють роль донорів, а в інших – захоплюють додатковий електрон і поведуться як акцептори. Такі домішки називаються *амфотерними*. Глибокі центри характеризуються енергіями, що набагато перевищують величину  $kT$ . Для їх вивчення існують спеціальні методи спектроскопії глибоких рівнів, які зазвичай називають DLTS (Deep Level Transient Spectroscopy).

Домішкові рівні зручно зображувати на енергетичних діаграмах напівпровідників лініями у забороненій зоні. При цьому донорні рівні відстоять від дна зони провідності на величину енергії іонізації донора  $E_C - E_d$ ; акцепторні – розташовуються вище стелі валентної зони на відстані енергії іонізації акцептора  $E_a - E_V$  (рис. 7.5).

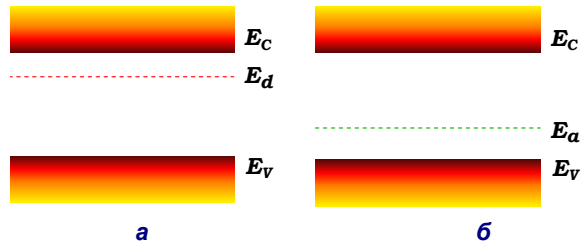


Рис. 7.5. Енергетичні діаграми напівпровідника із домішками: а – донорною; б – акцепторною

На рис. 7.6 подано енергії іонізації деяких домішок у германії та кремнії.

Таким чином, домішки, з одного боку, визначають тип провідності матеріалу та концентрацію носіїв, а з іншого – можуть впливати на процеси розсіювання та рекомбінації, тобто основні характеристики напівпровідників визначаються типом і кількістю домішок.

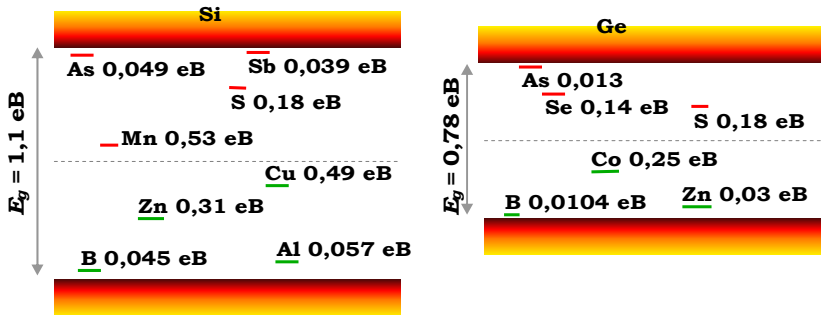


Рис. 7.6. Енергії іонізації деяких домішок у германії і кремнії.

Рівні, розташовані вище середини забороненої зони, відповідають донорам, нижче середини забороненої зони – акцепторам

## ■ 7.4. Задачі

1. У моделі воднеподібної домішки визначте енергію іонізації мілких домішок у германії та кремнії.
2. Оцініть максимальну рівноважну концентрацію мілких домішок ( $E_d = 0,01$  eВ) у напівпровіднику за кімнатної температури.
3. За якої температури варто проводити легування напівпровідника мілкою домішкою ( $E_d = 0,01$  eВ) для отримання їх рівноважної концентрації  $n = 5 \times 10^{15}$  eВ.

### **Список літератури**

1. *Келли А., Гровс Г.* Кристаллография и дефекты в кристаллах. – М.: Мир, 1974.
2. *Фридель Ж.* Дислокации. – М.: Мир, 1967.
3. *Бургуэн Ж., Ланно М.* Точечные дефекты в полупроводниках. Теория. – М.: Мир, 1984.
4. *Бургуэн Ж., Ланно М.* Точечные дефекты в полупроводниках. Экспериментальные аспекты. – М.: Мир, 1985.
5. *Ю Пітер, Кардона М.* Основы физики полупроводников. – М.: Физматлит, 2002.



# Розділ 8

## СТАТИСТИКА ЕЛЕКТРОНІВ У НАПІВПРОВІДНИКАХ В УМОВАХ ТЕРМОДИНАМІЧНОЇ РІВНОВАГИ

Повний набір параметрів, необхідний для опису квантових станів, складає чотири фізичних величини. В атомі це зазвичай енергія, модуль орбітального моменту, проекція орбітального моменту на вісь OZ, спіні. Для опису електронів у кристалі найзручніше використовувати набір проекцій квазіімпульсу (або хвильового вектора) та спіну.

Згідно із принципом Паулі у будь-якому квантовому стані із повним набором параметрів  $(L_1, L_2, L_3, s)$  може перебувати тільки один електрон (взагалі ферміон – частинка із напівцілим спіном). Оскільки набір проекцій квазіімпульсу визначає величину енергії електрона в кристалі, а проекція спіну на визначену вісь має два значення  $\pm 1/2$ , то на даному (невиродженому) енергетичному рівні може знаходитись не більше двох електронів.

Якщо кристал має форму паралелепіпеда із довжиною ребер  $N_1 a_1, N_2 a_2, N_3 a_3$  ( $N_i$  – кількість атомів уздовж ребра  $i$ , повною кількістю атомів у кристалі є  $N_1 N_2 N_3 = N$ ), то  $i$ -та проекція хвильового вектора  $\mathbf{k}$  набуває таких значень

$$k_i = \frac{2\pi}{a_i} \frac{n_i}{N_i}, \quad n_i = 0; \pm 1; \pm 2; \dots \pm \frac{N_i}{2},$$

тобто повне число станів у зоні Бриллюена дорівнює  $N$ . Це означає, що в невірродженій енергетичній зоні може бути не більше  $2N$  електронів. Якщо зона є  $g$ -разово виродженою, то в ній може бути не більше  $2Ng$  електронів.

### ■ 8.1. Статистика електронів і дірок

Шестивимірний простір координат та імпульсів називається фазовим. Елемент об'єму в цьому просторі

$$d\Gamma = d\mathbf{r} d\mathbf{p}. \quad (8.1)$$

Повний об'єм фазового простору задається інтегралом за всіма можливими значеннями координат та імпульсів системи

$$\Omega = \int d\Gamma. \quad (8.2)$$

Число електронів в елементі  $d\Gamma$  фазового простору має залежати від точки фазового об'єму, де воно визначається, бути пропорційним об'єму  $d\Gamma$  і, очевидно, змінюватись у часі

$$dN = F(\mathbf{r}, \mathbf{p}, t) d\Gamma. \quad (8.3)$$

Функція  $F(\mathbf{r}, \mathbf{p}, t)$  визначає число електронів в одиниці об'єму фазового простору. Таким чином, повне число електронів є інтегралом від (8.3) за всім фазовим простором

$$N = \int_{\Omega} d\Gamma F(\mathbf{r}, \mathbf{p}, t). \quad (8.4)$$

Якщо ввести функцію  $f(\mathbf{r}, \mathbf{p}, t) = \frac{1}{N} F(\mathbf{r}, \mathbf{p}, t)$ , (8.5)

то отримаємо ймовірність знаходження одного електрона в одиничному об'ємі фазового простору, тобто  $f(\mathbf{r}, \mathbf{p}, t)$  є функцією розподілу ймовірності, яка, зрозуміло, має нормуватись на одиницю

$$\int_{\Omega} d\Gamma f(\mathbf{r}, \mathbf{p}, t) = 1. \quad (8.6)$$

Зазначимо, що умови нормування не залежать від часу, у той час як функція розподілу у загальному випадку може залежати від часу. Нехай  $a$  – деяка величина, залежна від стану електрона, тобто один електрон у стані  $(\mathbf{r}, \mathbf{p})$  у момент часу  $t$  "продукує" величину  $a$ . В елементі фазового простору  $d\Gamma$  знаходиться  $dN = Nf(\mathbf{r}, \mathbf{p}, t)d\Gamma$  електронів, і сумарним значенням цієї фізичної величини є  $adN$ . Уся система електронів формує середню величину, що приходить на один електрон,

$$\langle a \rangle = \int_{\Omega} d\Gamma a(\mathbf{r}, \mathbf{p}, t) f(\mathbf{r}, \mathbf{p}, t). \quad (8.7)$$

Якщо кристал є однорідним, то ймовірність знаходження частинки у стані, що характеризується імпульсом  $\mathbf{p}$ , не залежить від координат, й інтегрування за координатами просто дає об'єм кристала. Тоді середня щільність частинок у цьому об'ємі визначається як

$$\bar{n} = \frac{N}{V} = \frac{1}{V} \frac{2}{8\pi^3 \hbar^3} \int_{\Omega} d\Gamma f(\mathbf{p}, \mathbf{r}, t) = \frac{1}{4\pi^3} \int_{V_{\mathbf{k}}} d\mathbf{k} f(\mathbf{k}, t) = n(t). \quad (8.8)$$

Іншими словами, в однорідних напівпровідниках середня щільність електронів дорівнює їхній концентрації. Оскільки хвильовий вектор та енергія електрона зв'язані визначеним співвідношенням – законом дисперсії, то можна вважати, що функція розподілу є функцією енергії. Для системи електронів (узагалі ферміонів) у рівноважному стані функція розподілу може бути отримана в рамках квантової статистики. Виходимо з того, що термодинамічний потенціал системи частинок, що знаходяться в  $k$ -му стані, записується у вигляді

$$\Omega_k = -k_B T \cdot \ln \sum_{n_k} \left( e^{\frac{\mu - \varepsilon_k}{k_B T}} \right). \quad (8.9)$$

Оскільки ми розглядаємо частинки, що підлягають принципу Паулі, то числа заповнення кожного стану можуть набувати тільки значень 0 або 1. Тоді із (8.9) отримуємо

$$\Omega_k = -k_B T \ln \left( 1 + e^{\frac{\mu - \varepsilon_k}{k_B T}} \right). \quad (8.10)$$

Згідно із загальними принципами статистичної фізики середнє число частинок у системі дорівнює похідній від термодинамічного потенціалу за хімічним потенціалом  $\mu$ , що береться зі знаком мінус

$$n_k = -\frac{\partial \Omega_k}{\partial \mu}. \quad (8.11)$$

Це середнє число частинок, які знаходяться у системі з енергією  $\varepsilon_k$  і є функцією розподілу частинок, що підлягають статистиці Фермі–Дірака.

Підстановкою до (8.11) термодинамічного потенціалу (8.10) отримуємо рівноважну функцію розподілу електронів за енергіями за заданої температури

$$f_0(E, T) = \frac{1}{e^{\frac{E - E_F}{k_B T}} + 1}. \quad (8.12)$$

Ця функція називається функцією розподілу Фермі–Дірака. Тут враховано такий факт: у фізиці напівпровідників історично склалось, що хімічний потенціал називають рівнем Фермі. Однак відомо, що у системі ферміонів енергія Фермі (найвище значення енергії, що заповнено за заданої концентрації частинок) збігається з хімічним потенціалом лише за нульової температури. Таким чином у (8.12) через  $E$  позначено повну енергію, що характеризує стан електронів, через  $E_F$  – рівень Фермі. Функція розподілу (8.12) визначає ймовірність того, що за температури  $T$  стан, який характеризується енергією  $E$ , зайнятий електроном. При  $T \rightarrow 0$  з (8.12) знаходимо

$$\lim_{T \rightarrow 0} f_0(E, T) = \begin{cases} 1, & E < E_F, \\ 0, & E > E_F. \end{cases} \quad (8.13)$$

При  $T = 0$  та  $E = E_F$  функція не визначена – має розрив. У цій точці функцію  $f_0(E, T)$  необхідно довизначити. Формула (8.13) показує, що при  $T = 0$  усі стани з енергією, меншою від енергії Фермі, зайняті, а стани з  $E > E_F$  – вільні. Таким чином, при  $T = 0$  енергія Фермі є максимальною, яку мають електрони у металах. Це енергетичний рівень, що відділяє вільні стани від зайнятих. При  $T \neq 0$  функція розподілу  $f_0(E_F, T) = 1/2$ . Таким чином, функцію розподілу Фермі при  $T = 0$  можна довизначити як  $f_0(E = E_F, T = 0) = 1/2$ . Вигляд функції розподілу електронів у металах подано на рис. 8.1. Завдяки тому, що у напівпровіднику в енергетичному спектрі є щілина (заборонена зона), розподіл електронів у власному напівпровіднику ма-

тиме інший вигляд (рис. 8.2). При  $T = 0$  у власному напівпровіднику всі електрони мали б заповнювати енергетичні рівні з енергіями  $E \leq E_V$ . При  $T \neq 0$  частина електронів може перейти у збуджений стан і заповнити стани  $E \geq E_C$  у зоні провідності. З іншого боку, у валентній зоні виникнуть незаповнені стани – дірки, які можна розглядати як деякі елементарні збудження, що мають усі властивості частинок.

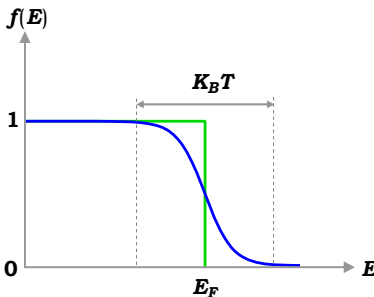


Рис. 8.1. Функція розподілу електронів у металах

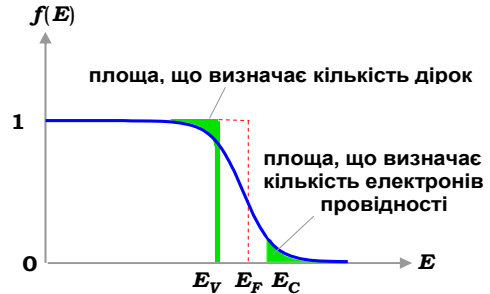


Рис. 8.2. Функція розподілу електронів у напівпровіднику

## ■ 8.2. Електрони та дірки

Із погляду статистики дірку можна визначити як вільний, не зайнятий електроном стан  $E(\mathbf{k})$ . Однак дуже важко зрозуміти, чому відсутність електрона на визначеному стані можна розглядати як частинку. Адже поняття дірки не тільки широко використовується у фізиці напівпровідників для моделювання тих чи інших властивостей, але існує ціла гама прямих експериментальних доказів наявності у напівпровідниках дірок – частинок, що характеризуються зарядом  $e^+ = -e > 0$  та ефективною масою, яка дорівнює ефективній масі електрона зі знаком мінус, хвильовим вектором  $\mathbf{k}_p = -\mathbf{k}$ , енергією  $E_p(\mathbf{k}_p) = -E(-\mathbf{k})$ . При цьому дрейфова швидкість дірок спрямована по полю, у той час як дрейфова швидкість електронів – проти поля. Цей факт означає, що основні принципи статистики можна формулювати для носіїв будь-якого одного типу, наприклад електронів, а для дірок усі вирази можна отримати шляхом вищенаведених перепозначень. Для того, щоб деяке збудження можна було вважати частинкою, необхідно переконатись, що воно має всі властивості частинки, й у першу чергу його динамічні характеристики однозначно можна описати як характеристики частинки. Для впевненості в тому, що незаповнені електронні стани в майже повністю заповненій (валентній) зоні можна розглядати як частинки – дірки, визначимо динамічні характеристики таких станів. Розглянемо двовимірну зону Брилюена, яка в невеликому околі поблизу центра зони зайнята електронними станами. Візьмемо для простоти систему за нульової температури, коли зайняті стани різко відділені від вільних. Нехай на



таку систему діє зовнішнє електричне поле  $\mathbf{E}$ , під дією якого всі стани зсуваються в оберненому просторі на величину

$$\Delta \mathbf{P} = e\mathbf{E}\tau, \quad (8.14)$$

де  $\tau$  – середній час вільного пробігу електронів. Оскільки зсув (8.14) не залежить від початкового стану  $\mathbf{P}_i(0)$ , то кожний зайнятий стан зсувається у зоні Бриллюена на однакову величину. Це означає, що за досить значного часу (що суттєво перевищує час вільного пробігу електронів,  $t \gg \tau$ ) всі зайняті стани як ціле зсунуться на величину (рис. 8.3)

$$\mathbf{P}_i(t) = \mathbf{P}_i(0) + e\mathbf{E}\tau. \quad (8.15)$$

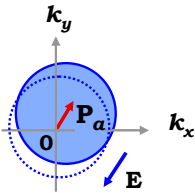


Рис. 8.3. Зсув електронних станів у зоні Бриллюена під дією електричного поля  $\mathbf{E}$

Оскільки число всіх заповнених станів дорівнює  $n$ , то повний імпульс усіх електронів задаватиметься сумою

$$\mathbf{P}_t = \sum_{i=1}^n \mathbf{P}_i(t). \quad (8.16)$$

Зважаючи на той факт, що завдяки симетрії зайнятих станів сума  $\sum_{i=1}^n \mathbf{P}_i(0) = 0$ , маємо

$$\mathbf{P}_t = ne\mathbf{E}\tau = n\mathbf{P}_\alpha. \quad (8.17)$$

Тут  $\mathbf{P}_\alpha$  – середній імпульс, що приходить на один електрон. Стану  $\mathbf{P}_i$  відповідає швидкість електрона  $\mathbf{v}_i = \mathbf{P}_i/m^*$  та елементарний струм  $\mathbf{J}_i = e\mathbf{v}_i$ . Повний струм, що створює система електронів під дією електричного поля  $\mathbf{E}$ , визначатиметься сумою елементарних струмів за всіма зайнятими станами

$$\mathbf{J}_t = \frac{e}{m^*} \sum_{i=1}^n \mathbf{P}_i = \frac{e}{m^*} ne\tau \mathbf{E}. \quad (8.18)$$

Якщо ввести рухливість електронів за допомогою формули

$$\mu_n = \frac{e\tau}{m^*}, \quad (8.19)$$

то повний струм у системі

$$\mathbf{J}_t = en\mu_n \mathbf{E}. \quad (8.20)$$

Розглянемо тепер ситуацію, коли зона майже повністю заповнена. Нехай у зоні  $N'$  станів заповнені електронами, а  $p$  станів – вільні, причому останні знаходяться поблизу екстремуму енергії у центрі зони Бриллюена в околі точки  $\mathbf{k} = 0$  (рис. 8.4 а). Під дією зовнішнього електричного поля всі стани зсунуться на однаковий вектор  $\Delta \mathbf{P} = e\mathbf{E}\tau$  так, що  $i$ -й стан характеризуватиметься імпульсом

$$\mathbf{P}_i(t) = \mathbf{P}_i(0) + e\mathbf{E}\tau. \quad (8.21)$$

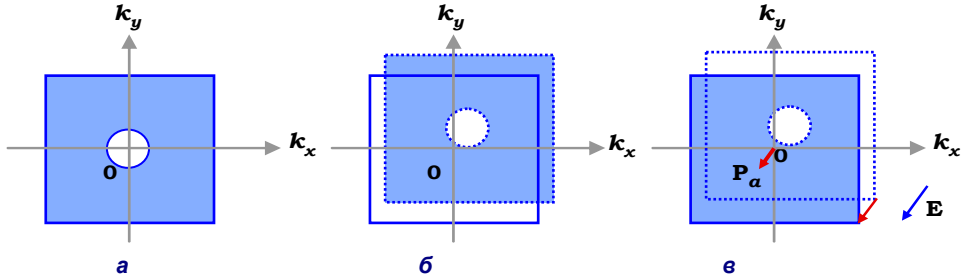


Рис. 8.4. Зсув електронних станів у майже повністю заповненій зоні Бриллюена під дією поля  $\mathbf{E}$

У результаті переміщення зайнятих станів частина з них перейде із першої зони Бриллюена до другої, а в першій з'явиться деяка частина додаткових незайнятих станів (рис. 8.4 б). Але всі стани із другої зони шляхом додавання до них вектора оберненої ґратки можна перевести до першої зони. Тому всі зайняті стани знаходяться в першій зоні Бриллюена, а незайняті – зсунуться щодо точки  $\mathbf{k} = 0$  на вектор  $\mathbf{P}_a$ . Знайдемо тепер повний імпульс, яким характеризується система електронів

$$\mathbf{P}_t = \sum_{i=1}^{N'} \mathbf{P}_i = \sum_{i=1}^{N'} \mathbf{P}_i + \sum_{i=1}^b \mathbf{P}_i - \sum_{i=1}^b \mathbf{P}_i = \sum_{i=1}^N \mathbf{P}_i - \sum_{i=1}^b \mathbf{P}_i .$$

Візьмемо до уваги, що сума всіх станів повністю заповненої зони Бриллюена дорівнює нулю, тобто перша сума у правій частині останньої рівності дає нуль. Тоді

$$\mathbf{P}_t = -\sum_{i=1}^b \mathbf{P}_i = -pe\mathbf{E}\tau . \tag{8.22}$$

Таким чином, завдяки періодичності кристалічного потенціалу сумарний імпульс сукупності  $N'$  електронних станів зсунеться не вздовж напрямку сили  $e\mathbf{E}$ , а проти неї. Середній квазіімпульс, що приходить на один електрон, дорівнює  $\mathbf{P}_i/N'$ , отже середня швидкість електронів

$$\mathbf{v}_a = -\frac{pe\tau}{m^*N'} \mathbf{E} . \tag{8.23}$$

Оскільки ми розглядаємо стани майже заповненої зони, то  $N \approx N'$  і  $p/N' \ll 1$ . Це означає, що під дією зовнішнього поля електрони набувають невеликої швидкості, тобто стани  $\mathbf{P}_a$  знаходяться поблизу точки  $\mathbf{P} = 0$  (рис. 8.4 в) або поблизу максимуму енергії, де ефективна маса є від'ємною. Таким чином можна визначити струм, що утворюється електронами майже заповненої зони Бриллюена

$$\mathbf{J}_{zB} = -eN'\mu_{zB} \mathbf{E} , \tag{8.24}$$

де формально введено рухливість електронів майже заповненої зони Бриллюена

$$\mu_{zB} = -\frac{e\tau}{m^* N'} \quad (8.25)$$

Якщо тепер ввести нові квазічастинки – дірки, згідно із правилом (дірки є квазічастинками, динамічні властивості яких тотожні динамічним властивостям сукупності електронів майже повністю заповненої зони), то статистику станів майже заповненої зони можна замінити статистикою елементарних збуджень – дірок. Дійсно, нехай

$$\mathbf{P}_t^{(p)} = \mathbf{P}_t, \quad (8.26)$$

де  $\mathbf{P}_t^{(p)}$  – сумарний квазіімпульс дірок. Природно вважати, що число дірок дорівнює числу всіх незайнятих станів. При цьому квазіімпульс дірки визначатиметься як квазіімпульс вільного стану, взятий зі знаком мінус

$$\mathbf{P}_p = -\mathbf{P}_0. \quad (8.27)$$

У цьому легко перекоонатися, якщо для простоти припустити, що в зоні існує єдиний незаповнений стан. Тоді сумарний імпульс електронів із майже заповненої зони

$$\mathbf{P}_t = \sum_{i=1}^{N'} \mathbf{P}_i$$

можна знайти, зважаючи на те, що сумарний імпульс повністю заповненої зони є нулем, тобто при додаванні до суми величини квазіімпульсу вільного стану отримаємо суму квазіімпульсів усіх заповнених станів і квазіімпульс вільного стану зі знаком мінус

$$\mathbf{P}_t = \sum_{i=1}^N \mathbf{P}_i - \mathbf{P}_0 = -\mathbf{P}_0.$$

Таке визначення квазіімпульсу дірки дозволяє зрозуміти, що передача квазіімпульсу діркою ґратці означає передачу квазіімпульсу ґратці сукупністю електронів, що майже повністю заповнюють стани у зоні Бриллюена. Оскільки  $d\mathbf{P}_0/dt = e\mathbf{E}$ , то

$$\frac{d\mathbf{P}_p}{dt} = -\frac{d\mathbf{P}_0}{dt} = -e\mathbf{E} = |e| \mathbf{E}. \quad (8.28)$$

Прирошення квазіімпульсу дірки відбувається так, начебто заряд частинки позитивний, тобто дірка є квазічастинкою із зарядом, що дорівнює заряду електрона, але позитивним. Запишемо тепер струм, що формується системою електронів, які майже повністю заповнюють стани в зоні Бриллюена

$$\mathbf{J}_{zB} = -|e| p \frac{e\tau}{m^*} \mathbf{E} = |e| p \frac{|e|\tau}{m_p} \mathbf{E}. \quad (8.29)$$

Якщо масу дірки визначити як масу електрона зі знаком мінус, то для струму, який виникає під дією поля  $\mathbf{E}$ , у системі електронів, що майже повністю заповнюють стани у зоні Бриллюена, можна записати вираз,

який означає струм, утворений позитивно зарядженими частинками з ефективною масою  $m_p^* = -m^*$ . Тобто ефективну масу дірки можна визначити як похідну другого порядку від закону дисперсії майже заповненої електронної зони.

$$\frac{1}{m_p^*} = -\frac{1}{m^*} = -\frac{1}{\hbar^2} \frac{d^2 E(\mathbf{k})}{dk^2}. \quad (8.30)$$

При цьому дірка характеризується позитивною ефективною масою, якщо знаходиться поблизу стелі валентної зони, і негативною – якщо поблизу дна зони. Із (8.29) випливає, що на протиположних електронам, які під дією електричного поля рухаються в напрямку проти поля, дірки рухаються в напрямку, що збігається з напрямком вектора поля, отже дрейфова швидкість дірки направлена по полю. Крім того, із (8.30) видно, що можна ввести закон дисперсії дірок через закон дисперсії електронів у валентній зоні  $E_p(\mathbf{p}) = -E(\mathbf{p})$ . Це правило показує, що закон дисперсії дірки визначається законом дисперсії електрона, але напрямок відрахування енергії буде протилежний щодо напрямку відрахування енергії електрона.

Таким чином, усі характеристики нової квазічастинки визначаються через характеристики електрона у валентній зоні.

Тепер легко визначити функцію розподілу для дірок, оскільки вони визначаються як квазічастинки, що займають у зоні Бриллюена вільні від електронів стани

$$f_p = 1 - f. \quad (8.31)$$

Підставляючи сюди формулу для функції розподілу електронів (8.12), отримуємо

$$f_p = \frac{1}{e^{(E_p - E_F)/kT} + 1}. \quad (8.32)$$

Зважаючи на (8.31), легко зрозуміти сенс указаної на рис. 8.2 площі, що характеризує кількість дірок у системі як кількість не заповнених електронами станів.

### ■ 8.3. Щільність станів

При розгляді властивостей електронної системи у напівпровіднику часто доводиться визначати експериментальні величини, що залежать від кількості дозволених станів електронів, які приходяться на одиничний інтервал енергії. В інтервалі енергії  $dE$  може міститися різна кількість станів  $dS$ , що має бути пропорційною  $dE$ , і коефіцієнт пропорційності має залежати від величини енергії, в околі якої визначається елементарний інтервал  $dE$ , тобто можна записати, що

$$dS = N(E)dE, \quad (8.33)$$

де  $N(E)$  – кількість станів в одиничному інтервалі енергії для одиничного об'єму кристала. Якщо ймовірність заповнення станів з енергією  $E$  дорівнює  $f_0(E, T)$ , то число електронів у цих станах

$$dn = 2 f_0(E, T) dS = 2 f_0(E, T) N(E) dE. \quad (8.34)$$

Фактор 2 означає, що рівень з однаковою енергією можуть займати два електрони із різними спінами. Тоді, інтегруючи (8.34) за всіма дозволеними значеннями енергії, отримаємо формулу, якою визначається концентрація електронів

$$n(T) = 2 \int_{(E)} f_0(E, T) N(E) dE. \quad (8.35)$$

Тут інтегрування треба проводити від дна зони провідності до її стелі. Але, враховуючи різку залежність функції розподілу від енергії, верхню межу інтегрування можна розповсюдити до нескінченності

$$n(T) = 2 \int_{E_c}^{\infty} f_0(E, T) N(E) dE. \quad (8.36)$$

Величина  $N(E)$  називається щільністю станів і визначається формою ізоенергетичних поверхонь. Дійсно, побудуємо в зоні Бриллюена дві ізоенергетичні поверхні  $E$  та  $E + dE$ , що виділяють деякий шар простору квазіімпульсів. Нехай об'єм цього шару  $d\mathbf{p}$ . Йому відповідає об'єм фазового простору

$$d\Gamma = V d\mathbf{p}. \quad (8.37)$$

Число фазових комірок у ньому  $d\Gamma/\hbar^3$ , а число станів для одиничного об'єму –  $dS = d\Gamma/\hbar^3 V$ . Таким чином

$$N(E) dE = \frac{1}{V} \frac{d\Gamma}{\hbar^3} = \frac{d\mathbf{p}}{\hbar^3} = \frac{d\mathbf{k}}{8\pi^3}. \quad (8.38)$$

Величину  $d\mathbf{k}$  можна визначити, якщо відомі рівняння ізоенергетичних поверхонь. Зважаючи на те, що щільність станів є дуже важливою характеристикою електронної системи напівпровідника, обчислимо її для різних систем, що характеризуються різними законами дисперсії.

### 8.3.1. Сферичні ізоенергетичні поверхні

Нехай мінімум енергії знаходиться у центрі зони Бриллюена

$$E_{\mathbf{k}} = E_c + \frac{\hbar^2 \mathbf{k}^2}{2m^*}. \quad (8.39)$$

Між ізоенергетичними поверхнями  $E$  та  $dE$  можна виділити сферичний шар товщиною  $dk$  та об'ємом  $d\mathbf{k}$

$$d\mathbf{k} = 4\pi \cdot k^2 dk = 2\pi k \cdot 2k dk. \quad (8.40)$$

Використовуючи (8.39), представляємо  $k$  через  $E$

$$k^2 = \frac{2m^*}{\hbar^2} (E - E_c) \Rightarrow k = \left( \frac{2m^*}{\hbar^2} \right)^{1/2} (E - E_c)^{1/2}. \quad (8.41)$$

Оскільки із першої рівності (8.41) маємо

$$2k dk = \frac{2m^*}{\hbar^2} dE, \quad (8.42)$$

то (8.40) дає

$$dk = 2\pi \left( \frac{2m^*}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2} dE. \quad (8.43)$$

Підставляючи (8.43) до (8.38), маємо

$$N(E)dE = \frac{2\pi}{8\pi^3} \left( \frac{2m^*}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2} dE. \quad (8.44)$$

Звідси отримуємо щільність станів для сферичної ізоенергетичної поверхні

$$N(E) = 2\pi \left( \frac{2m^*}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2}. \quad (8.45)$$

Таким чином, щільність станів для квадратичного закону дисперсії зі сферичними ізоенергетичними поверхнями визначається квадратним коренем із енергії, відрахованої від дна зони провідності, та є пропорційною квадратному кореню від куба ефективної маси зонного електрона: чим більша пологою є парабола, що описує закон дисперсії

квазічастинки поблизу дна зони провідності, тим більшими є ефективна маса частинок і щільність станів таких збуджень. Зрозуміло, що подібні вкладки можна виконати і для дірок. Результат збігається із (8.45), де замість ефективної маси електронів необхідно підставити ефективну масу дірки, а замість енергії дна зони провідності  $E_c$  – енергію стелі валентної зони –  $E_v$ .

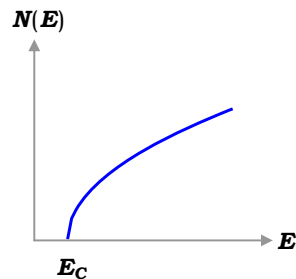


Рис. 8.5. Щільність станів для сферичної ізоенергетичної поверхні

### 8.3.2. Еліпсоїдальні ізоенергетичні поверхні

Енергія у випадку еліпсоїдальних енергетичних поверхонь залежить від компонентів хвильового вектора як

$$E(\mathbf{k}) = E_c + \frac{\hbar^2}{2} \left( \frac{k_x^2}{m_1^*} + \frac{k_y^2}{m_2^*} + \frac{k_z^2}{m_3^*} \right). \quad (8.46)$$

Закон дисперсії у канонічній формі має вигляд рівняння поверхні еліпсоїда в оберненому просторі

$$\frac{k_x^2}{a_1^2} + \frac{k_y^2}{a_2^2} + \frac{k_z^2}{a_3^2} = 1, \quad (8.47)$$

де

$$a_i = \left[ \frac{2m_i^*(E - E_c)}{\hbar^2} \right]^{1/2}.$$

(8.48)

Об'ємом еліпсоїда з напіввісями  $a_1$ ,  $a_2$ ,  $a_3$  є

$$v = \frac{4\pi}{3} a_1 a_2 a_3 = \frac{4\pi}{3\hbar^3} (8m_1^* m_2^* m_3^*)^{1/2} (E - E_c)^{3/2}. \quad (8.49)$$

Елементом об'єму між двома еліпсоїдами є еліпсоїдальний шар, об'єм якого визначається формулою  $dV = V(E + dE) - V(E) = (dV/dE)dE$ . Таким чином, у випадку еліпсоїдальних ізоенергетичних поверхонь

$$d\mathbf{k} = \frac{2\pi}{\hbar^3} (8m_1^* m_2^* m_3^*)^{1/2} (E - E_c)^{1/2} dE. \quad (8.50)$$

Підставляючи цей вираз до (8.38) і враховуючи, що в зоні Бриллюена може існувати кілька мінімумів енергії (напр., у германії їх чотири, у кремнії – шість), маємо

$$N(E) = M \frac{2\pi}{\hbar^3} (8m_1^* m_2^* m_3^*)^{1/2} (E - E_c)^{1/2}, \quad (8.51)$$

де  $M$  – кількість мінімумів у зоні. Введемо ефективну масу для щільності станів

$$(m_d^*)^3 = M^2 m_1^* m_2^* m_3^*, \quad (8.52)$$

тоді (8.51) можна переписати у вигляді, аналогічному (8.45)

$$N(E) = 2\pi \cdot \left( \frac{2m_d^*}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2}. \quad (8.53)$$

Аналогічну формулу можна отримати й для станів поблизу краю валентної зони. При цьому необхідно тільки зважати на те, що в цьому випадку елементи тензора ефективної маси електронів є від'ємними. У результаті матимемо ефективну масу для щільності станів дірок  $(m_{pd}^* = (M^2 m_{p1}^* m_{p2}^* m_{p3}^*)^{1/3})$ , де  $m_{pi}^* = -m_i^*$ ,  $E_V$  – стеля валентної зони)

$$N(E) = 2\pi \cdot \left( \frac{2m_{pd}^*}{h^2} \right)^{3/2} (E_V - E)^{1/2}. \quad (8.54)$$

Формули (8.50), (8.53) та (8.54) отримано за активного використання квадратичності закону дисперсії електронів поблизу екстремумів.

### 8.3.3. Щільність станів для неквадратичного закону дисперсії

У випадку, коли закон дисперсії носіїв не є квадратичним, можна запропонувати інші підходи до обчислення щільності станів. Використовуватимемо визначення

$$n = \int_{E_C}^{\infty} f_0(E, T) \frac{d\mathbf{k}}{8\pi^3}. \quad (8.55)$$

Для обчислення щільності станів визначимо елемент об'єму в імпульсному просторі як добуток елемента площі ізоенергетичної поверхні та нормального до нього компонента хвильового вектора

$$d\mathbf{k} = dS_E dk_n. \quad (8.56)$$

Для визначення  $dk_n$  маємо

$$dE = \frac{dE}{d\mathbf{k}} d\mathbf{k} = |\nabla_{\mathbf{k}} E| dk_n. \quad (8.57)$$

Тоді, згідно із (8.55), щільність електронів у системі

$$n = \frac{1}{4\pi^3} \int f(E, T) \frac{dS_E dE}{|\nabla_{\mathbf{k}} E|} = 2 \int f(E, T) dE \int_{S_E} \frac{dS_E}{8\pi^3 |\nabla_{\mathbf{k}} E|}. \quad (8.58)$$

Звідси для щільності станів отримаємо

$$N(E) = \frac{1}{8\pi^3} \int_{S_E} \frac{dS_E}{|\nabla_{\mathbf{k}} E|}. \quad (8.59)$$

Знайдемо на основі цієї формули щільність станів для системи, що підлягає закону дисперсії Кейна. Якщо для зони провідності за початок відліку енергії прийняти дно зони провідності, то після віднімання  $E_g$  маємо

$$E(\mathbf{k}) = -\frac{E_g}{2} + \frac{\hbar^2 \mathbf{k}^2}{2m} + \sqrt{\frac{E_g^2}{4} + \frac{2\mathbf{P}^2 \mathbf{k}^2}{3}}, \quad (8.60)$$

де параметр  $\mathbf{P}$  пов'язаний із матричним елементом імпульсу для взаємодіючих зон. Для багатьох напівпровідників оцінка модуля  $\mathbf{P}$  дає величину  $|\mathbf{P}| \sim 6 \cdot 10^{-8}$  еВ є. Це означає, що за малих значень імпульсу другий доданок у (8.60) можна відкинути. Тоді маємо



$$\left[ E(\mathbf{k}) - \frac{E_g}{2} \right]^2 = \frac{E_g^2}{4} + \frac{2\mathbf{P}^2 \mathbf{k}^2}{3}. \quad (8.61)$$

Звідси отримуємо

$$\frac{2\mathbf{P}^2 \mathbf{k}^2}{3} = E(\mathbf{k}) \left[ E(\mathbf{k}) + E_g \right]. \quad (8.62)$$

Із (8.60) знаходимо вираз для градієнта

$$|\nabla_{\mathbf{k}} E| = \frac{\hbar^2 k}{m} + \frac{(4\mathbf{P}^2/3)k}{\sqrt{\frac{2}{3}\mathbf{P}^2 \mathbf{k}^2 + \frac{E_g^2}{4}}}. \quad (8.63)$$

З іншого боку, для сферичних ізоенергетичних поверхонь маємо

$$N(E) = \frac{4\pi \cdot \mathbf{k}^2}{8\pi^3 |\nabla_{\mathbf{k}} E|} = \frac{\mathbf{k}^2}{2\pi^2 |\nabla_{\mathbf{k}} E|}. \quad (8.64)$$

Підставляючи сюди (8.63) і нехтуючи другим доданком у правій частині, маємо

$$N(E) = \frac{3k \sqrt{\frac{2}{3}\mathbf{P}^2 \mathbf{k}^2 + \frac{E_g^2}{4}}}{8\pi^2 \mathbf{P}^2}. \quad (8.65)$$

Виключимо звідси, згідно із (8.62), хвильовий вектор і знайдемо остаточно

$$N(E) = \frac{E^{1/2} (E_g + E)^{1/2} (E + E_g/2)}{4\pi^2 (2\mathbf{P}^2/3)^{3/2}} \quad (8.66)$$

або

$$N(E) = \frac{1}{2\pi^2} \left( \frac{2m_n^*}{\hbar^2} \right)^{3/2} E^{1/2} \left( 1 + \frac{E}{E_g} \right)^{1/2} \left( 1 + 2\frac{E}{E_g} \right), \quad (8.67)$$

де ефективна маса на дні зони провідності

$$m_n^* = \frac{3\hbar^2 E_g}{4\mathbf{P}^2}. \quad (8.68)$$

### 8.3.4. Щільність станів електронів напівпровідника в магнітному полі

Іноді для обчислення щільності станів можна використовувати підхід, що ґрунтується на прямому інтегруванні закону дисперсії. Ідея такого підходу полягає в наступному. Якщо необхідно обчислити інтеграл від деякої функції  $F(E(\mathbf{k}))$ , що непрямо залежить від хвильового вектора

$$I = \int \frac{d\mathbf{k}}{(2\pi)^3} F(E(\mathbf{k})), \quad (8.69)$$

то можна використати функцію щільності станів і записати інтеграл у вигляді

$$I = \int dE N(E) F(E). \quad (8.70)$$

Із порівняння цих формул маємо вираз для обчислення щільності станів прямим інтегруванням

$$N(E) = \int \frac{d\mathbf{k}}{(2\pi)^3} \delta(E - E(\mathbf{k})). \quad (8.71)$$

Тут через  $E(\mathbf{k})$  позначено функцію, що зв'язує енергію з хвильовим вектором – закон дисперсії.

Тепер за допомогою (8.71) обчислимо щільність станів електронів напівпровідника в магнітному полі. Як ми з'ясували раніше, у магнітному полі спектр енергії електронів характеризується появою рівнів Ландау

$$E = \frac{\hbar^2 k_z^2}{2m^*} + \hbar\omega_0 \left(n + \frac{1}{2}\right), \quad E_C = 0. \quad (8.72)$$

Підставимо закон дисперсії (8.72) до (8.71) і врахуємо, що інтегрування за  $k_x$  та  $k_y$  замінюється на підсумовування за  $n$ . Маємо

$$N(E) = \sum_n \int \frac{dk_z}{2\pi} \delta\left(E - \frac{\hbar^2 k_z^2}{2m^*} - \hbar\omega_0 \left(n + \frac{1}{2}\right)\right), \quad (8.73)$$

тобто для обчислення щільності станів за (8.73) потрібно обчислити інтеграл типу

$$J = \int dx \delta(\varepsilon - ax^2). \quad (8.74)$$

Зауважимо, що при  $\varepsilon < 0$  цей інтеграл тотожно дорівнює нулю. У випадку, коли  $\varepsilon > 0$  скористаємось відомими властивостями  $\delta$ -функції

$$\delta(A^2 - X^2) = \frac{1}{2A} [\delta(X - A) + \delta(X + A)] \quad (8.75)$$

та 
$$\delta(ax) = \frac{1}{a} \delta(x). \quad (8.76)$$

Тоді замість (8.74) маємо

$$J = \int dx \frac{1}{a} \delta(\varepsilon/a - x^2) = \frac{1}{a} \frac{1}{2\sqrt{\varepsilon/a}} \int dx [\delta(x - \sqrt{\varepsilon/a}) + \delta(x + \sqrt{\varepsilon/a})]. \quad (8.77)$$

Оскільки аргумент  $\delta$ -функції у другому доданку підінтегральної функції в (8.77) може дорівнювати нулю лише при  $x < 0$ , а в першому доданку – при  $x > 0$ , то можна зняти інтегрування, використавши властивості  $\delta$ -функції,

$$\int_{\Omega} dx F(x) \delta(x-a) = F(a), \quad a \in \Omega. \quad (8.78)$$

У результаті маємо 
$$J = \frac{1}{\sqrt{a}} \frac{1}{\sqrt{\varepsilon}}. \quad (8.79)$$

Пам'ятаючи що 
$$a = \frac{\hbar^2}{2m^*}, \quad \varepsilon = E - \hbar\omega_0(n+1/2), \quad (8.80)$$

із вимоги  $\varepsilon \geq 0$  знайдемо значення верхньої межі індексу підсумовування  $n(E)$

$$E - \hbar\omega_0(n+1/2) \geq 0 \Rightarrow n \leq \frac{E - \hbar\omega_0/2}{\hbar\omega_0} = \nu. \quad (8.81)$$

Підставляємо визначення (8.80) до (8.79). Із врахуванням нерівності (8.81) маємо із (8.73)

$$N(E) = \frac{1}{2\pi} \left( \frac{2m^*}{\hbar^2} \right)^{1/2} \sum_{n=0}^{\nu} (E - \hbar\omega_0(n+1/2))^{-1/2}. \quad (8.82)$$

Але ця формула не може вважатися остаточною, оскільки вона не враховує, що за зменшення магнітного поля до нуля має виконуватись граничний перехід

$$\lim_{B \rightarrow 0} N_B(E) = N_0(E) = 2\pi \left( \frac{2m^*}{\hbar^2} \right)^{3/2} \sqrt{E}. \quad (8.83)$$

Для виконання граничного переходу домножимо праву частину у (8.82) на деяку функцію магнітного поля  $g(B)$  і запишемо щільність станів у вигляді

$$N_B(E) = \sum_{n=0}^{\nu} N_B^{(n)}(E), \quad (8.84)$$

де позначено 
$$N_B^{(n)}(E) = \frac{g(B)}{2\pi} \left( \frac{2m^*}{\hbar^2} \right)^{1/2} (E - \hbar\omega_0(n+1/2))^{-1/2}. \quad (8.85)$$

Для визначення функції  $g(B)$  зауважимо: оскільки  $\omega_0 \sim B$ , при  $B \rightarrow 0$  число  $\nu$  для будь-якого обмеженого значення енергії прямує до нескінченності (див. (8.81)). Це означає, що за зменшення поля відстань між рівнями зменшується так, що від підсумовування в (8.84) можна перейти до інтегрування

$$\begin{aligned} \sum_{n=0}^{\nu} (E - \hbar\omega_0(n+1/2))^{-1/2} &\Rightarrow \int_0^{\nu} dn (E - \hbar\omega_0(n+1/2))^{-1/2} = \\ &= \frac{1}{\hbar\omega_0} \int_0^{E - \hbar\omega_0/2} (1/\sqrt{x}) dx = \frac{2}{\hbar\omega_0} \left( E - \frac{\hbar\omega_0}{2} \right)^{1/2}. \end{aligned} \quad (8.86)$$

Таким чином, у слабкому полі маємо

$$\lim_{B \rightarrow 0} g(B) \sum_{n=0}^{\nu} (E - \hbar\omega_0(n + 1/2))^{-1/2} = \lim_{\substack{B \rightarrow 0 \\ \omega_0 \rightarrow 0}} \frac{2g(B)}{\hbar\omega_0} (E)^{1/2}. \quad (8.87)$$

Для обмеження граничного значення (8.87) необхідно припустити, що  $g(B) = G\hbar\omega_0$ . Константу  $G$  знайдемо із (8.83)

$$N_B(E)|_{B \rightarrow 0} = \frac{G\hbar\omega_0}{2\pi} \left(\frac{2m^*}{\hbar^2}\right)^{1/2} \frac{2}{\hbar\omega_0} \left(E - \frac{\hbar\omega_0}{2}\right)^{1/2} = \frac{G}{\pi} \left(\frac{2m^*}{\hbar^2}\right)^{1/2} E^{1/2}. \quad (8.88)$$

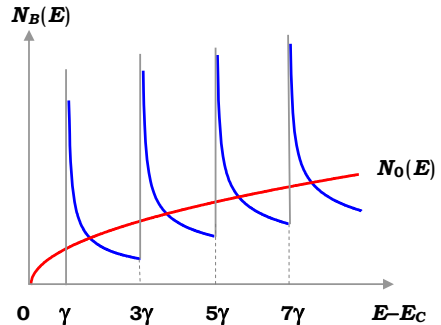
Порівнюючи цей результат з (8.83), знаходимо, що

$$\frac{G}{\pi} \left(\frac{2m^*}{\hbar^2}\right)^{1/2} = 2\pi \cdot \left(\frac{2m^*}{\hbar^2}\right)^{3/2} \Rightarrow G = \frac{2\pi \cdot m^*}{\hbar^2}. \quad (8.89)$$

Таким чином, щільність станів електронів напівпровідника в магнітному полі має вигляд

$$N_B(E) = \frac{\hbar\omega_0}{2} 4\pi \left(\frac{2m^*}{\hbar^2}\right)^{3/2} \sum_{n=0}^{\nu} (E - \hbar\omega_0(n + 1/2))^{-1/2}. \quad (8.90)$$

Щільність станів електронів напівпровідника в магнітному полі (8.90) подано на рис. 8.6. Хід  $N_B(E)$  в околі  $n$ -х особливих точок визначається тільки доданком  $N_B^{(n)}(E)$  у сумі (8.90). Удалині від цих точок ураховується вся сума  $\nu$  гіпербол. Унаслідок цього зі зростанням енергії  $E$  величина  $N_B(E)$  ліворуч від кожної особливої точки розташована тим вище, чим більшим є параметр  $\nu$ .



**Рис. 8.6.** Щільність станів електронів напівпровідника у магнітному полі.  $N_0(E) \sim \sqrt{E - E_c}$  є щільністю станів у нульовому магнітному полі;  $\gamma = \hbar\omega_0/2$

### 8.3.5. Щільність станів електронів, що локалізовані на атомах домішок

Вважатимемо, що кожен домішковий атом має один валентний електрон, що може бути відданий до зони провідності (донорна домішка) або один ненасичений зв'язок, на який електрон може бути захоплений із валентної зони напівпровідника – у такому випадку кажуть, що акцепторна домішка віддає до валентної зони дірку. Повним числом станів буде  $N_d$  – для атомів донорів і  $N_a$  – для атомів акцепторів. Якщо концентрація

домішкових атомів невелика й домішкова зона не формується, то для кожного рівня щільність станів має бути дуже значною – майже нескінченною. Але інтеграл від щільності станів за всіма можливими станами домішок, нормований на  $N_d$  або  $N_a$ , має дорівнювати одиниці (див. (8.6)). Таким вимогам задовольняє єдина функція –  $\delta$ -функція Дірака. Тоді для щільності станів можна записати

$$N_d(E) = N_d \delta(E - E_d), \quad (8.91)$$

$$N_a(E) = N_a \delta(E - E_a). \quad (8.92)$$

Зрозуміло, що завдяки процесам розсіювання електронів, які викликають (навіть за відсутності дисипації енергії) невизначеність імпульсу,  $\delta$ -функції у (8.91)–(8.92) розмиватимуться у вузькі високі піки. При цьому площі під піками мають дорівнювати одиниці.

## ■ 8.4. Концентрація носіїв заряду. Вироджені та невироджені напівпровідники

Як обговорювалось вище, концентрацію електронів у зоні провідності можна виразити через щільність станів і функцію розподілу. Таким чином,

$$n = \int_{E_C}^{\infty} f_0(E, T) N(E) dE. \quad (8.93)$$

Підставимо сюди щільність станів для еліпсоїдальних ізоенергетичних поверхонь

$$N(E) = 2\pi \cdot \left( \frac{2m_d^*}{h^2} \right)^{3/2} (E - E_C)^{1/2}. \quad (8.94)$$

Хоча (8.94) виведена для станів, енергія яких незначно перевищує енергію дна зони провідності, її можна використовувати при обчисленні концентрації електронів у напівпровіднику, оскільки функція розподілу Фермі швидко зменшується зі зростанням енергії, тобто внесок до загальної концентрації електронів від електронів з енергією, більшою за  $E_C$ , незначний. Підставимо до (8.93) вирази для щільності станів (8.94) і функцію розподілу Фермі–Дірака (8.12). У результаті маємо

$$n = 2 \int_{E_C}^{\infty} dE \frac{2\pi \cdot \left( \frac{2m_d^*}{h^2} \right)^{3/2} (E - E_C)^{1/2}}{e^{\frac{E - E_F}{kT}} + 1}. \quad (8.95)$$

Введемо безрозмірні змінні

$$x = \frac{E - E_C}{kT}, \quad \xi = \frac{E_F - E_C}{kT}, \quad dx = \frac{dE}{kT}. \quad (8.96)$$

Формула (8.95) у цих змінних набуває вигляду

$$n = 4\pi \left( \frac{2m_d^* kT}{h^2} \right)^{3/2} \int_0^\infty \frac{x^{1/2} dx}{e^{x-\xi} + 1}. \quad (8.97)$$

Якщо ввести позначення 
$$2 \left( \frac{2\pi \cdot m_d^* kT}{h^2} \right)^{3/2} = N_C \quad (8.98)$$

і 
$$\Phi_{1/2}(\xi) = \int_0^\infty \frac{x^{1/2} dx}{e^{x-\xi} + 1}, \quad (8.99)$$

то концентрація електронів визначатиметься такою формулою

$$n = \frac{2N_C}{\sqrt{\pi}} \Phi_{1/2}(\xi). \quad (8.100)$$

Тут  $N_C$  – ефективне число станів у зоні провідності, а  $\Phi_{1/2}(\xi)$  – інтеграл Фермі порядку 1/2.

Знайдемо тепер концентрацію вільних дірок у валентній зоні. Оскільки дірки у ній з'являються внаслідок того, що електрони при переході до зони провідності залишають незайняті електронні стани (розглядаються як такі, що заповнюються іншими квазічастинками – дірками), то щільність розподілу станів біля стелі валентної зони можна визначити як різницю між повною щільністю розподілу станів (одиноцею) і щільністю розподілу станів, що заповнені електронами,

$$f_{0p}(E, T) = 1 - f_0(E, T) = \frac{1}{e^{\frac{E_F - E}{kT}} + 1}. \quad (8.101)$$

Концентрацією дірок у валентній зоні є

$$p = 2 \int_{-\infty}^{E_V} f_{0p}(E, T) N(E) dE. \quad (8.102)$$

Тут нижню межу інтегрування розширено до  $-\infty$ , оскільки функція розподілу дуже швидко прямує до нуля за зменшення енергії. Введемо безрозмірні змінні

$$x = \frac{E_V - E}{kT}, \quad \eta = \frac{E_V - E_F}{kT}, \quad dx = -\frac{dE}{kT} \quad (8.103)$$

і запишемо концентрацію дірок у валентної зони

$$p = \frac{2N_V}{\sqrt{\pi}} \Phi_{1/2}(\eta), \quad (8.104)$$

де ефективне число станів  $N_V = 2 \left( \frac{2\pi \cdot m_{pd}^* kT}{h^2} \right)^{3/2}$ . (8.105)

Отже, концентрація електронів і дірок у напівпровіднику є функцією температури та залежить від рівня Фермі  $E_F$ .

Інтеграл Фермі залежно від величини параметра  $\xi$  можна наближено виразити таким чином<sup>2</sup>

$$\Phi_{1/2}(\xi) = \begin{cases} \frac{\sqrt{\pi}}{2} e^{\xi}, & -\infty < \xi < -1, \\ \frac{\sqrt{\pi}}{2} \frac{1}{1/4 + e^{-\xi}}, & -1 < \xi < 5, \\ \frac{2}{3} \xi^{3/2}, & 5 < \xi < \infty. \end{cases} \quad (8.106)$$

Тобто при  $\xi < -1$  або  $E_F < E_C - kT$  залежність концентрації електронів від енергії та температури відповідає класичній статистиці Больцмана. Тоді кажуть, що напівпровідник є невідродженим. Таким чином, якщо рівень Фермі лежить нижче зони провідності не менше, ніж на величину  $kT$ , електрони підлягають класичній статистиці, і напівпровідник називається невідродженим. Якщо ж рівень Фермі лежить вище дна зони провідності більше, ніж на  $5kT$ , то напівпровідник повністю вироджений. Згідно із (8.100) та (8.98), концентрацією електронів у невідродженому напівпровіднику є

$$n = \frac{2N_C}{\sqrt{\pi}} \Phi_{1/2}(\xi) \approx N_C e^{\xi} = 2 \left( \frac{2\pi \cdot m_d^* kT}{h^2} \right)^{3/2} e^{-\frac{E_C - E_F}{kT}}, \quad E_F < E_C - kT. \quad (8.107)$$

У випадку повністю виродженого напівпровідника концентрація електронів

$$n = \frac{8\pi}{3} \left( \frac{2m_d^*}{h^2} \right)^{3/2} (E_F - E_C)^{3/2} \quad (8.108)$$

не залежить від температури, що узгоджується з класичним визначенням виродженого електронного газу. У перехідній області від невідродженого до повністю виродженого напівпровідника концентрація електронів залежить від температури та задається формулою

<sup>2</sup> Число **5** подано лише для визначеності. Виродження у напівпровідниках може наставати за перевищення рівнем Фермі дна зони провідності на величину кількох  $kT$ .

$$n = N_C \frac{1}{1/4 + \exp\left(\frac{E_C - E_F}{kT}\right)}. \quad (8.109)$$

Ці результати легко узагальнити для випадку дірок. Для цього необхідно врахувати, що тепер інтеграл Фермі залежатиме від параметра  $\eta = (E_V - E_F)/kT$ . Таким чином можна дійти висновків: у невідродженому напівпровіднику концентрація дірок визначається статистикою Больцмана, а умовами її застосування є

$$\frac{E_V - E_F}{kT} < -1, \Rightarrow E_F > E_V + kT. \quad (8.110)$$

Отже:

**У невідродженому напівпровіднику рівень Фермі лежить нижче дна зони провідності та вище стелі валентної зони не менш ніж на величину  $kT$ .**

Замінюючи у (8.107) параметри електронів на параметри дірок, маємо концентрацію дірок у невідродженому напівпровіднику

$$p = N_V e^{-\frac{E_F - E_V}{kT}}, \quad N_V = 2 \left( \frac{2\pi \cdot m_{pd}^* kT}{h^2} \right)^{3/2}. \quad (8.111)$$

Умови повного виродження для дірок записуються у вигляді

$$\frac{E_V - E_F}{kT} > 5, \Rightarrow E_F < E_V - 5kT. \quad (8.112)$$

Це означає, що у повністю виродженому напівпровіднику рівень Фермі лежить у валентній зоні нижче її стелі на величину, не меншу за  $5kT$ . Концентрація дірок для виродженого напівпровідника задається формулою

$$p = \frac{8\pi}{3} \left( \frac{2m_{pd}^*}{h^2} \right)^{3/2} (E_V - E_F)^{3/2}. \quad (8.113)$$

Зауважимо, що добуток концентрацій електронів і дірок у невідродженому напівпровіднику не залежить від рівня Фермі

$$np = N_C N_V e^{-\frac{E_C - E_F}{kT}} e^{-\frac{E_F - E_V}{kT}} = N_C N_V e^{-\frac{E_C - E_V}{kT}}, \quad E_V + kT < E_F < E_C - kT. \quad (8.114)$$

Це дуже важливе співвідношення, яке показує, що:

**У невідродженому напівпровіднику в широкому інтервалі значень рівнів Фермі добуток концентрації електронів і дірок ви-**



**значається тільки температурою та шириною забороненої зони й не залежить від положення рівня Фермі.**

Часто (особливо в англomовній літературі) співвідношення (8.114) називають законом діючих мас. Це підкреслює, що добуток концентрацій електронів і дірок визначається шириною забороненої зони та добутком ефективних мас електронів і дірок.

## ■ 8.5. Залежність положення рівня Фермі від температури

Нам вже відомо, що статистика носіїв заряду (електронів і дірок) у напівпровіднику може змінюватись за зміни положення рівня Фермі. Далі детально розглянемо ефекти, що викликають зміни положення рівня Фермі.

### 8.5.1. Рівняння електронейтральності

Для визначення статистичних характеристик носіїв у напівпровіднику необхідно знати положення рівня Фермі  $E_F$ , для чого варто використати загальні закони, наприклад закон збереження заряду. Використовуватимемо так званий принцип електронейтральності – у кристалі в цілому та в будь-якому фізично малому об'ємі речовини сумарний заряд має бути нульовим.

Електрони виникають унаслідок іонізації донорної домішки та атомів основної речовини. Перехід електрона із валентної зони до зони провідності або до атомів акцепторної домішки викликає утворення вільних дірок у валентній зоні. Таким чином, від'ємний заряд, який дорівнює  $(n + N_A^-)e^-$ , утворюють вільні електрони у зоні провідності та іони акцепторів. Позитивний заряд, який дорівнює  $(p + N_D^+)e^+$ , утворюють вільні дірки та іони донорів. Тут  $N_A^-$  та  $N_D^+$  – кількість іонізованих акцепторів і донорів в одиниці об'єму, відповідно. Оскільки  $e^+ = -e^-$ , то умову електронейтральності можна записати як

$$(n + N_A^-) - (p + N_D^+) = 0. \quad (8.115)$$

Ясно, що наявність (або відсутність) іонізованих домішок відіграє вирішальну роль у процесах формування електронних властивостей напівпровідника, тому питання про розподіл електронів або дірок на локальних домішкових рівнях є важливим і потребує додаткового розгляду із застосуванням відомих принципів статистичної фізики.

### 8.5.2. Концентрація електронів і дірок на простих домішкових центрах

Розглянемо простий випадок, коли домішковий центр нейтральний або одноразово іонізований, тобто може знаходитись тільки у двох зарядових станах (випадок багатозарядових центрів – див. додаток С). Нехай такий центр (для визначеності вважатимемо його донором) характеризується рівнем енергії  $E_1$ . Тоді концентрацією нейтральних центрів буде  $N_1 = g_1 N_D f$ , де  $N_D$  – загальна концентрація донорів,  $f$  – функція розподілу Фермі–Дірака, а  $g_1$  – кратність виродження нейтрального центра. Концентрація іонізованих центрів, у свою чергу, виражається як  $N_D^+ = g_0 N_D (1 - f)$ , де  $g_0$  – кратність виродження іонізованого донора. Тоді маємо

$$\frac{N_1}{N_D^+} = \frac{g_1}{g_0} \exp\left(\frac{E_F - E_1}{kT}\right), \quad N_D^+ + N_1 = N_D. \quad (8.116)$$

Використовуючи (8.116), знайдемо функцію розподілу електронів за донорами

$$f_D = \frac{N_1}{N_D} = \left\{ 1 + \frac{g_0}{g_1} \exp\frac{E_1 - E_F}{kT} \right\}^{-1}. \quad (8.117)$$

Якщо вважати, що електронний рівень нейтрального донора вироджений тільки за спіном, а іонізований донор невироджений, тобто  $g_0 = 1$ , а  $g_1 = 2$ , то із (8.117) маємо, що розподіл електронів задається такою функцією:

♦ **за донорними рівнями** 
$$f_n^{(D)} = \left\{ \frac{1}{2} e^{\frac{E_D - E_F}{kT}} + 1 \right\}^{-1}; \quad (8.118)$$

♦ **за акцепторними рівнями** 
$$f_n^{(A)} = \left\{ 2e^{\frac{E_A - E_F}{kT}} + 1 \right\}^{-1}. \quad (8.119)$$

Коефіцієнт 2 перед експонентою вказує, що електронний рівень нейтрального акцептора невироджений, а рівень іонізованого акцептора вироджений за спіном. Для дірок маємо функції їх розподілу, відповідно:

♦ **за донорними рівнями** 
$$f_p^{(D)} = \frac{1}{2e^{\frac{E_F - E_D}{kT}} + 1}, \quad (8.120)$$

♦ **за акцепторними рівнями** 
$$f_p^{(A)} = \frac{1}{\frac{1}{2} e^{\frac{E_F - E_A}{kT}} + 1}. \quad (8.121)$$

Нехай  $N_D$  і  $N_A$  – кількість донорних та акцепторних домішок в одиниці об'єму напівпровідника, відповідно. Введемо позначення:  $n_D$ ,  $p_D$ ,  $n_A$ ,  $p_A$  –

кількість електронів і дірок, що знаходяться на донорних та акцепторних рівнях. Тоді концентрація електронів, що знаходяться на донорних рівнях, визначається різницею концентрацій донорів та іонізованих донорів. З іншого боку, іонізований донор можна розглядати як дірку, що локалізована на донорному рівні. Тоді

$$n_D = N_D - N_D^+ = N_D - p_D, \quad (8.122)$$

$$N_D^+ = N_D - n_D = p_D. \quad (8.123)$$

Аналогічні рівняння можна записати і для акцепторних рівнів

$$p_A = N_A - N_A^- = N_A - n_A, \quad (8.124)$$

$$N_A^- = N_A - p_A = n_A. \quad (8.125)$$

Підставимо  $N_A^-$  та  $N_D^+$  до (8.115) та отримаємо

$$(n + n_A) - (p + p_D) = 0, \quad (8.126)$$

або зважаючи на те, що  $n_A = N_A - p_A$  та  $p_D = N_D - n_D$ , знайдемо

$$n - p_A - p + n_D = N_D - N_A. \quad (8.127)$$

Враховуючи, що щільність домішкових станів є просто  $\delta$ -функцією,  $N_{d,a}(E) = N_{d,a}\delta(E - E_{d,a})$ , знайдемо:

♦ **число електронів, що знаходяться на донорних рівнях**

$$n_D = N_D \int dE \delta(E - E_D) \frac{1}{\frac{1}{2} e^{\frac{E - E_F}{kT}} + 1} = \frac{N_D}{\frac{1}{2} e^{\frac{E_D - E_F}{kT}} + 1}, \quad (8.128)$$

♦ **число дірок, що знаходяться на акцепторних рівнях**

$$p_A = N_A \int dE \delta(E - E_A) \frac{1}{\frac{1}{2} e^{\frac{E_F - E}{kT}} + 1} = \frac{N_A}{\frac{1}{2} e^{\frac{E_F - E_A}{kT}} + 1}. \quad (8.129)$$

Зауважимо, що отримані формули мають загальний вигляд  $n_D = N_D f_n^{(D)}(E_D)$  і  $p_A = N_A f_p^{(A)}(E_A)$ , що є прямим наслідком  $\delta$ -подібної щільності станів для домішкових центрів.

Повертаючись до (8.115), бачимо, що тепер рівняння електронейтральності набуває вигляду

$$4\pi \left( \frac{2m_{nd}^*}{h^2} \right)^{3/2} \int_{E_C}^{\infty} dE \frac{(E - E_C)^{1/2}}{e^{\frac{E - E_F}{kT}} + 1} + \frac{N_D}{\frac{1}{2} e^{\frac{E_D - E_F}{kT}} + 1} - 4\pi \left( \frac{2m_{pd}^*}{h^2} \right)^{3/2} \int_{\infty}^{E_V} dE \frac{(E_V - E)^{1/2}}{e^{\frac{E_F - E}{kT}} + 1} - \frac{N_A}{\frac{1}{2} e^{\frac{E_F - E_A}{kT}} + 1} = N_D - N_A. \quad (8.130)$$

Це рівняння записано для найпростішого випадку, коли є тільки по одному донорному та акцепторному рівню. Але воно однаково є дуже складним, і ми проаналізуємо лише деякі граничні випадки його розв'язків.

### 8.5.3. Рівняння електронейтральності для власного напівпровідника

Напівпровідник без домішок називається чистим або власним. Для нього рівнянням електронейтральності є

$$n = p, \quad (8.131)$$

що показує: перехід кожного електрона із валентної зони до зони провідності породжує дірку. У цьому випадку рівняння (8.130) зводиться до

$$\frac{2N_C}{\sqrt{\pi}} \Phi_{1/2}(\xi) = \frac{2N_V}{\sqrt{\pi}} \Phi_{1/2}(\eta). \quad (8.132)$$

Оскільки ширина забороненої зони в напівпровідниках є набагато більшою за  $2kT$  для температур порядку і вище кімнатної, то  $\xi \ll 1$  і  $\eta \ll 1$ . Тобто власний напівпровідник є невідродженим, й інтеграл Фермі можна представити у вигляді

$$\Phi_{1/2}(\xi) \cong \frac{\sqrt{\pi}}{2} e^{\xi} = \frac{\sqrt{\pi}}{2} e^{\frac{E_F - E_C}{kT}}, \quad (8.133)$$

$$\Phi_{1/2}(\eta) \cong \frac{\sqrt{\pi}}{2} e^{\eta} = \frac{\sqrt{\pi}}{2} e^{\frac{E_V - E_F}{kT}}. \quad (8.134)$$

Підставляючи ці функції до (8.132), маємо

$$N_C e^{\frac{E_F - E_C}{kT}} = N_V e^{\frac{E_V - E_F}{kT}}, \quad (8.135)$$

звідки за допомогою логарифмування отримаємо

$$\ln N_C + \frac{E_F - E_C}{kT} = \ln N_V + \frac{E_V - E_F}{kT} \quad (8.136)$$

або

$$\ln \frac{N_V}{N_C} = 2 \frac{E_F}{kT} - \frac{E_C + E_V}{kT}. \quad (8.137)$$

$$\text{Звідси остаточно маємо } E_F = \frac{E_C + E_V}{2} + kT \ln \left( \frac{N_V}{N_C} \right)^{1/2}. \quad (8.138)$$

Оскільки відношення

$$\frac{N_V}{N_C} = \left( \frac{m_{pd}^*}{m_{nd}^*} \right)^{3/2}, \quad (8.139)$$

то (8.138) дає

$$E_F = \frac{E_C + E_V}{2} + kT \ln \left( \frac{m_{pd}^*}{m_{nd}^*} \right)^{3/4}. \quad (8.140)$$

Бачимо, що при  $m_{nd}^* = m_{pd}^*$  рівень Фермі не залежить від температури та лежить рівно посередині забороненої зони. При  $m_{nd}^* \neq m_{pd}^*$  положення рівня Фермі залежить від температури. При  $T = 0$  рівень Фермі лежить посередині забороненої зони й зі зростанням температури лінійно залежить від  $T$ . Із подальшим зростанням температури рівень Фермі наближається до зони із меншою щільністю станів і тому заповнюється швидше (рис. 8.7). Концентрація власних носіїв заряду визначається формулою

$$n_i = n = p = \sqrt{N_C N_V} e^{-\frac{E_g}{2kT}}. \quad (8.141)$$

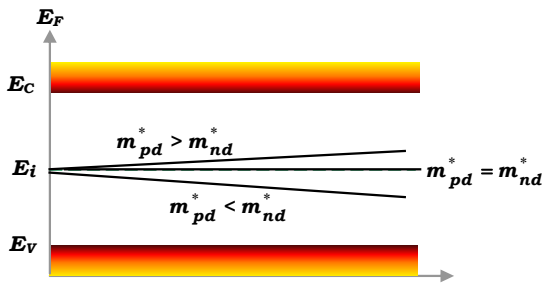


Рис. 8.7. Залежність рівня Фермі від температури у власному напівпровіднику.

$$E_F(0) = E_i$$

носіїв заряду від оберненої температури є прямою лінією з кутом нахилу, що пропорційний ширині забороненої зони

$$\operatorname{tg} \varphi = -\frac{E_g}{2k}, \quad (8.143)$$

звідки

$$E_g = 2k |\operatorname{tg} \varphi|. \quad (8.144)$$

Оцінка концентрації власних носіїв заряду для германію дає при  $T \approx 300$  К  $n_i \approx 2 \cdot 10^{13}$  см<sup>-3</sup> і, аналогічно, для кремнію  $n_i \approx 2 \cdot 10^{10}$  см<sup>-3</sup>. За допомогою побудови залежності логарифма концентрації носіїв заряду від оберненої температури в області власної провідності та визначення куту нахилу прямої можна отримати ширину забороненої зони напівпровідника. Однак визначення  $E_g$  за даними залежності концентрації носіїв від температури часто не дає правильних результатів, оскільки ширина забороненої зони в напівпровіднику залежить від температури. Зростання температури викликає збільшення коливальності атомів в ґратці, що спричиняє ефективне зменшення міжатомних відстаней у кристалі. Це викликає зміну кристалічного потенціалу, яка впливає на параметри зонної струк-

Логарифмуємо цей вираз, отримаємо (оскільки  $\sqrt{N_C N_V} \sim T^{3/2}$ )

$$\ln n_i = \operatorname{const} - \frac{3}{2} \ln \frac{1}{T} - \frac{E_g}{2k} \frac{1}{T}. \quad (8.142)$$

Логарифм змінюється набагато повільніше порівняно з лінійним (за  $1/T$ ) доданком, тому доданком  $\sim \ln(1/T)$  у правій частині (8.142) можна знехтувати. Таким чином, ми отримали дуже важливий результат: залежність логарифма концентрації власних

тури. Експериментально доведено, що у більшості напівпровідників в області температур 175-350 К зрос-

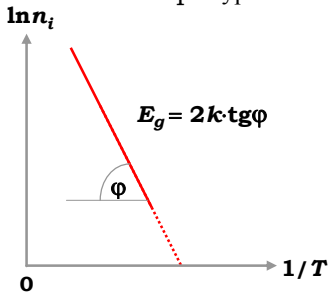


Рис. 8.8. Залежність  $\ln n_i$  від оберненої температури

тання температури призводить до зменшення ширини забороненої зони за лінійним законом

$$E_g(T) = E_g^{(0)} - \alpha T, \quad (8.145)$$

де  $E_g^{(0)}$  – ширина забороненої зони за нульової температури,  $\alpha$  – температурний коефіцієнт зміни ширини забороненої зони. У табл. 8.1 подано значення коефіцієнта  $\alpha$  для найпоширеніших напівпровідників.

Таблиця 8.1

Напівпровідник	Ge	Si	GaAs	InSb
$\alpha, \times 10^{-4} \text{ eВ} \cdot \text{K}^{-1}$	3,9	2,4	4,3	2,9

### 8.5.4. Напівпровідник, що має домішку одного типу

Нехай у напівпровідник введено домішку одного типу. Для визначеності вважатимемо її донором ( $N_D \neq 0$ , а  $N_A = 0$ ). У цьому випадку рівняння електронейтральності матиме такий вигляд

$$n + n_D - p = N_D \quad (8.146)$$

або

$$n = p + N_D^+. \quad (8.147)$$

Рівняння (8.147) вказує на існування двох каналів виникнення вільних електронів у зоні провідності завдяки:

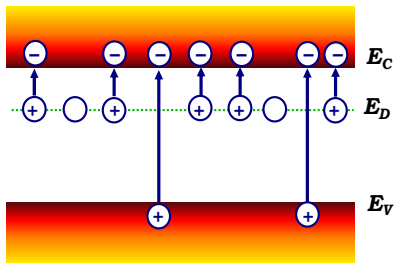


Рис. 8.9. Енергетична діаграма донорного напівпровідника

валентної зони до зони провідності (часто такий процес називають генерацією електрон-діркових пар при переході **зона-зона**); по-друге переходам електронів до зони провідності із домішкового рівня, у результаті чого з'являються  $N_D^+$  іонів донорної домішки (рис. 8.9). У невідродженому напівпровіднику, як ми знаємо,

$$np = N_C N_V e^{-E_g/kT}, \quad (8.148)$$

тобто добуток концентрації електронів і дірок не залежить від положення рівня Фермі, отже – від наявності домішок. Він дорівнює квадрату концентрації електронів у власному напівпровіднику. Із (8.147) доходимо та-

кого висновку: оскільки  $E_C - E_D < E_C - E_V$ , то за низьких температур основну роль відіграватимуть переходи електронів із домішкового рівня. Можна припустити, що

$$p \ll N_D^+ . \quad (8.149)$$

Зі зростанням температури відбуватиметься подальша іонізація домішок, доки всі вони не будуть іонізовані, і зростання концентрації електронів проходитиме тепер за рахунок іонізації основної речовини. Цей процес супроводжується зростанням концентрації дірок у валентній зоні. За великих температур очевидно буде досягнуто умови  $p \gg N_D^+ = N_D$ , і напівпровідник стає *власним*. А величина критичної температури очевидно визначатиметься концентрацією домішки. Згідно із цими міркуваннями розглядатимемо різні випадки температур. Почнемо з *низьких температур*. У цьому граничному випадку маємо  $p \ll N_D^+$ . Тоді рівняння електронейтральності набуває вигляду

$$n = N_D^+ \quad \text{або} \quad n = p_D . \quad (8.150)$$

Звідси випливає

$$N_C e^{-\frac{E_C - E_F}{kT}} = \frac{N_D}{2e^{\frac{E_F - E_D}{kT}} + 1} . \quad (8.151)$$

Розв'яжемо це рівняння щодо енергії Фермі, для чого введемо позначення  $e^{\frac{E_F}{kT}} = x$ . Тоді маємо

$$N_C e^{\frac{E_C}{kT}} x \left( 1 + 2e^{-\frac{E_D}{kT}} x \right) = N_D .$$

Перепишемо це квадратне рівняння у зручному вигляді

$$x^2 + \frac{1}{2} e^{\frac{E_D}{kT}} x - \frac{N_D}{2N_C} e^{\frac{E_C + E_D}{kT}} = 0 , \quad (8.152)$$

звідки

$$x = \frac{1}{4} e^{\frac{E_D}{kT}} \left[ -1 + \sqrt{1 + \frac{8N_D}{N_C} e^{\frac{\Delta E_D}{kT}}} \right] , \quad \Delta E_D = E_C - E_D . \quad (8.153)$$

Ми вибрали знак + перед коренем, оскільки тільки позитивні розв'язки мають фізичний сенс. Таким чином, із (8.153) маємо

$$e^{\frac{E_F}{kT}} = \frac{1}{4} e^{\frac{E_D}{kT}} \left[ \sqrt{1 + \frac{8N_D}{N_C} e^{\frac{\Delta E_D}{kT}}} - 1 \right] , \quad (8.154)$$

звідки

$$E_F = kT \ln \left\{ \frac{1}{4} e^{\frac{E_D}{kT}} \left[ \sqrt{1 + \frac{8N_D}{N_C} e^{\frac{\Delta E_D}{kT}}} - 1 \right] \right\} . \quad (8.155)$$

Зауважимо ще раз, що результат (8.155) отримано для випадку, коли електрони у зоні провідності виникають в основному за рахунок іонізації донорів. Можна сподіватись, що для досить малих температур існує температурний інтервал, в якому

$$\frac{8N_D}{N_C} e^{\frac{E_C - E_D}{kT}} \gg 1. \quad (8.156)$$

У цьому випадку із (8.154) отримаємо

$$e^{\frac{E_F}{kT}} = \sqrt{\frac{N_D}{2N_C}} e^{\frac{E_C + E_D}{2kT}}, \quad (8.157)$$

звідки

$$E_F = \frac{E_C + E_D}{2} + kT \ln \left( \frac{N_D}{2N_C} \right)^{1/2}. \quad (8.158)$$

При  $T = 0$  це рівняння дає

$$E_F = \frac{E_C + E_D}{2}, \quad (8.159)$$

тобто, на відміну від власного, у донорному напівпровіднику при  $T = 0$  рівень Фермі  $E_F^{(0)}$  знаходиться посередині між дном зони провідності та рівнем донорної домішки. Зі зростанням температури рівень Фермі починає зростати, досягає максимуму й далі зменшується так, що за температури  $T^*$  (для якої  $2N_C = N_D$ ) знов досягає енергії  $E_F^{(0)}$ . Концентрація електронів у цьому випадку (використано (8.151), де одиницею знехтовано)

$$n = N_C e^{\frac{E_C - E_F}{kT}} = \sqrt{\frac{N_C N_D}{2}} e^{\frac{\Delta E_D}{kT}}. \quad (8.160)$$

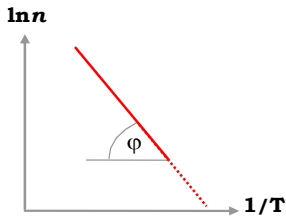
Іншими словами, за малих температур концентрація електронів визначається коренем квадратним із концентрації домішок. Із цієї формули отримуємо

$$\ln n = \text{const} + \ln T^{3/4} - \frac{\Delta E_D}{2k} \frac{1}{T}. \quad (8.161)$$

Оскільки за зміни температури логарифм змінюється дуже повільно, то в основному поведінку концентрації визначатиме доданок, пропорційний  $1/T$ . При цьому тангенс кута нахилу прямої (8.161) визначатиме енергію іонізації донора (рис. 8.10)

$$\text{tg} \varphi = \frac{\Delta E_D}{2k}. \quad (8.162)$$





**Рис. 8.10.** Визначення концентрації донорів із залежності  $\lg n = f(1/T)$

Розглянемо тепер випадок

$$\frac{8N_D}{N_C} e^{\frac{E_C - E_D}{kT}} \ll 1, \text{ тобто } N_C \gg 8N_D. \quad (8.163)$$

З одного боку, температура тут має бути достатньо високою для генерування достатньої кількості електронів, а з іншого – досить малою для виконання нерівності  $p \ll N_D^+$ . Тепер розкладемо у ряд підкореневу функцію у (8.153)

$$x \cong \frac{1}{4} e^{\frac{E_D}{kT}} \left( 1 + \frac{4N_D}{N_C} e^{\frac{\Delta E_D}{kT}} - 1 \right) = \frac{N_D}{N_C} e^{\frac{E_C}{kT}}, \quad (8.164)$$

звідки маємо

$$E_F = E_C + kT \ln \frac{N_D}{N_C}. \quad (8.165)$$

Оскільки ми розглядаємо область температур, де  $N_C \gg N_D$ , то логарифм є від'ємним. Зростання температури приводить до зменшення рівня Фермі. Концентрація електронів у цьому випадку

$$n = N_C e^{-\frac{E_C - E_F}{kT}} = N_C e^{\ln(N_D/N_C)} = N_D \quad (8.166)$$

не залежить від температури та дорівнює концентрації домішок. Це область збіднення (виснаження) домішки, де концентрація основних носіїв заряду лишається сталою, а концентрація неосновних носіїв заряду (дірок) швидко зростає за зростання температури

$$p = \frac{n_i^2}{n} = \frac{n_i^2}{N_D} = \frac{N_C N_V}{N_D} e^{-\frac{E_g}{kT}}. \quad (8.167)$$

Але треба пам'ятати, що цей вираз є справедливим, доки виконується нерівність  $p \ll N_D^+ = N_D$ .

Часто носії заряду називають *основними*, якщо їхня концентрація більша за концентрацію власних носіїв заряду  $n_i$  за заданої температури; якщо ж їхня концентрація менша за  $n_i$ , то їх називають *неосновними*. Але таке визначення не є коректним для компенсованого напівпровідника – такого, що має домішки обох типів: як донорні, так і акцепторні. У подібному випадку концентрація електронів і дірок може бути більшою за концентрацію власних носіїв за заданої температури. Тому для коректного визначення основних і неосновних носіїв заряду в напівпровідниках варто скористатись правилом діючих мас і записати  $pn = n_i^2$ . Звідси видно, що при  $n \gg n_i$  завжди  $p \ll n_i$  (коректніше можна сказати, що у випадку, коли  $n > p$  електрони є основними носіями заряду, а дірки – неосновними). У ви-

падку зворотної нерівності  $n < p$  основними носіями є дірки, а неосновними – електрони.

Далі розглянемо випадок високих температур. Зі зростанням температури концентрація дірок зростає й може зрівнятися із концентрацією електронів. Тоді маємо знову записати рівняння електронейтральності у вигляді

$$n = p + N_D^+ \quad (8.168)$$

Якщо всі атоми домішки іонізовані, тобто  $N_D^+ = N_D$ , можна брати до уваги іонізацію основної речовини. Перепишемо (8.168) для невідродженого напівпровідника

$$n = \frac{n_i^2}{n} + N_D \quad (8.169)$$

Звідси отримуємо квадратне рівняння

$$n^2 - N_D n - n_i^2 = 0 \quad (8.170)$$

Зважаючи на те, що величина  $n$  позитивно визначена, маємо розв'язок цього рівняння у вигляді

$$n = \frac{N_D}{2} \left( 1 + \sqrt{1 + \frac{4n_i^2}{N_D^2}} \right) \quad (8.171)$$

Використовуючи цей вираз, із (8.168) отримуємо для концентрації дірок

$$p = \frac{2n_i^2}{N_D \left( 1 + \sqrt{1 + \frac{4n_i^2}{N_D^2}} \right)} \quad (8.172)$$

Тепер, оскільки  $n = N_C \exp\{-(E_C - E_F)/kT\}$ , то

$$E_F = E_C + kT \ln \frac{n}{N_C} \quad (8.173)$$

Підставляємо сюди вираз (8.171) для концентрації електронів. Маємо

$$E_F = E_C + kT \ln \left[ \frac{N_D}{2N_C} \left( 1 + \sqrt{1 + \frac{4n_i^2}{N_D^2}} \right) \right]$$

або, записуючи явно концентрацію носіїв у власному напівпровіднику,

$$E_F = E_C + kT \ln \left[ \frac{N_D}{2N_C} \left( 1 + \sqrt{1 + \frac{4N_C N_V}{N_D^2} e^{-E_g/kT}} \right) \right] \quad (8.174)$$

Проаналізуємо отримані вирази:

• при  $\frac{4n_i^2}{N_D^2} \ll 1$   $n \cong N_D$ ,  $p \cong \frac{n_i^2}{N_D}$ , а для рівня Фермі  $E_F = E_C + kT \ln \frac{N_D}{N_C}$ ,

тобто маємо результат для області збіднення, коли концентрація носіїв визначається концентрацією донорів;

• при  $\frac{4n_i^2}{N_D^2} \gg 1$   $n = p = n_i$ , а для рівня Фермі  $E_F = \frac{E_C + E_V}{2} + kT \ln \frac{N_V}{N_C}$ ,

тобто маємо результат для власного напівпровідника.

Таким чином ми отримали вирази, які описують поведінку рівня Фермі у всьому температурному інтервалі від  $T = 0$  до температури збіднення  $T_d$

$$E_F = kT \ln \left\{ \frac{1}{4} e^{\frac{E_D}{kT}} \left[ \sqrt{1 + \frac{8N_D}{N_C} e^{\frac{\Delta E_D}{kT}}} - 1 \right] \right\}. \quad (8.175)$$

За температур, більших від  $T_d$

$$E_F = E_C + kT \ln \left[ \frac{N_D}{2N_C} \left( 1 + \sqrt{1 + \frac{4N_C N_V}{N_D^2} e^{-E_g/kT}} \right) \right]. \quad (8.176)$$

Перехід від збіднення домішки до власної провідності відбувається в деякому температурному інтервалі. Зрозуміло, що за заданої величини енергетичної щільності (для конкретного напівпровідника) цей інтервал визначається концентрацією домішки. За заданої концентрації домішки інтервал визначається величиною щільності. Позначимо температуру нижньої межі виснаження домішки через  $T_d^-$ , а температуру верхньої межі – через  $T_d^+$ . На рис. 8.11 проілюстровано поведінку рівня Фермі як функції температури в напівпровіднику із донорною домішкою. Літерами  $A$  та  $B$  позначено області виморожування та виснаження домішок,  $C$  – область власної провідності. Тепер можна визначити поведінку концентрації носіїв заряду від температури у широкому інтервалі температур. Використовуючи (8.142, 8.161 та 8.171), можна побудувати графік залежності логарифма концентрації електронів від оберненої температури (рис. 8.12). Оскільки нижня межа області виснаження визначається нерівністю  $N_C \gg 8N_D$ , а верхня –  $4n_i^2 \ll N_D^2$ , то можна вважати, що ці температури визначаються такими нерівностями:

$$\frac{N_C(T_d^-)}{8} \gg N_D > 2n_i(T_d^+). \quad (8.177)$$

Для чисельних оцінок використаємо, що

$$N_c(T) = 2,5 \cdot 10^{19} \left( \frac{m_d^*}{m} \right)^{3/2} \left( \frac{T}{300} \right)^{3/2}. \quad (8.178)$$

Нижню оцінку межі виснаження домішки знайдемо із рівняння  $N_c(T_d^-) = N_D$ .

$$\frac{T_d^-}{300} = \frac{m}{m_d^*} \cdot 2,5 \cdot \left( \frac{N_D}{10^{18}} \right)^{2/3}. \quad (8.179)$$

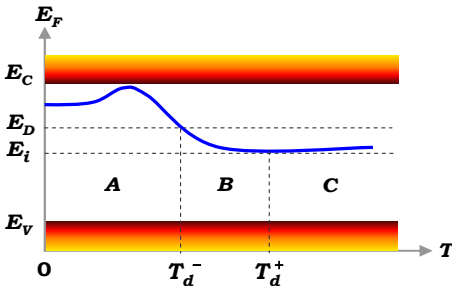


Рис. 8.11. Залежність рівня Фермі від температури в донорному напівпровіднику

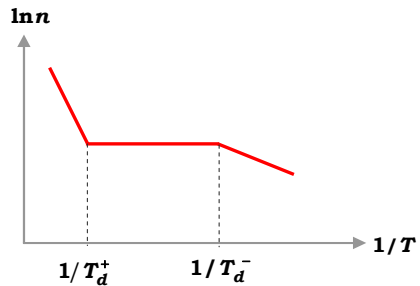


Рис. 8.12. Залежність концентрації носіїв заряду від температури в донорному напівпровіднику

Беручи до уваги, що для Ge  $m_d^* = 0,56m$  ( $E_d \approx 0,01$  eV) із (8.179) отримуємо, що для концентрації домішки  $N_D = 10^{18} \text{ см}^{-3}$  нижню межу області збіднення можна оцінити температурою  $T_d^- = 1340$  К. Для Si, де  $m_d^* = 1,08m$ , нижня межа збіднення домішки з концентрацією  $N_D = 10^{18} \text{ см}^{-3}$  становить 694 К. За зменшення концентрації донорів до рівня  $N_D = 10^{15} \text{ см}^{-3}$  температура нижньої межі області збіднення оцінюється для Ge як  $T_d^- = 13,4$  К і для Si – як  $T_d^- = 6,9$  К. Іноді температуру  $T_d^-$  визначають з умови  $E_F = E_D$ , звідки

$$T_d^- = \frac{\Delta E_D}{k \ln(N_c/N_d)}. \quad (8.180)$$

Оцінка за цією формулою дає такий самий порядок величин, як і оцінка за (8.179). Для визначення верхньої межі області виснаження домішки можна покласти  $p = N_D$  та  $n = 2N_D$ . Звідси, використовуючи формулу для добутку концентрацій електронів і дірок, маємо

$$np = 2N_D^2 = N_v(T_d^+)N_c(T_d^+)e^{-\frac{E_g}{kT_d^+}}. \quad (8.181)$$

Тоді

$$T_d^+ = \frac{E_g}{k \ln \left( \frac{N_c(T_d^+)N_v(T_d^+)}{2N_D^2} \right)}. \quad (8.182)$$

Використовуючи (8.178) (таку саму залежність від температури має й  $N_V$ ), із (8.182) отримуємо

$$T_d^+ \approx \frac{E_g}{k} \frac{1}{2 \ln \left( \frac{2,5 \cdot 10^{19}}{\sqrt{2} N_D} \right) + \frac{3}{2} \ln \left( \frac{m_d^* m_{dp}^*}{m^2} \right) + 3 \ln \left( \frac{T_d^+}{300} \right)}.$$

Звідси маємо рівняння щодо невідомої  $\zeta = 300/T_d^+$

$$\zeta - P = Q \ln \zeta, \quad (8.183)$$

де 
$$P = \frac{200k}{E_g} \left[ 3 \ln \left( \frac{2,5 \times 10^{19}}{\sqrt{2} N_D} \right) + \ln \left( \frac{m_d^* m_{dp}^*}{m^2} \right) \right], \quad Q = 900k/E_g.$$

Рівняння (8.183) можна розв'язати чисельно, наприклад графічним методом. Будуємо дві функції  $F_1 = \zeta - P$  та  $F_2 = Q \ln \zeta$ . Точка перетину функцій є розв'язком рівняння. Видно (рис. 8.13), що пряма може перетинати логарифмічну криву у двох точках. Для вибору одного розв'язку, який має фізичний сенс, зауважимо, що зі зростанням концентрації донорів (що означає зменшення параметра  $P$ ) температура  $T_d^+$  має зростати, тобто змінна  $\zeta$  зменшуватись. Такій вимозі відповідає розв'язок  $\zeta_2$ .

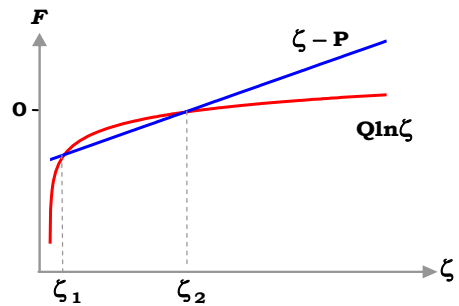


Рис. 8.13. Графічне визначення верхньої температури області виснаження

Проведений аналіз дозволяє дійти таких висновків: поведінка рівня Фермі в донорному напівпровіднику сильно залежить від температури і визначається матеріалом напівпровідника, типом і концентрацією домішок. Наприклад, за не дуже великих концентрацій домішки та нульової температури  $E_F$  лежить посередині між дном зони провідності й рівнем донорної домішки. Зі зростанням температури рівень Фермі спочатку зростає до деякого максимального значення, а потім починає зменшуватись так, наче він зрівнюється зі значенням рівня донорної домішки. Ця область температур називається областю слабкої іонізації, або областю вимороження донора. За подальшого зростання температури рівень Фермі зменшується до енергій, що ненабагато перевищують середину забороненої зони, після чого починає повільно збільшуватись. Ця область температур називається областю виснаження (збіднення) донорів й характеризується тим, що всі донори за таких температур є іонізованими. За подальшого збільшення температури основну роль у генерації носіїв відіграють переходи **зона-зона**, тобто напівпровідник поводить себе як власний. При цьому, згідно із (8.182),

температура, за якої настає власна провідність, буде тим меншою, чим меншими є ширина забороненої зони та концентрація домішки. Вона також зменшується для важких носіїв заряду.

На рис. 8.14 та 8.15 подано залежність положення рівня Фермі від температури в донорному та акцепторному напівпровідниках.

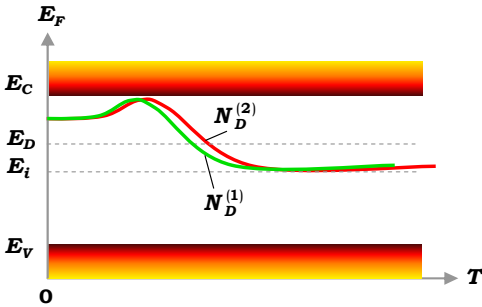


Рис. 8.14. Залежність рівня Фермі від температури в напівпровіднику  $n$ -типу за різних концентрацій домішок.  $N_D^{(2)} > N_D^{(1)}$

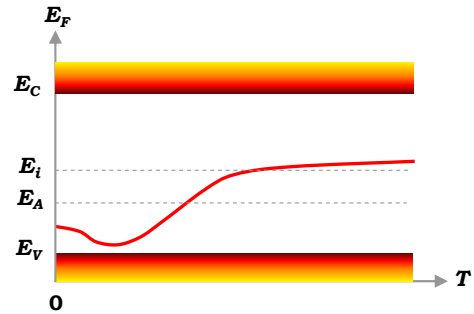


Рис. 8.15. Залежність рівня Фермі від температури в напівпровіднику  $p$ -типу

### 8.5.5. Напівпровідник, що містить акцепторну та донорну домішки

Розглянемо тепер випадок, коли напівпровідник одночасно легований домішкою як донорного, так і акцепторного типу. Припустимо, що  $T = 0$ . Тоді валентна зона буде повністю заповнена, а зона провідності – порожня. Оскільки в напівпровіднику міститься  $N_A$  вільних станів і  $N_D$  електронів, то електрони намагатимуться зайняти рівні з найменшими енергіями та перейдуть від донорів до акцепторів. Це пов'язано з тим, що енергетичні рівні акцепторів лежать нижче донорних рівнів. Якщо  $N_A = N_D$ , то в кристалі утворюється однакова кількість позитивно та негативно заряджених іонів  $N_D^+ = N_A^-$ . Нехай тепер температура зростає. Оскільки на донорному рівні електронів немає, то до зони провідності можливі переходи тільки із валентної зони та з рівня  $E_A$ . Але відстань  $E_C - E_A$  майже дорівнює ширині забороненої зони, тому концентрація  $n$  зростатиме майже як у власному напівпровіднику. При  $T = 0$  рівень Фермі лежатиме майже на середині забороненої зони, як й у власному напівпровіднику. Такий напівпровідник називається компенсованим, оскільки відбувається повна компенсація домішок, які не є постачальниками електронів і дірок (рис. 8.16).

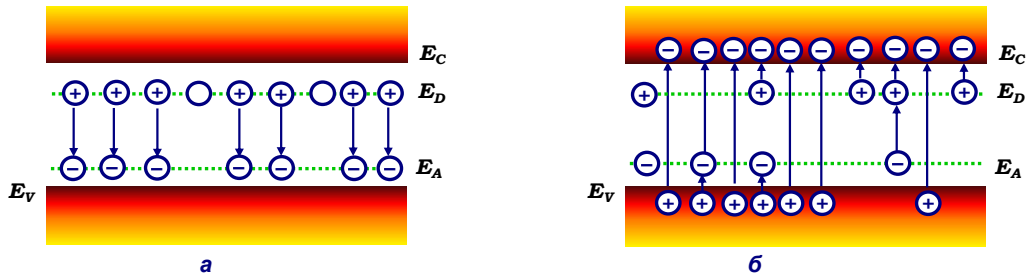


Рис. 8.16. Енергетична діаграма компенсованого напівпровідника ( $N_D > N_A$ ):  
а –  $T = 0$ ; б –  $T \neq 0$

Якщо концентрації домішок різного типу різні, то компенсація не буде повною. Нехай  $N_D > N_A$ , тоді величина  $N'_D = N_D - N_A$  відіграватиме роль домішки одного типу. Запишемо рівняння електронейтральності

$$n + n_D - p - p_A = N_D - N_A = N'_D. \quad (8.184)$$

При  $T \rightarrow 0$  концентрації  $n$  та  $p$  обнулюються, і рівняння електронейтральності спрощується

$$\frac{N_D}{\frac{1}{2} e^{\frac{E_D - E_F}{kT}} + 1} - \frac{N_A}{\frac{1}{2} e^{\frac{E_F - E_A}{kT}} + 1} = N'_D. \quad (8.185)$$

Припустимо, що рівень Фермі лежить вище енергії акцепторної домішки, тобто  $E_F > E_A$ . Тоді  $p_A = 0$  при  $T = 0$ , і рівняння електронейтральності спрощується

$$n_D = N'_D. \quad (8.186)$$

Звідси отримуємо

$$\frac{N_D}{\frac{1}{2} e^{\frac{E_D - E_F}{kT}} + 1} = N'_D \quad (8.187)$$

або

$$e^{\frac{E_D - E_F}{kT}} = 2 \frac{N_D - N'_D}{N'_D} = 2 \frac{N_A}{N_D - N_A}. \quad (8.188)$$

Із цього рівняння знаходимо рівень Фермі

$$E_F = E_D + kT \ln \frac{N_D - N_A}{2N_A}. \quad (8.189)$$

Видно, що за нульової температури  $E_F = E_D$ , тобто у напівпровіднику із домішками двох типів за нульової температури рівень Фермі збігається з рівнем енергії тієї домішки, концентрація якої є більшою. Для визначення концентрації електронів запишемо загальний вираз для концентрації електронів у невідродженому напівпровіднику

$$n = N_C e^{-\frac{E_C - E_F}{kT}} = N_C e^{-\frac{E_C}{kT}} e^{\frac{E_F}{kT}}. \quad (8.190)$$

Із (8.188) маємо

$$e^{\frac{E_F}{kT}} = \frac{N_D - N_A}{2N_A} e^{\frac{E_D}{kT}}. \quad (8.191)$$

Підставимо цей вираз до (8.190)

$$n = N_C \frac{N_D - N_A}{2N_A} e^{-\frac{E_C - E_D}{kT}}, \quad (8.192)$$

тобто енергія активації ( $\delta E_a = E_C - E_D$ ) дорівнює енергії іонізації донорної домішки, у той час як у чисто донорному напівпровіднику енергія активації удвічі менша. Ці формули можна використовувати не тільки за дуже низьких температур  $T \rightarrow 0$ , але й при  $T \neq 0$ . Із (8.189) видно, що зі зростанням температури рівень Фермі зсувається від  $E_F = E_D$  вгору або вниз залежно від співвідношення між  $N_A$  і  $N_D$ . Якщо  $N_D = 3N_A$ , то  $E_F = E$ , і рівень Фермі не залежить від температури (доки виконуються вихідні умови). При  $N_D > 3N_A$  рівень Фермі піднімається вгору тим швидше, чим меншим є  $N_A$ .

Детальніший аналіз показує, що наявність компенсуючої домішки викликає швидкіші зміни положення рівня Фермі. При цьому, якщо концентрації домішок донорів і акцепторів не значно різняться, при  $T = 0$  рівень Фермі збігається з енергією тієї домішки, концентрація якої є більшою. Підвищення температури приводить до переходу до власного напівпровідника, причому температура такого переходу тим менша, чим меншою є різниця між концентраціями домішок. Із (8.192)

випливає: тангенс кута нахилу графіка  $\ln n$  як функції  $1/T$  визначається величиною  $\Delta E_D/k$  для електронів (та  $\Delta E_A/k$  – для дірок), що вдвічі перевищує тангенс кута нахилу у випадку, коли напівпровідник вміщує домішку тільки одного типу. На рис. 8.17 подано температурну залежність рівня Фермі у напівпровіднику  $n$ -типу, що має донорну та акцепторну домішки за різних співвідношень між концентраціями донорів та акцепторів.

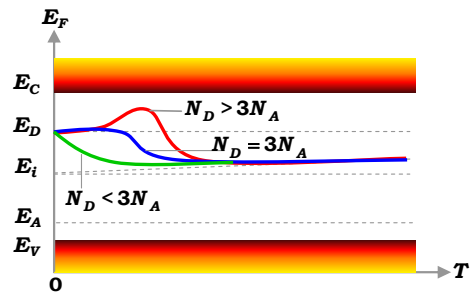


Рис. 8.17. Залежність рівня Фермі від температури в компенсованому напівпровіднику  $n$ -типу ( $N_D > N_A$ )

### 8.5.6. Вироджений напівпровідник

Як відомо, виродженим є напівпровідник, де рівень Фермі лежить усередині дозволеної енергетичної смуги на відстані, що перевищує кілька  $kT$  від її границі. Концентрація електронів при цьому не залежить від температури. Таку ситуацію можна відтворити належним легуванням напівпровідника за визначених значень температури. Дійсно, як ми бачили вище, зі зростанням температури рівень Фермі у домішковому напівпровіднику наближається до смуги зони провідності. Знайдемо поло-



ження максимуму рівня Фермі зі зміною температури. При цьому для визначеності розглядатимемо донорний напівпровідник. Виходимо із рівняння для рівня Фермі у напівпровіднику, що має тільки донорну домішку

$$E_F = \frac{E_C + E_D}{2} + \frac{kT}{2} \ln \frac{N_D}{2N_C}. \quad (8.193)$$

Умова максимуму положення рівня Фермі як функції температури дає

$$\frac{dE_F}{dT} = \frac{k}{2} \ln \frac{N_D}{2N_C} - \frac{kT}{2} \frac{N_D}{2N_C} \frac{2}{N_D} \frac{dN_C}{dT} = 0. \quad (8.194)$$

Звідси маємо рівняння  $\ln \frac{N_D}{2N_C} - \frac{T}{N_C} \frac{dN_C}{dT} = 0$ . (8.195)

Оскільки  $\frac{dN_C}{dT} = \frac{3}{2} \frac{N_C}{T}$ , (8.196)

то із (8.195) маємо  $\ln(N_D/2N_C) = 3/2$  або

$$2N_C(T_{\max}) = \frac{N_D}{e^{3/2}}. \quad (8.197)$$

Підставляючи цей вираз до (8.193), отримуємо

$$E_F = \frac{E_C + E_D}{2} + \frac{kT_{\max}}{2} \ln \frac{N_D}{2N_C(T_{\max})} = \frac{E_C + E_D}{2} + \frac{3}{4} kT_{\max}. \quad (8.198)$$

Концентрація домішки, за якої рівень Фермі збігається із дном зони провідності  $E_F^{\max} = E_C$ , називається критичною. Знайти її можна із такої умови

$$E_C = \frac{E_C + E_D}{2} + \frac{3}{4} kT_{\max}. \quad (8.199)$$

Пригадаємо, що  $N_C$  пов'язано з температурою через співвідношення  $N_C = \text{const} \cdot T^{3/2}$ . Тоді з (8.197) знаходимо

$$T_{\max} = \left( \frac{N_D}{2\text{const} \cdot e^{3/2}} \right)^{2/3}. \quad (8.200)$$

Далі із (8.199) отримуємо, що критична концентрація домішки, за якої починається виродження напівпровідника,

$$N_D^{cr} = 2\text{const} \cdot e^{3/2} \left( \frac{E_C - E_D}{3k} \right)^{3/2}. \quad (8.201)$$

Ця критична концентрація початку виродження напівпровідника суттєво залежить від енергії іонізації домішки та через  $\text{const} \sim (m_d^*/m)^{3/2}$  – від ефе-

ктивної маси. Таким чином, збільшуючи концентрацію домішки, можна ефективно зсунути рівень Фермі високо до зони провідності, і таким чином перетворити напівпровідник на вироджений. Оцінка критичної концентрації (8.201) для Ge, де  $(m_d^*/m) = 0,56$ , легованого миш'яком ( $E_C - E_D = 0,013$  eV), дає  $N_D^{cr} = 1,2 \cdot 10^{19} \text{ см}^{-3}$ .

## ■ 8.6. Задачі

1. Як зміниться концентрація електронів власного напівпровідника, якщо температура зростає з  $T_1$  до  $3/2T_1$ , а залежність забороненої зони від температури є лінійною  $E_g = E_g^{(0)} - \alpha T$ .

2. Нехай у власному напівпровіднику  $n(T_1 = 400 \text{ K}) = 1,3 \cdot 10^{16} \text{ см}^{-3}$  та  $n(T_2 = 350 \text{ K}) = 6,2 \cdot 10^{15} \text{ см}^{-3}$ . Припустимо, що ширина забороненої зони змінюється зі зміною температури за лінійним законом  $E_g = E_g^{(0)} - \alpha \cdot T$ . Знайдіть ширину забороненої зони напівпровідника при  $T = 0 \text{ K}$ . (Використайте результат попередньої задачі.)

3. У випадку вироджених валентних зон закон дисперсії дірок поблизу стелі валентної зони має вигляд

$$E_p(\mathbf{k}) = E_v - \left( \frac{\hbar^2}{2m_e} \right) \left[ Ak^2 \pm \sqrt{B^2 k^4 + C^2 (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2)} \right]. \quad (8.202)$$

Обчисліть щільність станів та ефективні маси легких і важких дірок.

**Розв'язок.** Згадаємо, що основна ідея введення щільності станів полягає в необхідності обчислення інтегралів типу

$$I = 2 \int \frac{d\mathbf{k}}{(2\pi)^3} F(\mathbf{k}). \quad (8.203)$$

Припустимо, що функція  $F(\mathbf{k})$  є неявною від змінних  $k$ ,  $\vartheta$ ,  $\phi$ , де  $\vartheta$ ,  $\phi$  – полярний та азимутальний кути, які визначають напрямок хвильового вектора. З іншого боку, ця функція є явною функцією змінних  $E$ ,  $\vartheta$ ,  $\phi$ , де  $E$  – енергія частинки, що пов'язана з модулем хвильового вектора співвідношенням  $k = k(E)$ . Тоді в інтегралі (8.203) можна перейти до нових змінних

$$I = 2 \int \frac{d\mathbf{k}}{(2\pi)^3} F(E(\mathbf{k})) \rightarrow I = \int_{-\infty}^{E_v} dE \left[ \frac{2}{(2\pi)^3} \int_0^\pi \sin \vartheta d\vartheta \int_0^{2\pi} d\phi k^2(E, \vartheta, \phi) |J| \right] F(E), \quad (8.204)$$

де  $J$  – якобіан переходу від змінних  $k$ ,  $\vartheta$ ,  $\phi$  до змінних  $E$ ,  $\vartheta$ ,  $\phi$ . Запишемо закон дисперсії (8.202) у вигляді

$$E_p(\mathbf{k}) = E_v - \left( \frac{\hbar^2 k^2 C}{2m_e} \right) \left[ A/C \pm \sqrt{(B/C)^2 + \sin^2 \vartheta \left( (1/4) \sin^2 2\phi \sin^2 \vartheta + \cos^2 \vartheta \right)} \right].$$

Звідси неважко отримати енергію, що записана у стандартному вигляді

$$E_p(k) = E_V - \frac{\hbar^2 k^2}{2m_{\pm}(\vartheta, \phi)}, \quad (8.205)$$

де через 
$$m_{\pm}(\vartheta, \phi) = \frac{m_e}{C} \Phi_{\pm}^{-1}(\vartheta, \phi) \quad (8.206)$$

позначено ефективні маси легких і важких дірок. Тут

$$\Phi_{\pm}(\vartheta, \phi) = A/C \pm \sqrt{(B/C)^2 + \sin^2 \vartheta \left( (1/4) \sin^2 2\phi \cdot \sin^2 \vartheta + \cos^2 \vartheta \right)}. \quad (8.207)$$

Для обчислення якобіана переходу скористаємось (8.205) і запишемо

$$k = \sqrt{\frac{2m_{\pm}(\vartheta, \phi)}{\hbar^2} (E_V - E)}. \quad (8.208)$$

Як відомо, якобіан переходу записується у вигляді визначника

$$J = \begin{vmatrix} \partial k / \partial E & \partial \vartheta / \partial E & \partial \phi / \partial E \\ \partial k / \partial \vartheta & \partial \vartheta / \partial \vartheta & \partial \phi / \partial \vartheta \\ \partial k / \partial \phi & \partial \vartheta / \partial \phi & \partial \phi / \partial \phi \end{vmatrix} = \begin{vmatrix} \partial k / \partial E & 0 & 0 \\ \partial k / \partial \vartheta & 1 & 0 \\ \partial k / \partial \phi & 0 & 1 \end{vmatrix} = \partial k / \partial E. \quad (8.209)$$

Використовуючи (8.208), отримаємо

$$|J| = \frac{1}{2} \sqrt{\frac{2m_{\pm}(\vartheta, \phi)}{\hbar^2}} \cdot \frac{1}{\sqrt{E_V - E}}. \quad (8.210)$$

Запишемо тепер добуток

$$k^2 |J| = \frac{1}{2} \left( \frac{2m_{\pm}(\vartheta, \phi)}{\hbar^2} \right)^{3/2} \sqrt{E_V - E},$$

який підставимо до виразу у квадратних дужках (8.204). Маємо щільність станів легких і важких дірок

$$N_{\pm}^{(p)}(E) = \frac{1}{(2\pi\hbar)^3} \int_0^{\pi} \sin \vartheta d\vartheta \int_0^{2\pi} d\phi (2m_{\pm}(\vartheta, \phi))^{3/2} \sqrt{E_V - E}. \quad (8.211)$$

Введемо ефективні маси щільності станів дірок, згідно із формулою

$$m_{\pm}^p = \frac{m_e}{C} \left[ \int_0^{\pi} \sin \vartheta d\vartheta \int_0^{2\pi} d\phi \Phi_{\pm}^{-3/2}(\vartheta, \phi) \right]^{2/3}. \quad (8.212)$$

Тоді щільність станів (8.211) запишемо у стандартному вигляді

$$N_{\pm}^{(p)}(E) = \frac{(2m_{\pm}^{(p)})^{3/2}}{(2\pi\hbar)^3} \sqrt{E_V - E}. \quad (8.213)$$

Таким чином у напівпровіднику із кейновським законом дисперсії щільність станів легких і важких дірок можна записати у стандартному вигляді з ефективними масами щільності станів (8.213).

4. Закон дисперсії дірок у валентній зоні Ge можна представити у вигляді

$$E(\mathbf{k}) = E_v - \frac{\hbar^2 k^2}{2m} \left\{ A \pm B \sqrt{1 + 6\delta \left( \frac{k_x^2 k_y^2 + k_x^2 k_z^2 + k_z^2 k_y^2}{k^4} - \frac{1}{6} \right)} \right\},$$

де  $B = 10,03$ ,  $A = 13,3$ ,  $\delta = 0,27$ . Знайдіть ефективні маси щільності станів для зон важких і легких дірок та ефективну щільність станів. Обчисліть частку легких дірок.

5. Враховуючи можливість виродження носіїв у донорному напівпровіднику, оцініть концентрацію донорів  $N_D$ , за якої рівень Фермі попадає до зони провідності. Вважайте, що

при  $\zeta = \frac{E_F - E_C}{kT} \leq 1,3$  інтеграл Фермі можна апроксимувати виразом

$$\Phi_{1/2}(\zeta) = e^\zeta (1 + 0,3e^\zeta)^{-1}$$

6. Домішка утворює два акцепторних рівня у забороненій зоні. Знайдіть залежність концентрації нейтральних, одно- та дворазово заряджених домішкових центрів від положення рівня Фермі та температури.

**Розв'язок.** У результаті перезарядки центрів можуть відбуватись тільки такі переходи з одноразово зарядженого центра:  $N_1 \rightarrow N_0$  та  $N_1 \rightarrow N_{II}$ . Оскільки концентрація  $n$ -разово заряджених домішок пропорційна функції розподілу  $f_{(n)}$  (див. додаток С), можна записати, що

$$\frac{N_1}{N_0} = \frac{g_1}{g_0} e^{\frac{E_F - E_1}{kT}}, \quad (8.214)$$

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} e^{\frac{2E_F - E_1 - E_2}{kT}} e^{-\frac{E_F - E_1}{kT}} \rightarrow \frac{N_2}{N_1} = \frac{g_2}{g_1} e^{\frac{E_F - E_1 - U}{kT}}, \quad U = E_2 - E_1. \quad (8.215)$$

Повна концентрація домішок дорівнює сумі нейтральних, одно- та дворазово заряджених домішок:  $N = N_0 + N_1 + N_2$ . Звідси маємо

$$\frac{N}{N_1} = \frac{N_0}{N_1} + 1 + \frac{N_2}{N_1}$$

або, підставляючи сюди (8.214) та (8.215), отримаємо

$$\frac{N}{N_1} = \frac{g_0}{g_1} e^{\frac{E_F - E_1}{kT}} + 1 + \frac{g_2}{g_1} e^{\frac{E_F - E_1 - U}{kT}} \quad (8.216)$$

і знаходимо, що 
$$N_1 = \frac{Ng_1 e^{\frac{E_F - E_1}{kT}}}{g_0 + g_1 e^{\frac{E_F - E_1}{kT}} + g_2 e^{\frac{2E_F - 2E_1 - U}{kT}}}. \quad (8.217)$$

Підставляючи (8.217) до (8.214) та (8.215), отримуємо

$$N_0 = \frac{Ng_0}{g_0 + g_1 e^{\frac{E_F - E_1}{kT}} + g_2 e^{\frac{2E_F - 2E_1 - U}{kT}}}, \quad (8.218)$$

$$N_2 = \frac{Ng_2 e^{\frac{2E_F - 2E_1 - U}{kT}}}{g_0 + g_1 e^{\frac{E_F - E_1}{kT}} + g_2 e^{\frac{2E_F - 2E_1 - U}{kT}}}. \quad (8.213)$$

Проаналізуємо отриманий результат. При  $E_1 - E_F \gg kT$  маємо: якщо  $E_F < E_1$ , то  $N_0 \approx N$ ,  $N_1 \approx 0$ ,  $N_2 \approx 0$ . Усі домішки знаходяться в нейтральному стані. Коли рівень Фермі потрапляє до інтервалу  $E_1 < E_F < E_1 + U$  за умови, що  $E_F - E_1 \gg kT$  та  $E_F - E_1 - U \gg kT$ , маємо  $N_0 \approx 0$ ,  $N_2 \approx 0$ ,  $N_1 \approx N$ , тобто всі домішки є одноразово іонізованими. За умови, коли  $E_F - E_1 - U \gg kT$ , маємо, що в інтервалі значень рівнів Фермі  $E_F > E_1 + U$  концентрація заряджених центрів розподілена таким чином:  $N_0 \approx 0$ ,  $N_1 \approx 0$ ,  $N_2 \approx N$ , тобто всі домішки є дворазово іонізованими.

### Список літератури

1. Ландау Л.Д., Лифшиц Е.М. Статистическая физика. – М.: Наука, 1976. – Ч. 1.
2. Куреев П.С. Физика полупроводников. – М.: Высшая шк., 1975.
3. Ю Пугер, Кардона Мануэль. Основы физики полупроводников. – М.: Физматлит, 2002.
4. Balkanski M., Wallis R.F. Semiconductor Physics and Applications. – Oxford: Oxford Univ.Press., 2000.
5. Ming-Fu Li. Modern Semiconductor Quantum Physics. – World Scientific, 1995.



# Розділ 9

## ЯВИЩА ПЕРЕНЕСЕННЯ В НАПІВПРОВІДНИКАХ

У цьому розділі розглянемо фізичні явища у напівпровідниках, зумовлені направленим рухом носіїв заряду під дією зовнішніх і внутрішніх сил. Такі явища мають загальну назву явищ перенесення або кінетичних, серед яких – електро- та теплопровідність, гальваноманітні, термомагнітні та термоелектричні явища. Розгляд усіх цих явищ базується на розв'язанні рівняння для функції розподілу носіїв. Це рівняння називається кінетичним рівнянням Больцмана.

### ■ 9.1. Кінетичне рівняння Больцмана

В ідеальному кристалі, на який не діють зовнішні сили, деякий стан, що описується хвильовою функцією  $\Psi_{\mathbf{k}}(\mathbf{r})$ , залишається незмінним протягом скільки завгодно великого проміжку часу. Цьому стану відповідає функція розподілу електронів  $f(\mathbf{r}, \mathbf{k}, t)$ , яка також не змінюється із часом. Якщо систему електронів розташувати у зовнішньому полі  $V(\mathbf{r})$ , то стан кожного електрона у зоні Бриллюена зміниться, оскільки на електрон діятиме сила  $\mathbf{F}_a = -\nabla V(\mathbf{r})$ , обумовлена зовнішнім полем, яка викличе зміну імпульсу електрона

$$\frac{d\mathbf{k}}{dt} = \frac{1}{\hbar} \mathbf{F}_a. \quad (9.1)$$

Звідси

$$\mathbf{k}(t) = \int_0^t \mathbf{F}_a(\xi) d\xi + \mathbf{k}_0(0). \quad (9.2)$$

Ця формула показує: якщо в деякий початковий момент часу  $t = 0$  електрони розподілені за станами згідно із функцією розподілу  $f(\mathbf{r}, \mathbf{k}_0)$ , то в момент  $t$  їхній розподіл визначається іншою функцією

$$f(\mathbf{r}, \mathbf{k}) = f\left(\mathbf{r}, \mathbf{k}_0(t) + \frac{1}{\hbar} \int_0^t d\xi \mathbf{F}_a(\xi)\right), \quad (9.3)$$

тобто зовнішні щодо періодичного поля кристала сили викликають зміни функції розподілу за електронними станами. Таким чином повну похідну функції розподілу за часом можна записати так

$$\frac{df(\mathbf{r}, \mathbf{k}, t)}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{r}} \frac{\partial \mathbf{r}}{\partial t} + \frac{\partial f}{\partial \mathbf{k}} \frac{\partial \mathbf{k}}{\partial t} = \frac{\partial f}{\partial t} + \nabla_{\mathbf{r}} f \cdot \mathbf{v} + \frac{1}{\hbar} \nabla_{\mathbf{k}} f \cdot \mathbf{F}_a. \quad (9.4)$$

Оскільки кількість станів не змінюється із часом (що є відображенням теореми Ліувілля про збереження фазового об'єму за руху системи

вздовж фазових траєкторій), то повна похідна функції розподілу має дорівнювати нулю

$$\frac{\partial f}{\partial t} + \nabla_{\mathbf{r}} f \cdot \mathbf{v} + \frac{1}{\hbar} \nabla_{\mathbf{k}} f \cdot \mathbf{F}_a = 0. \quad (9.5)$$

Сила, що входить до цього рівняння, складається із двох частин – сили  $\mathbf{F}$ , породженої зовнішніми макроскопічними полями, які викликають зміну  $f$ , і сили, що намагається повернути  $f$  до  $f_0$ , природою якої є взаємодія електрона з неоднорідностями кристалічного потенціалу (розсіювання на дефектах, коливаннях ґратки тощо).

$$\mathbf{F}_a = \mathbf{F} + \mathbf{F}_i. \quad (9.6)$$

Сили, викликані зовнішніми полями, приводять до відносно повільних змін у стані частинки. Інша справа – внутрішні сили, спричинені швидкими взаємодіями частинок у кристалі із сильно локалізованими збуреннями кристалічного потенціалу. Характерні розміри області локалізації складають величини порядку кількох сталих ґратки ( $\sim 10^{-7}$  см), а характерні теплові швидкості частинок –  $10^7$  см/с. Таким чином, час взаємодії частинки з локальним збуренням складатиме величину  $\sim 10^{-14}$  с. Така швидкодіюча взаємодія частинки приводить до помітної зміни швидкості та квазіімпульсу частинки. Це означає, що вплив короткодіючої сили, породженої неоднорідностями кристалічного потенціалу, аналогічний зіткненню електрона із розсіювачем, тому ефекти, породжені цією силою, називатимемо ефектами зіткнень. Таким чином, (9.5) можна переписати у вигляді

$$\frac{\partial f}{\partial t} + \nabla_{\mathbf{r}} f \cdot \mathbf{v} + \frac{1}{\hbar} \nabla_{\mathbf{k}} f \cdot \mathbf{F} = -\frac{1}{\hbar} \nabla_{\mathbf{k}} f \cdot \mathbf{F}_i. \quad (9.7)$$

Ясно, що на відміну від дії зовнішнього поля, яке може викликати впорядкований рух частинок, зіткнення мають хаотичний у просторі та часі характер, для опису якого варто користуватися статистичними підходами. Права частина (9.7) визначає швидкість зміни функції розподілу, що зумовлена процесами зіткнень і часто позначається як  $(\partial f / \partial t)_{col}$ . Для визначення правої частини, яка має спеціальну назву інтеграла зіткнень, припустимо, що в результаті зіткнень частинки переходять зі станів  $(\mathbf{r}, \mathbf{k})$  до станів  $(\mathbf{r}, \mathbf{k}')$ . Нехай ймовірністю такого переходу в одиницю часу є  $w(\mathbf{r}, \mathbf{k}')$ . Ясно, що внаслідок зіткнень координата не змінюється так сильно, як імпульс, тому можна вважати, що ймовірність є функцією тільки імпульсів (або, що теж саме, – хвильових векторів). Виділимо два елементарні об'єми в околицях точок  $\mathbf{k}$  і  $\mathbf{k}'$  –  $d\mathbf{k}$  і  $d\mathbf{k}'$ , відповідно. Числом станів у цих елементарних об'ємах в імпульсному просторі із врахуванням спіну є  $d\mathbf{k}/4\pi^3$  та  $d\mathbf{k}'/4\pi^3$ , відповідно. Із них число зайнятих станів до й після зіткнення



$$N_{\text{зайн.}} = f(\mathbf{r}, \mathbf{k}) \frac{d\mathbf{k}}{4\pi^3} \quad \text{та} \quad N'_{\text{зайн.}} = f(\mathbf{r}, \mathbf{k}') \frac{d\mathbf{k}'}{4\pi^3}, \quad (9.8)$$

а число вільних (незайнятих) станів

$$N_{\text{вільн.}} = [1 - f(\mathbf{r}, \mathbf{k})] \frac{d\mathbf{k}}{4\pi^3} \quad \text{та} \quad N'_{\text{вільн.}} = [1 - f(\mathbf{r}, \mathbf{k}')] \frac{d\mathbf{k}'}{4\pi^3}. \quad (9.9)$$

У результаті зіткнень частина електронів переходить зі станів  $d\mathbf{k}$  до станів  $d\mathbf{k}'$  і навпаки, частина електронів перейде зі станів  $d\mathbf{k}'$  до станів  $d\mathbf{k}$ . Із врахуванням принципу Паулі електрони можуть переходити тільки із зайнятих станів до вільних. Тому кількістю електронів, що за час  $dt$  залишають об'єм  $d\mathbf{k}$ , буде

$$dN \uparrow = dt \{ w(\mathbf{k}, \mathbf{k}') N_{\text{зайн.}} N'_{\text{вільн.}} \}, \quad (9.10)$$

а кількістю електронів, які за той самий час залишили об'єм  $d\mathbf{k}'$  і перейшли до станів  $d\mathbf{k}$ ,

$$dN \downarrow = dt \{ w(\mathbf{k}', \mathbf{k}) N'_{\text{зайн.}} N_{\text{вільн.}} \}. \quad (9.11)$$

Таким чином, унаслідок прямих переходів (які зменшують кількість електронів в об'ємі  $d\mathbf{k}$ ) і зворотних переходів (які збільшують кількість електронів) за час  $dt$  число станів в об'ємі  $d\mathbf{k}$  зміниться на величину

$$dN = dN \downarrow - dN \uparrow. \quad (9.12)$$

Повне число змін станів за час  $dt$  в об'ємі  $d\mathbf{k}$  можна знайти за допомогою інтегрування за всіма можливими  $\mathbf{k}'$ , тобто за всією зоною Бриллюена

$$dN_{dk} = dt \frac{d\mathbf{k}}{4\pi^3} \int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} \{ w(\mathbf{k}', \mathbf{k}) f(\mathbf{r}, \mathbf{k}') [1 - f(\mathbf{r}, \mathbf{k})] - w(\mathbf{k}, \mathbf{k}') f(\mathbf{r}, \mathbf{k}) [1 - f(\mathbf{r}, \mathbf{k}')] \}. \quad (9.13)$$

З іншого боку, число зайнятих станів у кожний момент часу визначається величиною  $f(\mathbf{r}, \mathbf{k})(d\mathbf{k}/4\pi^3)$ , і за короткий проміжок часу  $dt$  змінюється як  $(df/dt)_{col} dt(d\mathbf{k}/4\pi^3)$ . Таким чином,

$$\left( \frac{\partial f}{\partial t} \right)_{col} = \int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} \{ w(\mathbf{k}', \mathbf{k}) f(\mathbf{r}, \mathbf{k}') [1 - f(\mathbf{r}, \mathbf{k})] - w(\mathbf{k}, \mathbf{k}') f(\mathbf{r}, \mathbf{k}) [1 - f(\mathbf{r}, \mathbf{k}')] \}. \quad (9.14)$$

Якщо ймовірності прямих і зворотних переходів є рівними, то підінтегральна функція суттєво спрощується. Дійсно, підінтегральний вираз можна представити у вигляді

$$\begin{aligned} w(\mathbf{k}', \mathbf{k}) f(\mathbf{r}, \mathbf{k}') - w(\mathbf{k}, \mathbf{k}') f(\mathbf{r}, \mathbf{k}) - [w(\mathbf{k}', \mathbf{k}) - w(\mathbf{k}, \mathbf{k}')] f(\mathbf{r}, \mathbf{k}) f(\mathbf{r}, \mathbf{k}') = \\ = w(\mathbf{k}, \mathbf{k}') [f(\mathbf{r}, \mathbf{k}') - f(\mathbf{r}, \mathbf{k})]. \end{aligned}$$

Тоді інтегралом зіткнень є

$$\left(\frac{\partial f}{\partial t}\right)_{col} = \int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} w(\mathbf{k}', \mathbf{k}) [\mathbf{f}(\mathbf{r}, \mathbf{k}') - \mathbf{f}(\mathbf{r}, \mathbf{k})], \quad (9.15)$$

а питоме кінетичне рівняння запишеться як

$$\frac{\partial f}{\partial t} = -\mathbf{v} \cdot \nabla_{\mathbf{r}} f - \frac{1}{\hbar} \mathbf{F} \cdot \nabla_{\mathbf{k}} f + \int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} w(\mathbf{k}', \mathbf{k}) [f(\mathbf{r}, \mathbf{k}') - f(\mathbf{r}, \mathbf{k})]. \quad (9.16)$$

Це рівняння називається кінетичним рівнянням Больцмана та пов'язує еволюцію функції розподілу в часі із рухом частинок, дією зовнішніх полів і зіткненнями носіїв заряду з локальними неоднорідностями потенціалу ґратки. У стаціонарному випадку похідна в лівій частині (9.16) занулюється, і в результаті маємо рівняння балансу

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f + \frac{1}{\hbar} \mathbf{F} \cdot \nabla_{\mathbf{k}} f = \int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} w(\mathbf{k}', \mathbf{k}) [f(\mathbf{r}, \mathbf{k}') - f(\mathbf{r}, \mathbf{k})], \quad (9.17)$$

яке показує, що в стаціонарному випадку зміна функції розподілу, обумовлена дією зовнішніх полів і рухом частинок, компенсується зіткненнями носіїв із локальними порушеннями періодичності поля ґратки.

## ■ 9.2. Час релаксації

Розв'язок (9.16) у загальному випадку є дуже складною проблемою, тому часто використовують наближені методи отримання його розв'язків. Найпоширенішим наближенням є наближення часу релаксації, загальна ідея якого полягає в наступному. Припустимо, що в момент часу  $t = 0$  у системі, що описується нерівноважною функцією розподілу, вимикаються зовнішні поля. Після цього система внаслідок зіткнень приходить у рівноважний стан, який описується рівноважною функцією розподілу  $f_0$ . Це означає, що після вимкнення зовнішнього поля функція розподілу змінюється тільки завдяки зіткненням електронів із дефектами ґратки

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{col}. \quad (9.18)$$

Зрозуміло, що чим більшим є відхилення від рівноваги, тим довше система релаксуватиме до рівноважного стану. Отже, можна стверджувати, що швидкість зміни функції розподілу є функцією різниці нерівноважної і рівноважної функцій розподілу, тобто

$$\frac{\partial f}{\partial t} = F(f - f_0). \quad (9.19)$$

За невеликих відхилень функції розподілу від рівноважної функцію в правій частині (9.19) можна розкласти в ряд Тейлора за малим параметром  $f - f_0$ . Оскільки причиною повернення системи до стану рівноваги є тільки процеси зіткнення, то  $F(f - f_0 = 0) = 0$ . Таким чином, якщо відхилення розподілу носіїв від рівноважного незначне, то за відсутності зовнішнього поля швидкість зміни функції розподілу буде пропорційною величині відхилення функції від рівноважної, тобто

$$\frac{\partial f}{\partial t} = \frac{f - f_0}{\tau(\mathbf{k})}. \quad (9.20)$$

Звідси знайдемо 
$$\Delta f = \Delta f|_{t=0} e^{-t/\tau}, \quad (9.21)$$

тобто після припинення дії зовнішніх полів різниця  $\Delta f = f - f_0$  зменшується за експоненціальним законом зі сталою часу  $\tau$ . Час релаксації  $\tau$  є середнім часом, протягом якого в системі існує нерівноважний розподіл носіїв після припинення дії зовнішніх полів. Оскільки в наближенні, яке ми використовуємо, зовнішні сили не викликають великих відхилень розподілу носіїв від рівноважного, то нерівноважну функцію розподілу представимо як суму

$$f(\mathbf{r}, \mathbf{k}) = f_0(\mathbf{r}, \mathbf{k}) + f_1(\mathbf{r}, \mathbf{k}). \quad (9.22)$$

Оскільки функція розподілу є скалярною функцією хвильового вектора, запишемо поправку до рівноважної функції розподілу у вигляді

$$f_1(\mathbf{r}, \mathbf{k}) = -\frac{\partial f_0}{\partial E} \mathbf{k} \cdot \chi(E), \quad (9.23)$$

де  $\chi(E)$  – деякий вектор, що залежить тільки від енергії носія заряду.

Зробимо ще одне важливе припущення. Вважатимемо, що час релаксації не залежить від зовнішніх полів, тобто  $\tau(\mathbf{k})$  описує процеси розсіювання також за наявності зовнішніх сил. Тоді рівняння для стаціонарного випадку можна записати як

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f + \frac{1}{\hbar} \mathbf{F} \cdot \nabla_{\mathbf{k}} f = -\frac{f - f_0}{\tau(\mathbf{k})} = -\frac{f_1(\mathbf{k})}{\tau(\mathbf{k})}. \quad (9.24)$$

Записуючи інтеграл зіткнень в явному вигляді, із цього рівняння маємо

$$-\frac{f_1(\mathbf{r}, \mathbf{k})}{\tau(\mathbf{k})} = \int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} w(\mathbf{k}', \mathbf{k}) [f(\mathbf{r}, \mathbf{k}') - f(\mathbf{r}, \mathbf{k})], \quad (9.25)$$

яке перепишемо у вигляді

$$\frac{1}{\tau(\mathbf{k})} = -\int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} w(\mathbf{k}', \mathbf{k}) \frac{f_1(\mathbf{r}, \mathbf{k}') - f_1(\mathbf{r}, \mathbf{k})}{f_1(\mathbf{r}, \mathbf{k})} \quad (9.26)$$

і далі, використовуючи визначення (9.23),

$$\frac{1}{\tau(\mathbf{k})} = \int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} w(\mathbf{k}', \mathbf{k}) \left[ 1 - \frac{\frac{\partial f_0}{\partial E'} \mathbf{k}' \cdot \boldsymbol{\chi}(E')}{\frac{\partial f_0}{\partial E} \mathbf{k} \cdot \boldsymbol{\chi}(E)} \right]. \quad (9.27)$$

Зіткнення є пружними, якщо кінетична енергія частинок у процесі зіткнень зберігається. За непружних зіткнень енергія частинки змінюється. Якщо зіткнення електронів зі структурними порушеннями є пружними, то в електрона змінюється лише напрямок хвильового вектора та виконуються рівності:  $k = k'$ ,  $v = v'$ ,  $E = E'$ . Це означає, що  $\partial f_0 / \partial E' = \partial f_0 / \partial E$ . Тоді із (9.27) випливає, що час релаксації для пружних зіткнень визначається як

$$\frac{1}{\tau(\mathbf{k})} = \int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} w(\mathbf{k}', \mathbf{k}) \left[ 1 - \frac{\mathbf{k}' \cdot \boldsymbol{\chi}(E')}{\mathbf{k} \cdot \boldsymbol{\chi}(E)} \right]. \quad (9.28)$$

У межах використаного наближення припустимо тепер, що вектори  $\mathbf{k}$  та  $\boldsymbol{\chi}$  мають однаковий напрямок, а вектор  $\mathbf{k}'$  складає з  $\mathbf{k}$  кут  $\vartheta$ . Тоді  $\mathbf{k}' \cdot \boldsymbol{\chi}(E) = k\chi \cos \vartheta$ , а  $\mathbf{k} \cdot \boldsymbol{\chi}(E) = k\chi$  і (9.28) запишеться як

$$\frac{1}{\tau(\mathbf{k})} = \int_{\Omega} \frac{d\mathbf{k}'}{4\pi^3} w(\mathbf{k}', \mathbf{k}) [1 - \cos \vartheta]. \quad (9.29)$$

Для обчислення часу релаксації необхідно знати механізм розсіювання носіїв. Як відомо з теорії розсіювання, основною характеристикою процесу розсіювання є його диференціальний переріз: відношення числа частинок, які відхиляються одним розсіювачем за 1 с на кут  $d\vartheta$  в одиничний об'ємний кут до потоку частинок, які проходять через одиничну площину за 1 с. Наприклад, за розсіювання електрона на іонах домішки диференціальний переріз розсіювання описується формулою Резерфорда (див. додаток D)

$$\sigma(\vartheta) = \left( \frac{Ze^2}{2\epsilon_r m^* v^2} \right)^2 \frac{1}{\sin^4(\vartheta/2)}. \quad (9.30)$$

Такою самою формулою описується розсіювання носіїв на будь-якому зарядженому центрі. Зрозуміло, що в загальному випадку задача обчислення диференціального перерізу розсіювання на зарядженому центрі є досить складною. Але якщо вважати, що рух електрона в кристалі може бути врахований введенням ефективної маси, а екранування потенціалу центру в кристалі описується введенням діелектричної функції, задача суттєво спрощується та як результат отримуємо (9.30). На прикладі розсіювання на іонізованих домішках проаналізуємо зв'язок між часом релаксації та ефективним перерізом розсіювання. Припустимо, що в кристалі одиничного об'єму є  $N_i$  іонів

домішки. Тоді число носіїв, що розсіяні всіма центрами в тілесний кут  $d\Omega$ , складає

$$w(\mathbf{k}, \mathbf{k}')nd\Omega = \sigma(\vartheta)N_I n v d\Omega, \quad (9.31)$$

звідки

$$w(\mathbf{k}, \mathbf{k}') = \sigma(\vartheta)N_I v \quad (9.32)$$

або  $\tau = 1/\sigma N v$  – звичайний вигляд для характерних часів у фізиці. Ймовірність переходу зі стану  $\mathbf{k}$  до стану  $\mathbf{k}'$  за розсіювання електронів на іонах домішки визначається диференціальним перерізом розсіювання, кількістю центрів розсіювання і швидкістю руху носіїв. Оскільки маса іона на чотири порядки більша за масу електрона, а швидкість іона набагато менша за швидкість електрона, розсіювання електрона на іоні можна вважати пружним. Якщо при зіткненні носії розсіюються в тілесний кут  $d\Omega$ , то часом релаксації є

$$\frac{1}{\tau(\mathbf{k})} = \int_{(\Omega)} w(\mathbf{k}', \mathbf{k}) [1 - \cos\vartheta] d\Omega \quad (9.33)$$

або, враховуючи (9.32),

$$\frac{1}{\tau(\mathbf{k})} = 2\pi \cdot N_I v(\mathbf{k}) \int_{(\vartheta)} \sigma(\vartheta) [1 - \cos\vartheta] \sin\vartheta \cdot d\vartheta. \quad (9.34)$$

Іншими словами, для обчислення часу релаксації необхідно знати ефективний переріз розсіювання і за подальшого проведення інтегрування, згідно із (9.34), можна отримати  $\tau(\mathbf{k})$ . Величину

$$\sigma_C = 2\pi \cdot \int_{(\vartheta)} \sigma(\vartheta) [1 - \cos\vartheta] \sin\vartheta \cdot d\vartheta \quad (9.35)$$

називають ефективним перерізом розсіювання або транспортним ефективним перерізом за ізотропного розсіювання. Формулу, аналогічну (9.35), можна отримати й у більш загальному випадку анізотропного пружного розсіювання

$$\sigma_C = 2\pi \int_0^{\pi} \int_0^{2\pi} \sigma(\vartheta, \varphi) [1 - \cos\vartheta] \sin\vartheta \cdot d\vartheta \cdot d\varphi. \quad (9.36)$$

Таким чином, час релаксації визначається ефективним перерізом розсіювання та концентрацією іонів домішок, на яких відбувається розсіювання носіїв.

$$\tau(\mathbf{k}) = \frac{1}{N_I \sigma_C v(\mathbf{k})}, \quad (9.37)$$

$$\text{а довжина вільного пробігу } l(\mathbf{k}) = \tau(\mathbf{k})v(\mathbf{k}) = \frac{1}{\sigma_C N_I}. \quad (9.38)$$

Формулу, аналогічну (9.37), можна отримати і в загальному випадку розсіювання частинок на центрах розсіювання будь-якої природи. У кристалі носії можуть брати участь у процесах розсіювання на

- ♦ іонах домішок;
- ♦ нейтральних атомах домішок;
- ♦ теплових коливаннях ґратки;
- ♦ вакансіях і точкових дефектах;
- ♦ дислокаціях;
- ♦ границях кристалитів, площинах двійникування, границях кристала;
- ♦ носіях заряду.

Можна показати (див. додаток D), що час релаксації для розсіювання носіїв заряду на іонах домішок пропорційний енергії у ступені 3/2 та кореню квадратному з ефективної маси

$$\tau_I = \tau_{0I} E^{3/2}, \quad (9.39)$$

де

$$\tau_{0I} = \frac{(2m^*)^{1/2} \epsilon_r^2}{\pi Z^2 e^4 N_I \ln \left[ 1 + \left( \frac{\epsilon_r E}{Ze^2 N_I^{1/3}} \right)^2 \right]}. \quad (9.40)$$

При цьому довжина вільного пробігу

$$l_I = \sqrt{2/m^*} \tau_{0I} E^2. \quad (9.41)$$

За розсіювання носіїв заряду на нейтральних домішкових атомах (актуальне тільки в області низьких температур за концентрації іонізованих атомів домішки, істотно меншої від концентрації нейтральних атомів) варто зважати на те, що тепер процес розсіювання може проходити двома каналами (пряме пружне розсіювання та обмін падаючого електрона на електрон атома). У моделі, де електрони розсіюються на воднеподібній домішці в середовищі з ефективною діелектричною сталою  $\epsilon_r$ , час релаксації визначається формулою

$$\tau_A = \frac{e^2 (m^*)^2}{20 \epsilon_r \hbar^3} \frac{1}{N_A}, \quad (9.42)$$

тобто він при розсіюванні на нейтральних атомах домішок не залежить від температури та енергії носіїв заряду, а визначається тільки концентрацією домішок. Розрахунки показують, що у випадку розсіювання носіїв на дислокаціях час релаксації не залежить від температури, а визначається щільністю дислокацій на одиницю поверхні  $N_D$  і швидкістю носіїв

$$\tau_D = \frac{3}{8Rv} \frac{1}{N_D}. \quad (9.43)$$

У цій формулі  $R$  є ефективним діаметром ядра дислокації. При цьому характерні часи релаксації  $\tau_D \sim 10^{-9}$  с. За розсіювання носіїв заряду на теплових коливаннях ґратки виявляється, що час релаксації за розсіювання носіїв на довгохвильових повздожжніх фононах визначається формулою

$$\tau_l = \frac{9\pi \cdot \hbar^4 v_S^2}{\sqrt{8g^2 a^3 k}} \frac{M}{(m^*)^{3/2}} T^{-1} E^{-1/2}, \quad (9.44)$$

де  $g$  – константа електрон-фононного зв'язку,  $a$  – стала ґратки,  $M$  – маса ядра атома. Для германію за кімнатної температури  $\tau_l \sim 10^{-12} - 10^{-13}$  с.

Таким чином, ми бачимо, що за підвищення енергії носіїв час релаксації за розсіювання на іонізованих домішках збільшується, а за розсіювання на акустичних коливаннях ґратки – зменшується.

### ■ 9.3. Електропровідність напівпровідників

Визначення нерівноважної функції розподілу дає можливість описати у напівпровідниках усі явища перенесення. Для прикладу розглянемо невідроджений напівпровідник, що знаходиться в електричному та магнітному полях і характеризується деяким розподілом температури за об'ємом так, що градієнт температури не є нульовим. За наявності зовнішніх полів на електрон діє сила

$$\mathbf{F} = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (9.45)$$

Тоді стаціонарне кінетичне рівняння запишеться як

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f - \frac{e}{\hbar} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{k}} f = -\frac{1}{\tau_e} f_1(\mathbf{k}). \quad (9.46)$$

Якщо система знаходиться в майже рівноважному стані, то нерівноважна функція розподілу слабо відрізняється від рівноважної. Скористаємось цією обставиною і запишемо (9.46) у вигляді

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 - \frac{e}{\hbar} \mathbf{E} \cdot \nabla_{\mathbf{k}} f_0 - \frac{e}{\hbar} (\mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{k}} (f_0 + f_1(\mathbf{k})) = -\frac{1}{\tau_e} f_1(\mathbf{k}). \quad (9.47)$$

Подальший аналіз цього рівняння проведемо окремо для електронів і дірок. Виконаємо деякі перетворення. Почнемо з першого доданку лівої частини (9.47), пам'ятаючи про те, що рівноважна функція розподілу явно залежить від температури та рівня Фермі

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 = \mathbf{v} \cdot \left( \frac{\partial f_0}{\partial T} \nabla T + \frac{\partial f_0}{\partial E_F} \nabla E_F \right). \quad (9.48)$$

А оскільки для невідродженого електронного газу в напівпровіднику можна записати

$$\frac{\partial f_0}{\partial T} = \frac{E - E_F}{kT^2} e^{-\frac{(E - E_F)}{kT}} = -\frac{E - E_F}{T} \frac{\partial f_0}{\partial E} \quad (9.49)$$

та

$$\frac{\partial f_0}{\partial E_F} = -\frac{\partial f_0}{\partial E}, \quad (9.50)$$

то (9.48) перепишеться у вигляді

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 = \frac{\partial f_0}{\partial E} \mathbf{v} \cdot \left( \frac{E_F - E}{T} \nabla T - \nabla E_F \right). \quad (9.51)$$

Крім того, маємо

$$\nabla_{\mathbf{k}} f_0 = \frac{\partial f_0}{\partial E} \frac{\hbar^2}{m^*} \mathbf{k} = \frac{\partial f_0}{\partial E} \hbar \mathbf{v}. \quad (9.52)$$

Розберемося тепер із доданком у лівій частині (9.47), що містить магнітне поле

$$\mathbf{v} \times \mathbf{B} \cdot \nabla_{\mathbf{k}} f = \hbar \frac{\partial f_0}{\partial E} \mathbf{v} \times \mathbf{B} \cdot \mathbf{v} + \mathbf{v} \times \mathbf{B} \cdot \nabla_{\mathbf{k}} f_1 = \mathbf{v} \times \mathbf{B} \cdot \nabla_{\mathbf{k}} f_1. \quad (9.53)$$

Тут ми використали той факт, що доданок, пропорційний похідній функції розподілу за енергією, містить змішаний векторний добуток  $\mathbf{v} \times \mathbf{B} \cdot \mathbf{v} = (\mathbf{v} \times \mathbf{v}) \cdot \mathbf{B} = 0$ . Для подальшого аналізу (9.47) знову запишемо нерівноважний доданок функції розподілу  $f_1$  через вектор  $\chi_n(E)$  (індекс  $n$  підкреслює, що йдеться про електронну компоненту носіїв заряду)

$$f_1 = -\frac{\partial f_0}{\partial E} \mathbf{k} \cdot \chi_n(E). \quad (9.54)$$

Далі обчислення похідної від цієї функції за хвильовим вектором  $\mathbf{k}$  проводимо із врахуванням того, що  $\nabla_{\mathbf{k}}$  є векторним диференціальним оператором, який діє на скалярний добуток двох векторів

$$\begin{aligned} \nabla_{\mathbf{k}} f_1 &= -\nabla_{\mathbf{k}} \left[ \frac{\partial f_0}{\partial E} \mathbf{k} \cdot \chi_n(E) \right] = -\frac{\partial f_0}{\partial E} \chi_n(E) - \left[ \mathbf{k} \cdot \frac{\partial}{\partial E} \left( \frac{\partial f_0}{\partial E} \chi_n(E) \right) \right] \frac{\partial E}{\partial \mathbf{k}} = \\ &= -\frac{\partial f_0}{\partial E} \chi_n(E) - \left[ \hbar \mathbf{k} \cdot \frac{\partial}{\partial E} \left( \frac{\partial f_0}{\partial E} \chi_n(E) \right) \right] \mathbf{v}. \end{aligned} \quad (9.55)$$

Підставимо цей вираз до (9.53). Зважаючи на те, що в останньому доданку у (9.55) вираз у дужках є деяким скаляром (позначимо його тимчасово як  $\Gamma$ ), що помножений на вектор  $\mathbf{v}$ , маємо



$$\mathbf{v} \times \mathbf{B} \cdot \nabla_{\mathbf{k}} f = -\frac{\partial f_0}{\partial E} \mathbf{v} \times \mathbf{B} \cdot \boldsymbol{\chi}_n - \Gamma(\mathbf{v} \times \mathbf{B} \cdot \mathbf{v}) = -\frac{\partial f_0}{\partial E} \mathbf{B} \times \boldsymbol{\chi}_n \cdot \mathbf{v}. \quad (9.56)$$

При обчисленні цього виразу враховуємо, що змішаний добуток трьох векторів, два з яких однакові, дорівнює нулю. Підставимо тепер вирази із (9.51), (9.52) та (9.56) до рівняння (9.47)

$$f_1(\mathbf{k}) = -\tau_e \frac{\partial f_0}{\partial E} \mathbf{v} \left\{ -\frac{E - E_F}{T} \nabla T - \nabla E_F - eE + \frac{e}{\hbar} \mathbf{B} \times \boldsymbol{\chi}_n \right\}. \quad (9.57)$$

Враховуючи визначення (9.54) й те, що  $\mathbf{E} = -\nabla\varphi$ , отримуємо рівняння для визначення вектора  $\boldsymbol{\chi}_n$ , отже і нерівноважного доданку функції розподілу  $f_1(\mathbf{k})$

$$\boldsymbol{\chi}_n = -\tau_e \frac{\hbar}{m_n^*} \left\{ \frac{E - E_F}{T} \nabla T + \nabla(E_F - e\varphi) - \frac{e}{\hbar} \mathbf{B} \times \boldsymbol{\chi}_n \right\}. \quad (9.58)$$

Оскільки у явищах електроперенесення беруть участь як електрони, так і дірки, то необхідно також знати нерівноважний доданок до рівноважної функції розподілу дірок. Врахуємо тепер, що енергія дірок записується як

$$E' = -\frac{\hbar^2 \mathbf{k}^2}{2m_p^*}. \quad (9.59)$$

Концентрація вільних носіїв у невиродженому напівпровіднику визначається формулами

$$n = N_C \Phi_{1/2}(\zeta), \quad p = N_V \Phi_{1/2}(-\zeta - \varepsilon_i), \quad (9.60)$$

де  $\varepsilon_i = E_g / k_B T$  і  $\zeta = (E_F - E_C) / k_B T$  ( $E_g$  – ширина забороненої зони). Для дірок можна використати такий самий вираз, як і для електронів із заміною  $m_n^*$  на  $m_p^*$ ,  $E$  на  $E'$  та  $E_F$  на

$$E'_F = -E_g - E_F. \quad (9.61)$$

Таким чином, аналогічно (9.58), для дірок можна записати

$$\boldsymbol{\chi}_p = -\tau_p \frac{\hbar}{m_p^*} \left\{ \frac{E' + E_g + E_F}{T} \nabla T - \nabla(E_F - e\varphi) + \frac{e}{\hbar} \mathbf{B} \times \boldsymbol{\chi}_p \right\}. \quad (9.62)$$

Для розв'язку рівнянь (9.58) та (9.62) візьмемо до уваги, що вони є векторними лінійними рівняннями типу

$$\mathbf{x} = \mathbf{a} + \mathbf{b} \times \mathbf{x}. \quad (9.63)$$

Для розв'язку (9.63) зазначимо, що скалярний добуток обох його частин із вектором  $\mathbf{b}$  дає

$$\mathbf{b} \cdot \mathbf{x} = \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} \times \mathbf{x} = \mathbf{b} \cdot \mathbf{a}, \quad (9.64)$$

а векторний добуток обох частин із вектором  $\mathbf{b}$  приводить до рівності

$$\mathbf{b} \times \mathbf{x} = \mathbf{b} \times \mathbf{a} + \mathbf{b} \times \mathbf{b} \times \mathbf{x} = \mathbf{b} \times \mathbf{a} + \mathbf{b}(\mathbf{b} \cdot \mathbf{x}) - \mathbf{x}(\mathbf{b} \cdot \mathbf{b}). \quad (9.65)$$

Таким чином, (9.63) зводиться до

$$\mathbf{x} = \mathbf{a} + \mathbf{b} \times \mathbf{a} + \mathbf{b}(\mathbf{b} \cdot \mathbf{a}) - b^2 \mathbf{x},$$

звідки маємо розв'язок у вигляді

$$\mathbf{x} = \frac{1}{1+b^2} [\mathbf{a} + \mathbf{b} \times \mathbf{a} + \mathbf{b}(\mathbf{b} \cdot \mathbf{a})], \quad (9.66)$$

використовуючи який, маємо із (9.58) і (9.62)

$$\begin{aligned} \chi_n = & \frac{-\xi_n}{1+\xi_n^2\beta^2} \left\{ \frac{E-E_F}{T} \nabla T + \nabla(E_F - e\varphi) + \right. \\ & \left. + \xi_n \beta \left( \frac{E-E_F}{T} \nabla T + \nabla(E_F - e\varphi) \right) + \xi_n^2 \beta \left( \frac{E-E_F}{T} \nabla T + \nabla(E_F - e\varphi) \right) \beta \right\}, \end{aligned} \quad (9.67)$$

та

$$\begin{aligned} \chi_p + & \frac{-\xi_p}{1+\xi_p^2\beta^2} \left\{ \frac{E'+E_g+E_F}{T} \nabla T - \nabla(E_F - e\varphi) - \right. \\ & \left. - \xi_p \beta \left( \frac{E'+E_g+E_F}{T} \nabla T - \nabla(E_F - e\varphi) \right) + \xi_p^2 \beta^2 \left( \frac{E'+E_g+E_F}{T} \nabla T - \nabla(E_F - e\varphi) \right) \beta \right\}. \end{aligned} \quad (9.68)$$

Тут використано позначення

$$\xi_n = \frac{\hbar \tau_e}{m_n^*}, \quad \xi_p = \frac{\hbar \tau_p}{m_p^*}, \quad \beta = \frac{e}{\hbar} \mathbf{B}. \quad (9.69)$$

Таким чином, рівняння (9.67)–(9.68) визначають нерівноважний доданок до рівноважної функції розподілу носіїв заряду у напівпровіднику за незначного відхилення системи від рівноваги під дією зовнішньої сили.

Спочатку, використовуючи знайдені розв'язки для функції розподілу електронів і дірок напівпровідника, обчислимо питому електричну провідність однорідного неvirодженого напівпровідника за відсутності магнітного поля та градієнта температури. Для простоти вважатимемо, що ізоенергетичні поверхні напівпровідника є сферами, і дія зовнішнього електричного поля не приводить до міждолинного розсіювання. Нехай на напівпровідник діє постійне електричне поле  $\mathbf{E}$ . В елементі фазового простору  $d\mathbf{k}$  кристала одиничного об'єму знаходимо

$$dn = f_n(\mathbf{k}) \frac{d\mathbf{k}}{4\pi^3} \quad (9.70)$$

електронів, які завдяки дії зовнішнього поля рухаються зі швидкістю  $\mathbf{v}$ , тим самим створюючи електричний струм з елементарною густиною

$$d\mathbf{J}_n = -e\mathbf{v}f_n(\mathbf{k})\frac{d\mathbf{k}}{4\pi^3}. \quad (9.71)$$

З іншого боку, внесок до струму дає і рух дірок

$$d\mathbf{J}_p = e\mathbf{v}f_p(\mathbf{k})\frac{d\mathbf{k}}{4\pi^3} \quad (9.72)$$

так, що повною густиною струму є

$$\mathbf{J} = -\frac{e}{4\pi^3} \int_{\Omega} d\mathbf{k} \mathbf{v} f_n(\mathbf{k}) + \frac{e}{4\pi^3} \int_{\Omega} d\mathbf{k} \mathbf{v} f_p(\mathbf{k}). \quad (9.73)$$

Якщо до підінтегральних виразів у (9.73) підставити рівноважну функцію розподілу, то в результаті матимемо нульову густину струму. Дійсно, оскільки рівноважна функція розподілу є парною функцією хвильового вектора, а швидкість просто пропорційна  $\mathbf{k}$ , то матимемо непарну підінтегральну функцію, а межі інтегрування будуть симетричними. Фізична причина цього факту стає прозорою, якщо згадати, що кожному стану, який характеризується хвильовим вектором  $\mathbf{k}$ , відповідає стан  $-\mathbf{k}$ , отже кожному елементарному струму  $d\mathbf{J}$  відповідає елементарний струм  $-d\mathbf{J}$ . У результаті сумарний струм виявиться нульовим. Таким чином, використовуючи представлення для нерівноважного доданку функції розподілу через вектори  $\chi_{n(p)}$ , знайдемо

$$\mathbf{J} = \frac{e\hbar}{4\pi^3 m_n^*} \int_{\Omega} \frac{\partial f_0}{\partial E} (\chi_n \cdot \mathbf{k}) \mathbf{k} d\mathbf{k} - \frac{e\hbar}{4\pi^3 m_p^*} \int_{\Omega} \frac{\partial f_0}{\partial E} (\chi_p \cdot \mathbf{k}) \mathbf{k} d\mathbf{k}. \quad (9.74)$$

Оскільки напівпровідник є однорідним, градієнт від функції розподілу дорівнює нулю. З іншого боку, відсутність градієнта температур вимагає тепер, щоб нульовим був і градієнт рівня Фермі (див. (9.51)). Тоді, враховуючи, що  $\mathbf{V} = 0$ , із рівнянь (9.67) і (9.68) знайдемо

$$\chi_n = -\xi_n \mathbf{E}, \quad (9.75)$$

$$\chi_p = \xi_p \mathbf{E}. \quad (9.76)$$

Обчислення інтегралів у (9.74) проводитимемо окремо для електронів і дірок. Для похідної функції розподілу електронів запишемо

$$\frac{\partial f_0}{\partial E} = -\frac{1}{k_B T} e^{-\frac{E-E_F}{k_B T}}. \quad (9.77)$$

Пам'ятаючи, що якщо за початок відліку енергії брати дно зони провідності, то щільністю електронів для невідродженого напівпровідника є

$$n = N_c e^{\frac{E_F}{k_B T}}, \quad (9.78)$$

перепишемо (9.77) у вигляді

$$\frac{\partial f_0}{\partial E} = -\frac{1}{k_B T} \frac{n}{N_c} e^{-\frac{E}{k_B T}}. \quad (9.79)$$

Після підстановки виразів із (9.79) та (9.75) до (9.74) для електронної складової струму отримаємо

$$\mathbf{J}_n = \frac{e^2 \hbar^2 n}{4\pi^3 (m_n^*)^2 k_B T N_c} \int_{\Omega} \tau_e e^{-\frac{E}{k_B T}} (\mathbf{k} \cdot \mathbf{E}) \cdot \mathbf{k} d\mathbf{k}. \quad (9.80)$$

Якщо перейдемо до сферичної системи координат, де полярна вісь направлена вздовж напрямку електричного поля, то можна записати

$$\begin{aligned} \int_{\Omega} F(\mathbf{k}) (\mathbf{k} \cdot \mathbf{E}) \mathbf{k} d\mathbf{k} &= \int_{-A}^A dk_x \int_{-B}^B dk_y \int_{-C}^C dk_z F(\mathbf{k}) (E_z k_z) (\mathbf{e}_x k_x + \mathbf{e}_y k_y + \mathbf{e}_z k_z) = \\ &= \int_{-A}^A dk_x \int_{-B}^B dk_y \int_{-C}^C dk_z F(\mathbf{k}) (E_z k_z) (\mathbf{e}_z k_z) = \mathbf{E} \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta \int_0^{\infty} k^2 dk F(\mathbf{k}) k^2 \cos^2\theta, \end{aligned} \quad (9.81)$$

де  $F(\mathbf{k}) = \tau_e \exp(-E/k_B T)$  – парна функція хвильового вектора  $\mathbf{k}$ . Тепер врахуємо, що інтеграл

$$\int_0^{\pi} \sin\theta \cos^2\theta d\theta = \frac{2}{3}.$$

Після інтегрування за кутами отримуємо  $4\pi/3$ , і як результат маємо

$$\mathbf{J}_n = \frac{e^2 \hbar^2 n \mathbf{E}}{4\pi^3 (m_n^*)^2 k_B T N_c} \frac{4\pi}{3} \int_0^{\infty} \tau_e e^{-\frac{E}{k_B T}} k^4 dk. \quad (9.82)$$

Для подальшого обчислення інтеграла у (9.82) зазначимо, що

$$E/k_B T = (\hbar^2/2m_n^* k_B T) k^2. \quad (9.83)$$

Зручно ввести позначення  $E/k_B T = x$ ,  $\beta = (\hbar^2/2m_n^* k_B T)^{-1}$ . Тоді (9.83) можна записати у вигляді рівняння  $k^2 = \beta x$ , звідки  $k^4 dk = (1/2)\beta^{5/2} x^{3/2} dx$ . До того ж врахуємо, що  $N_c = 2(2\pi \cdot m_n^* k_B T/\hbar^2)^{3/2}$ . Тоді із (9.82) отримуємо

$$\mathbf{J}_n = \frac{e^2 n \mathbf{E}}{m_n^*} \frac{4}{3\sqrt{\pi}} \int_0^{\infty} \tau_e e^{-x} x^{3/2} dx. \quad (9.84)$$

З елементарного курсу електрики відомо, що електричний струм виражається через заряд, концентрацію електронів та їхню середню швидкість направленного руху  $\langle \mathbf{v} \rangle$  за допомогою формули  $\mathbf{J} = en \langle \mathbf{v} \rangle$ , порівнювання якої із (9.84) показує, що

$$\langle \mathbf{v} \rangle = \frac{e \cdot 4}{m_n^{*2} 3\sqrt{\pi}} \int_0^{\infty} \tau_e e^{-x} x^{3/2} dx \cdot \mathbf{E} = \mu \mathbf{E}.$$

Рухливість  $\mu$  має значення швидкості електронів, набутої ними в одиничному електричному полі. Виходячи із наведених міркувань, введемо тепер стандартні позначення:

- ♦ **для середнього часу релаксації електронів**

$$\langle \tau_e \rangle = \frac{4}{3\sqrt{\pi}} \int_0^{\infty} \tau_e e^{-x} x^{3/2} dx; \quad (9.85)$$

- ♦ **для дрейфової рухливості електронів**

$$\mu_n = \frac{e \langle \tau_e \rangle}{m_n^*}; \quad (9.86)$$

- ♦ **для питомої провідності, обумовленої електронами**

$$\sigma_n = en\mu_n. \quad (9.87)$$

Тоді електронною складовою струму є

$$\mathbf{J}_n = en\mu_n \mathbf{E} = \sigma_n \mathbf{E}. \quad (9.88)$$

Аналогічно для діркової складової струму можна записати

$$\mathbf{J}_p = ep\mu_p \mathbf{E} = \sigma_p \mathbf{E}, \quad (9.89)$$

де:

- ♦ **середній час релаксації дірок**

$$\langle \tau_p \rangle = \frac{4}{3\sqrt{\pi}} \int_0^{\infty} \tau_p e^{-x} x^{3/2} dx, \quad (9.90)$$

- ♦ **дрейфова рухливість**

$$\mu_p = \frac{e \langle \tau_p \rangle}{m_p^*}, \quad (9.91)$$

- ♦ **питома провідність, обумовлена рухом дірок,**

$$\sigma_p = ep\mu_p. \quad (9.92)$$

Остаточно густину струму невідродженого напівпровідника, що має два типи носіїв заряду, запишемо у вигляді

$$\mathbf{J} = (\sigma_n + \sigma_p) \mathbf{E}. \quad (9.93)$$

Зауважимо, що завдяки використанню найпростішої моделі сферичних ізоенергетичних поверхонь ми отримали струм, направлений уздовж електричного поля. У загальнішому випадку анізотропних ізоенергетичних поверхонь зв'язок між густиною струму та полем є

тензорним. Зважаючи на (9.86) та (9.91) бачимо, що електропровідність напівпровідника визначається середнім часом релаксації і концентрацією носіїв заряду. Середній час релаксації визначається процесами розсіювання носіїв на неоднорідностях ґратки. Таким чином, перед аналізом поведінки провідності напівпровідників як функції температури можна стверджувати, що всі механізми розсіювання носіїв даватимуть внесок у провідність. З'ясуємо, які механізми та за яких умов дають основні внески до формування провідності. Вважатимемо, що всі механізми розсіювання формують функцію розподілу носіїв і діють незалежно один від одного. Тоді можна стверджувати, що повна ймовірність розсіювання дорівнює сумі ймовірностей за кожним каналом розсіювання (на атомах домішок, на іонізованих домішках, коливаннях ґратки тощо) і записати для часу релаксації

$$\tau^{-1} = \frac{1}{4\pi^2} \int_{\Omega} d\mathbf{k}' \sum_i w_i(\mathbf{k}, \mathbf{k}') (1 - \cos\theta) = \sum_i \tau_i^{-1}. \quad (9.94)$$

Звідси маємо

$$\tau = \left( \sum_i \tau_i^{-1} \right)^{-1}. \quad (9.95)$$

Визначимо тепер рухливість носіїв

$$\mu = \frac{e \langle \tau(E) \rangle}{m^*} = \frac{e}{m^*} \left\langle \left( \sum_i \tau_i^{-1}(E) \right)^{-1} \right\rangle. \quad (9.96)$$

Формула показує, що повний час релаксації в основному визначається парціальним процесом, який має менший час релаксації. Дійсно, припустимо, що всі часи релаксації  $\tau_i$  однакові. Тоді, якщо розсіювання визначається  $s$  незалежними каналами, то

$$\sum_i \tau_i^{-1} = \frac{s}{\tau_i}. \quad (9.97)$$

Підставляючи (9.97) до (9.96), отримаємо

$$\mu = \frac{e \langle \tau_i(E) \rangle}{sm^*}, \quad (9.98)$$

тобто якщо функція розподілу носіїв формується  $s$  каналами розсіювання з однаковим часом релаксації, то рухливість носіїв буде в  $s$  разів меншою від парціальної рухливості. Уявимо, що деякий механізм розсіювання характеризується часом релаксації  $\tau_m$ , значно меншим за часи релаксації інших механізмів, яких для визначеності нехай буде два – із часами релаксації  $\tau_1$  та  $\tau_2$ , відповідно. Причому  $\tau_1 \sim \tau_2$ ,  $\tau_m \ll \tau_1, \tau_2$ . Тоді із (9.95) отримуємо, що час релаксації буде обмежений зверху

$$\tau = \tau_m \left( 1 - \tau_m/\tau_1 - \tau_m/\tau_2 \right) < \tau_m. \quad (9.99)$$

Іншими словами, повний час релаксації визначається парціальним процесом із найменшим часом релаксації. Виходячи з цього, можна стверджувати, що за високих температур основним механізмом розсіювання є розсіювання на теплових коливаннях ґратки, для якого  $\tau \sim E^{-1/2}$ . За зменшення температури роль теплових коливань зменшується, а роль розсіювання на іонах домішки – підвищується, і поведінка часу релаксації визначається співвідношенням  $\tau \sim E^{3/2}$ .

## ■ 9.4. Залежність рухливості напівпровідника від температури

Як видно з отриманих результатів, залежність провідності від температури визначається двома чинниками – залежністю від температури концентрації носіїв і залежністю від температури середнього часу релаксації. Температурна залежність  $n$  та  $p$  розглядалась у попередньому розділі. Тепер же розглянемо залежність рухливості носіїв напівпровідника від  $T$ . Оскільки час релаксації суттєво визначається механізмом розсіювання, розглянемо ці головні механізми окремо.

### 9.4.1. Розсіювання на коливаннях ґратки

Як нам уже відомо (див. (9.44)), за розсіювання носіїв на довгохвильових акустичних фононах час релаксації

$$\tau_1 = \frac{\tau_{0l}}{(m^*)^{3/2}} T^{-1} E^{-1/2}. \quad (9.100)$$

Тоді, використовуючи (9.85) і беручи до уваги, що  $x = E/kT$  для середнього часу релаксації, який визначає рухливість та електропровідність, маємо

$$\langle \tau_1 \rangle = \frac{4}{3\sqrt{\pi}} \frac{\tau_{0l}}{(m^*)^{3/2}} T^{-1} (kT)^{-1/2} \int_0^{\infty} e^{-x} x dx = \frac{4}{3\sqrt{\pi}} \frac{\tau_{0l}}{(m^*)^{3/2} k^{1/2}} T^{-3/2}. \quad (9.101)$$

Таким чином рухливість, що зумовлена розсіюванням носіїв на коливаннях ґратки, визначається формулою

$$\mu_l = \frac{4}{3\sqrt{\pi}} \frac{e\tau_{0l}}{k^{1/2} (m^*)^{5/2}} T^{-3/2}. \quad (9.102)$$

Вона зменшується зі зростанням температури і збільшується – зі зменшенням ефективної маси. Оскільки ефективна маса електрона зазвичай менша за ефективну масу дірок, то рухливість електронів більша за рухливість дірок.

### 9.4.2. Розсіювання на іонізованих домішках

Згадаємо, що за розсіювання на іонізованих домішках час релаксації залежить від енергії та ефективної маси як

$$\tau_I = \frac{\sqrt{2}\epsilon_r^2 (m^*)^{1/2} E^{3/2}}{\pi \cdot Z^2 e^4 N_I \ln\{1 + (\epsilon_r / Ze^2 N_I^{1/3})^2 E^2\}}. \quad (9.103)$$

При обчисленні середнього часу релаксації у цьому випадку необхідно обчислити інтеграл

$$I_\tau = (kT)^{3/2} \int_0^\infty dx x^3 e^{-x} \frac{1}{\ln(1 + \gamma^2 x^2)}, \quad \gamma = \frac{\epsilon_r kT}{Ze^2 N_I^{1/3}}. \quad (9.104)$$

Оскільки в межах інтегрування логарифм повільно змінюється, можна винести його за знак інтеграла, взявши у точці, де функція  $x^3 e^{-x}$  має максимум, тобто в точці  $x = 3$  або  $E = 3kT$ . У результаті, зважаючи на те, що  $\int_0^\infty dx x^3 e^{-x} = 6$ , отримаємо

$$\langle \tau_I \rangle = \frac{8\sqrt{2}\epsilon_r^2 k^{3/2} (m^*)^{1/2} T^{3/2}}{\pi^{3/2} Z^2 e^4 N_I \ln[1 + (3\epsilon_r kT / Ze^2 N_I^{1/3})^2]}. \quad (9.105)$$

Таким чином, за розсіювання носіїв на іонізованих домішках рухливість визначається формулою

$$\mu_I = \frac{8\sqrt{2}\epsilon_r^2 k^{3/2} T^{3/2}}{\pi^{3/2} Z^2 e^3 N_I (m^*)^{1/2} \ln[1 + (3\epsilon_r kT / Ze^2 N_I^{1/3})^2]}. \quad (9.106)$$

За не дуже низьких температур поведінку рухливості в основному визначає величина  $T^{3/2}$ . Таким чином, можна стверджувати, що

$$\mu_I \sim T^{3/2}, \quad (9.107)$$

тобто рухливість носіїв, що обумовлена розсіюванням на іонізованих домішках, зменшується за зменшення температури.

### 9.4.3. Розсіювання на нейтральних атомах домішок

Оскільки час релаксації за розсіювання на нейтральних атомах домішок не залежить ні від температури, ні від енергії, то частина рухливості носіїв, що визначається такими процесами розсіювання, також не залежатиме ні від  $T$ , ні від  $E$  та визначатиметься формулою

$$\mu_A = \frac{e^2 m^*}{20\epsilon_r \hbar^3 N_A}. \quad (9.108)$$



Зрозуміло, що цей компонент рухливості має сенс лише за низьких температур, коли основна кількість домішок не іонізована. Для кремнію та германію за не дуже низьких температур ( $T > 10$  К) внеском до рухливості від процесів розсіювання на нейтральних домішках можна знехтувати.

#### 9.4.4. Розсіювання на дислокаціях

Оскільки час релаксації за розсіювання носіїв на дислокаціях обернено пропорційний до швидкості носіїв, що розсіюються на дислокації і визначаються формулою  $\tau_D = (3/8RN_D)v^{-1}$ , а швидкість визначається енергією частинки, яка в свою чергу пов'язується із температурою

$$v^{-1} = \frac{(m^*)^{1/2}}{\sqrt{2}} E^{-1/2} = \frac{(m^*)^{1/2} k_B^{-1/2} T^{-1/2}}{\sqrt{2}} x^{-1/2},$$

то для середнього часу релаксації для цього випадку маємо

$$\langle \tau_D \rangle = \frac{4}{3\sqrt{2}\pi} \frac{3(m^*)^{1/2}}{8RN_D k_B^{1/2}} T^{-1/2} \int_0^\infty e^{-x} x dx.$$

Звідси рухливість носіїв за їхнього розсіювання на дислокаціях визначається формулою

$$\mu_D = \frac{e}{\sqrt{8\pi}} \frac{(m^*)^{-1/2}}{RN_D k_B^{1/2}} T^{-1/2}. \quad (9.109)$$

Тобто із підвищенням температури рухливість носіїв, що визначається розсіюванням на дислокаціях, зменшується, а за зменшення ефективної маси носія – збільшується.

#### 9.4.5. Залежність рухливості та електропровідності носіїв заряду в реальних напівпровідниках від температури

У реальних напівпровідниках розсіювання носіїв відбувається одночасно за багатьма механізмами і, строго кажучи, вони не є незалежними. Однак для оцінки залежності рухливості від температури можна припустити, що всі механізми розсіювання незалежні. У такому разі для оцінок можна застосовувати залежності рухливості від температури за різних механізмів розсіювання. Внесок різних механізмів релаксації по-різному залежить від температури, і зі зміною останньої роль одних механізмів зменшується, інших – підсилюється. За дуже низьких температур домінуючим механізмом розсіювання є розсію-

вання на нейтральних домішках і дислокаціях. Зі зростанням температури починає домінувати розсіювання на іонах домішки. Подальше зростання температури приводить до домінування розсіювання на фононах. Таким чином можна вважати, що комбінована рухливість за не дуже низьких температур визначається розсіюванням на іонах домішки та фононах, тобто

$$\mu^{-1} = \mu_I^{-1} + \mu_T^{-1} \tag{9.110}$$

або, враховуючи залежність парціальної провідності від температури, маємо

$$\mu^{-1} = aT^{-3/2} + bT^{3/2}. \tag{9.111}$$

Рухливість напівпровідника немонотонно залежить від температури. Цю складну залежність можна пояснити тим, що в різних температурних інтервалах до формування рухливості роблять внесок різні механізми розсіювання носіїв.

На рис. 9.1 подано характерну поведінку рухливості як функції температури у *p*-антимоніді індію із концентрацією електронів  $n = 10^{14} \text{ см}^{-3}$  та концентрацією дірок  $p = 4 \cdot 10^{15} \text{ см}^{-3}$ .

У табл. 9.1 подано чисельні значення рухливості деяких напівпровідників. Рухливість змінюється в широкому інтервалі значень, що визначається зонною структурою (власних) напівпровідників і значною залежністю від температури.

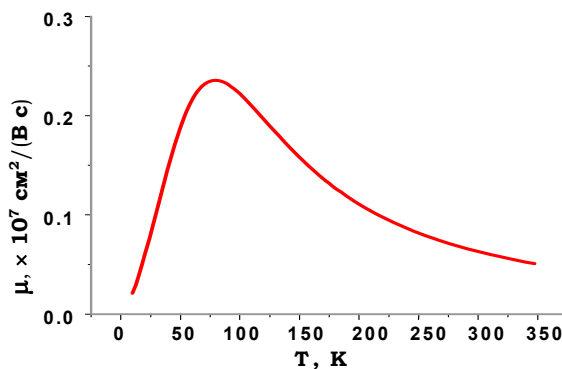


Рис. 9.1. Залежність рухливості від температури в InSb

Таблиця 9.1

	$\mu_n, \text{см}^2/(\text{В}\cdot\text{с})$		$\mu_p, \text{см}^2/(\text{В}\cdot\text{с})$		Показник ступеня $T^a$ (теор. $a=1,5$ )	
	300 К	77 К	300 К	77 К	<i>n</i>	<i>p</i>
Ge	3900	37100	1900	43700	1,66	2,33
Si	1500	45500	600	11600	2,5	2,7
InSb	78000	$1,2 \times 10^6$	750	10000	1,6	2,1
GaAs	8500	21000	420	4200	1,0	2,1

Оскільки електрони та дірки мають заряди протилежного знаку, то повна електропровідність є сумою парціальної провідності електронів і дірок. Питому провідність напівпровідника можна записати у вигляді

$$\sigma = e(p\mu_p + n\mu_n) = e\mu_p(p + bn), \tag{9.112}$$

де  $b = \mu_n/\mu_p$ . Для власного напівпровідника, де  $n = p = n_i$  ця формула переходить у наступну

$$\sigma = en_i\mu_p(1+b). \quad (9.113)$$

Із врахуванням залежності рухливості та концентрації носіїв від температури,  $\sigma$  можна записати у вигляді

$$\sigma = \text{const} \cdot T^{3/2+\gamma} e^{-\frac{E_g}{2kT}}. \quad (9.114)$$

Таким чином, якщо подавати залежність електропровідності від оберненої температури в логарифмічному масштабі (рис. 9.2), то видно, що в області власної провідності ця залежність майже лінійна. Зі зниженням температури провідність переходить в область, де концентрація носіїв визначається концентрацією іонізованих домішок і майже не змінюється з температурою (область виснаження див. на рис. 8.12). Тут температурна залежність провідності визначається в основному температурною залежністю рухливості й може бути різною залежно (як ми бачили раніше) від механізму розсіювання. За подальшого зменшення температури відбувається "вимержання" носіїв, де провідність знов визначається експоненціальною залежністю числа іонізованих домішок. Ми бачимо, що на відміну від металів провідність напівпровідників принципово залежить від температури, оскільки від останньої залежить концентрація носіїв заряду. І тільки в області виснаження домішок, де концентрація носіїв майже не залежить від температури, поведінка провідності визначається механізмом розсіювання носіїв і має ступеневий характер.

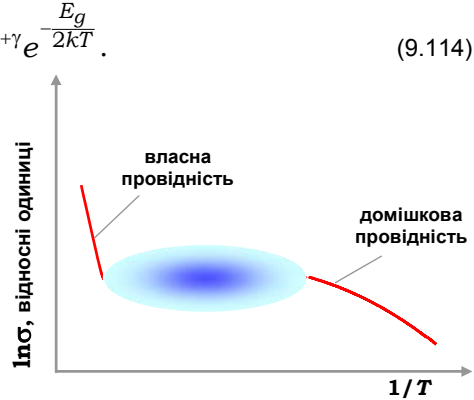


Рис. 9.2. Схематична залежність електропровідності від оберненої температури

## ■ 9.5. Гальваномагнітні явища в напівпровідниках

Кінетичні явища, що відбуваються в напівпровідниках за одночасної дії електричного та магнітного полів, називаються гальваномагнітними. Основною їх причиною є викривлення траєкторії носіїв заряду в магнітному полі під дією сили Лоренца. Траєкторії носіїв заряду у напівпровідниках під дією магнітного поля та вільного носія в магнітному полі – кругової спіралі, що навивається на магнітну силову лінію –

можуть суттєво відрізнитись. Різноманітність траєкторій носіїв заряду у різних матеріалах є основною причиною розмаїття гальваномагнітних явищ. Найбільш ефективно гальваномагнітні ефекти виявляються в геометрії, коли магнітне та електричне поля перпендикулярні.

До гальваномагнітних ефектів належать, зокрема:

- ♦ **ефект Холла** – виникнення різниці потенціалів за рахунок відхилення носіїв заряду до бічних граней зразка у магнітному полі, перпендикулярному до напрямку електричного струму;

- ♦ **магніторезистивний ефект** – зміна швидкості дрейфу носіїв заряду внаслідок дії магнітного поля;

- ♦ **ефект Етінгсхаузена** – виникнення поперечної різниці температур у магнітному полі, перпендикулярному до напрямку електричного поля;

- ♦ **ефект Нернста** – виникнення повздовжньої різниці температур у магнітному полі, перпендикулярному до напрямку електричного поля.

### 9.5.1. Ефект Холла

Розглянемо деякі кінетичні ефекти, що спостерігаються за одночасної дії на напівпровідник електричного та магнітного полів. Єдиним гальваномагнітним ефектом, що може спостерігатись у випадку, коли час релаксації не залежить від енергії, є Ефект Холла. Розглянемо його для слабкого магнітного поля. Спочатку з'ясуємо, що розумітимемо під терміном **слабке поле**. Вважатимемо поле слабким, якщо періоди обертання  $T_C$  за круговою орбітою в магнітному полі є набагато більшими за час релаксації  $\tau$ , тобто

$$\tau \ll T_C. \quad (9.115)$$

Частотою обертання носія в магнітному полі  $\mathbf{B}$  є циклотронна частота

$$\omega_c = \frac{2\pi}{T_C} = \frac{e}{m^*} B. \quad (9.116)$$

Тоді можна записати  $\frac{\tau}{T_C} = \frac{e\tau}{2\pi \cdot m^*} B = \frac{\mu B}{2\pi} \ll 1$ , (9.117)

тобто умова слабкого поля  $\frac{\mu B}{2\pi} \ll 1$  (9.118)

визначається не тільки величиною індукції поля, а й рухливістю носіїв. Якщо ввести довжину вільного пробігу  $l = v\tau$  і радіус циклотронної орбіти  $r = v/\omega_c$ , то

$$l/r = \tau\omega_c. \quad (9.119)$$

Із (9.119) та (9.117) знайдемо

$$\frac{\tau\omega_c}{2\pi} = \frac{\mu B}{2\pi} = \frac{1}{2\pi \cdot r} \ll 1, \quad (9.120)$$

тобто поле є слабким, якщо носії, рухаючись круговою орбітою у площині, перпендикулярній  $\mathbf{B}$ , встигають пройти до зіткнення малу частину циклотронної орбіти. Припустимо, що у зразку напівпровідника, що має форму прямокутної пластинки (рис. 9.3), під дією електричного поля  $\mathbf{E} = (E, 0, 0)$  протікає струм із густиною

$$\mathbf{J} = -en\mathbf{v} = \sigma\mathbf{E}. \quad (9.121)$$

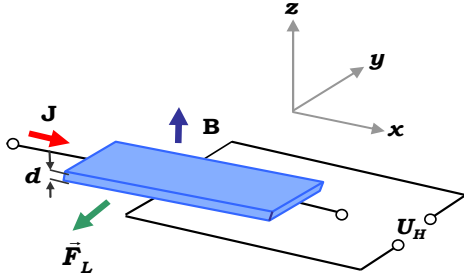


Рис. 9.3. Схема вимірювання ефекту Холла

Вважатимемо, що час релаксації сталий, тобто не залежить від енергії, а це можливо, якщо розсіювання носіїв відбувається в основному на нейтральних домішках. Нехай на систему діє магнітне поле  $\mathbf{B} = (0, 0, B)$ . Тоді на носій заряду, що рухається із дрейфовою швидкістю  $\mathbf{v}$ , діятиме сила Лоренца

$$\mathbf{F} = \pm e(\mathbf{v} \times \mathbf{B}), \quad (9.122)$$

перпендикулярна до  $\mathbf{v}$  та  $\mathbf{B}$  (тут знаки + та – відповідають дірці та електрону, відповідно). Оскільки швидкість дрейфового руху носіїв та електричне поле пов'язані співвідношенням

$$\mathbf{v} = \mu\mathbf{E} = \pm \frac{e \langle \tau \rangle}{m^*} \mathbf{E}, \quad (9.123)$$

то (9.122) переписеться у вигляді

$$\mathbf{F} = \frac{e^2 \langle \tau \rangle}{m^*} (\mathbf{E} \times \mathbf{B}). \quad (9.124)$$

Ця формула показує, що напрямок сили Лоренца не залежить від знаку заряду, а визначається напрямком векторів  $\mathbf{E}$  та  $\mathbf{B}$  ( $\mathbf{J}$  та  $\mathbf{B}$ ). Таким чином,

**Якщо швидкість носіїв визначається зовнішнім електричним полем, то електрони та дірки під дією сили Лоренца відхиляються в один і той самий бік.**

Для вибраної геометрії полів сила Лоренца спрямована проти вісі ОУ. Під дією цієї сили електрони в донорному напівпровіднику та дірки в акцепторному – відтіснятимуться до краю ( $-Y$ ) пластини. У результаті на протилежному боці ( $+Y$ ) виникне їхній дефіцит, що зумовить протилежний за знаком заряд порівняно із зарядом на боці ( $-Y$ ) пластини.

Унаслідок такого розділення зарядів уздовж вісі – ОУ з'явиться електричне поле  $\mathbf{E}_H$ , напрямком якого залежить від знаку носіїв заряду. В обраній геометрії це поле спрямовано вздовж вісі ОУ для зразка із провідністю  $p$ -типу, і проти вісі ОУ – у зразках із провідністю  $n$ -типу. Виникнення такого електричного поля і називається ефектом Холла. Напруженість поля  $\mathbf{E}_H$  зростатиме, доки сила, що обумовлена цим полем, не компенсує силу Лоренца, тобто доки не виконуватиметься рівність

$$e\mathbf{E}_H + \mathbf{F} = 0. \tag{9.125}$$

Із цієї формули випливає, що  $\mathbf{E}_H = -(1/e) \cdot \mathbf{F} = -(\mathbf{v} \times \mathbf{B})$ . У геометрії, що розглядається, в стаціонарному стані за розірваних контактів  $U$  маємо

$$-eE_H = evB. \tag{9.126}$$

У цьому випадку носії рухаються вздовж зразка під дією тільки по-вздовжнього електричного поля  $\mathbf{E}$ , тобто напрямком густини струму  $\mathbf{J}$  збігається із напрямком поля  $\mathbf{E}$  (рис. 9.4 – магнітне поле нормальне до площині рисунка і спрямовано на спостерігача). Вектор напруженості сумарного електричного поля  $\mathbf{E}' = \mathbf{E} + \mathbf{E}_H$  щодо напрямку струму  $\mathbf{J}$  тепер повернутий на деякий кут, який називається кутом Холла (причому, якщо  $|\mathbf{E}_H| = vB$  та  $v = \mu|\mathbf{E}|$ , то  $\text{tg}\varphi = |\mathbf{E}_H|/|\mathbf{E}| = \mu B$ ). А оскільки тепер еквіпотенціальні поверхні повернуті на кут  $\varphi$  щодо вісі ОУ, то вздовж неї з'явиться різниця потенціалів – електрорушійна сила Холла. Припустимо, що вектори  $\mathbf{E}_H$ ,  $\mathbf{J} = \sigma \mathbf{E}$  та  $\mathbf{B}$  пов'язані між собою деяким співвідношенням із невідомим коефіцієнтом  $R$

$$\mathbf{E}_H = R\sigma(\mathbf{E} \times \mathbf{B}). \tag{9.127}$$

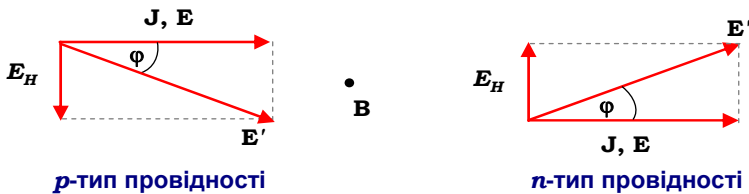


Рис. 9.4. Визначення кута Холла

З іншого боку (9.125) дає

$$\mathbf{E}_H = -(\mathbf{v} \times \mathbf{B}) = -\mu \cdot (\mathbf{E} \times \mathbf{B}). \tag{9.128}$$

Порівняння цих виразів дає для напівпровідника:

♦ з електронним типом провідності

$$R = -\frac{\mu}{\sigma} = -\frac{1}{en}; \tag{9.129}$$

♦ із дірковим типом провідності

$$R = \frac{1}{ep}. \quad (9.130)$$

Таким чином, коефіцієнт Холла обернено пропорційний концентрації носіїв заряду, а його знак збігається зі знаком носіїв. Якщо ширина зразка, на якому досліджується ефект Холла, дорівнює  $b$ , то повний струм через зразок можна записати як  $I = Jbd$ . Напругу Холла можна записати у вигляді  $V_H = E_H b$  або з використанням (9.127)  $V_H = RBI(b/bd)RBI$ . Таким чином, при вимірюванні  $V_H$  можна визначити знак і концентрацію носіїв заряду за певної температури.

Враховуючи, що струм через зразок визначається концентрацією носіїв заряду та їхньою рухливістю  $J = en\mu E$  і вимірюючи в одному експерименті як струм, так і напругу Холла, можна визначити рухливість носіїв заряду. Досліджуючи одночасно електропровідність та ефект Холла залежно від температури, можна знайти:

- ♦ залежність концентрації носіїв від температури;
- ♦ енергію активації домішок;
- ♦ залежність рухливості носіїв від температури;
- ♦ ширину забороненої зони,

отже визначити основні характеристики напівпровідника, у тому числі й механізми розсіювання носіїв заряду.

### **9.5.2. Ефект Холла в напівпровідниках із двома типами носіїв заряду**

Припустимо тепер, що струм в напівпровіднику формується за рахунок руху електронів і дірок, які мають концентрації та рухливості  $n$ ,  $\mu_n$  та  $p$ ,  $\mu_p$ , відповідно. Вважатимемо магнітне поле, спрямоване вздовж вісі  $OZ$  системи координат, слабким. При цьому вектор електричного поля  $\mathbf{E}$  та вектор струму  $\mathbf{J}$  спрямовані вздовж вісі  $OX$  системи координат. Крім того, не станемо накладати обмежень на тип механізмів розсіювання носіїв, тобто вважатимемо, що час релаксації може залежати від енергії. Під дією електричного поля носії заряду прискорюватимуться, і за збільшенні дрейфової швидкості на носії діятиме сила Лоренца, що спричиняє відхилення траєкторії носіїв заряду. Траєкторії носіїв заряду являтимуть собою частини дуг між кожними двома зіткненнями носіїв із дефектами ґратки. Оскільки носії при зіткненні із дефектами передають енергію, придбану ними в полі, то рух

носіїв буде в середньому рівномірним, і в результаті густина струму дірок відхилиться на кут  $\varphi_p$  щодо напрямку поля (рис. 9.5). Електрони під дією поля відхилитимуться у тому самому напрямку, що й дірки, а густина електронного струму – у протилежному напрямку на кут  $\varphi_n$ . У результаті сумарна густина струму

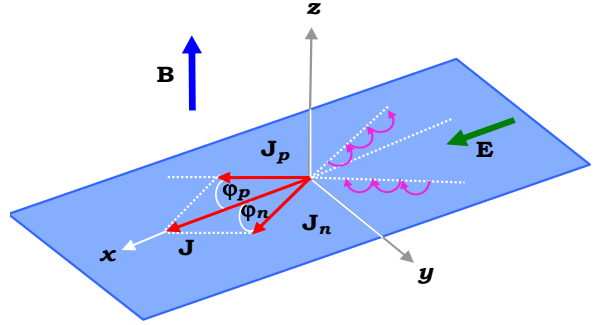


Рис. 9.5. Ефект Холла в напівпровіднику із двома типами носіїв

$$\mathbf{J} = \frac{e\hbar}{4\pi^3 m_n^*} \int_{\Omega} \frac{\partial f_0}{\partial E} (\chi_n \cdot \mathbf{k}) \mathbf{k} d\mathbf{k} - \frac{e\hbar}{4\pi^3 m_p^*} \int_{\Omega} \frac{\partial f_{0,p}}{\partial E} (\chi_p \cdot \mathbf{k}) \mathbf{k} d\mathbf{k}. \quad (9.131)$$

Подальші обчислення проводитимемо для однорідного напівпровідника зі сферичними ізоенергетичними поверхнями за відсутності градієнта температури. Це означає, що

$$\nabla f = \frac{E - E_F}{T} \nabla T + \nabla E_F = 0. \quad (9.132)$$

Використовуючи формули (9.40)–(9.41), запишемо для однорідного напівпровідника в слабкому магнітному полі ( $\zeta_{n,p}^2 \beta^2 \ll \zeta_{n,p} \beta \ll 1$ ,  $\beta = (e/\hbar)\mathbf{B}$ )

$$\chi_n = -\zeta_n e \{ \mathbf{E} + \xi_n e \boldsymbol{\beta} \times \mathbf{E} \}, \quad (9.133)$$

$$\chi_p = \zeta_p e \{ \mathbf{E} - \xi_p e \boldsymbol{\beta} \times \mathbf{E} \}. \quad (9.134)$$

Після підстановки (9.133) до (9.131) запишемо густина електронного струму

$$\mathbf{J}_n = \frac{e\hbar}{4\pi^3 m_n^*} \int_{\Omega} \frac{\partial f_0}{\partial E} (\chi_n \cdot \mathbf{k}) \mathbf{k} d\mathbf{k} = -\frac{e^2 \hbar^2}{4\pi^3 (m_n^*)^2} \int_{\Omega} \frac{\partial f_0}{\partial E} \left[ \tau_e \mathbf{E} \cdot \mathbf{k} + \frac{e\tau_n^2}{m_n^*} (\mathbf{B} \times \mathbf{E}) \cdot \mathbf{k} \right] \mathbf{k} d\mathbf{k}. \quad (9.135)$$

Зважаючи на те, що  $\partial f_0 / \partial E = -(n/k_B T N_c) e^{-E/k_B T}$  і приймаючи напрямок поля  $\mathbf{E}$  за вісь OZ, запишемо перший інтеграл в (9.135) у вигляді (див. вивід формули (9.82))

$$\int_{\Omega} \frac{\partial f_0}{\partial E} \tau_e (\mathbf{E} \cdot \mathbf{k}) \mathbf{k} d\mathbf{k} = \mathbf{E} \frac{4\pi}{3} \int_0^{\infty} \frac{\partial f_0}{\partial E} \tau_e k^4 dk. \quad (9.136)$$

Аналогічно вибираємо напрямок вектора  $\mathbf{B} \times \mathbf{E}$  за вісь OZ та отримуємо для другого інтеграла в (9.135)

$$\int_{\Omega} F(E) ((\mathbf{B} \times \mathbf{E}) \cdot \mathbf{k}) \mathbf{k} d\mathbf{k} = \mathbf{B} \times \mathbf{E} \frac{4\pi}{3} \int_0^{\infty} F(E) k^4 dk. \quad (9.137)$$



Перейдемо тепер до змінної  $t = E/kT$ . Беручи до уваги (9.136)–(9.137), а також те, що  $k^4 dk = ((2m_n^* kT)^{5/2}/2\hbar^5) t^{3/2} dt$ , отримаємо для електронної частини струму

$$\mathbf{J}_n = \frac{e^2 n}{m_n^*} \mathbf{E} \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-t} t^{3/2} dt + \frac{e^3 n}{(m_n^*)^2} (\mathbf{B} \times \mathbf{E}) \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e^2 e^{-t} t^{3/2} dt. \quad (9.138)$$

Визначимо середній час релаксації електронів формулою

$$\langle \tau_e \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e e^{-t} t^{3/2} dt, \quad (9.139)$$

а величину середнього квадрата часу релаксації електронів як

$$\langle \tau_e^2 \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e^2 e^{-t} t^{3/2} dt. \quad (9.140)$$

Врахуємо тепер, що рухливність пов'язана із середнім часом релаксації формулою  $\mu_n = e \langle \tau_e \rangle / m_n^*$ . Використовуючи визначення (9.139)–(9.140), маємо для електронної компоненти струму

$$\mathbf{J}_n = en\mu_n \mathbf{E} + r_n en\mu_n^2 (\mathbf{B} \times \mathbf{E}), \quad (9.141)$$

де введено позначення

$$r_n = \frac{\langle \tau_n^2 \rangle}{|\langle \tau_n \rangle|^2}. \quad (9.142)$$

Аналогічно для густини діркової складової струму маємо

$$\mathbf{J}_p = ep\mu_p \mathbf{E} - r_p ep\mu_p^2 (\mathbf{B} \times \mathbf{E}) \quad (9.143)$$

де

$$r_p = \frac{\langle \tau_p^2 \rangle}{|\langle \tau_p \rangle|^2}. \quad (9.144)$$

У результаті запишемо повну густину струму

$$\mathbf{J} = e(n\mu_n + p\mu_p) \mathbf{E} - r_n e \left( \frac{r_p}{r_n} p\mu_p^2 - n\mu_n^2 \right) (\mathbf{B} \times \mathbf{E}). \quad (9.145)$$

У цій формулі коефіцієнти  $r_n$  та  $r_p$  залежать від механізмів розсіювання носіїв. У випадку пружного розсіювання час релаксації як функцію енергії можна представити у вигляді

$$\tau_j = \tau_{0j} \left( \frac{E}{kT} \right)^\alpha. \quad (9.146)$$

Тоді середній час релаксації і середній квадрат часу релаксації запишуться у вигляді ( $j = n, p$ )

$$\langle \tau_j \rangle = \frac{4(\tau_{0j})}{3\sqrt{\pi}} \int_0^{\infty} e^{-t} t^{3/2+\alpha} dt \quad \text{та} \quad \langle \tau_j^2 \rangle = \frac{4(\tau_{0j})^2}{3\sqrt{\pi}} \int_0^{\infty} e^{-t} t^{3/2+2\alpha} dt. \quad (9.147)$$

Для обчислення інтегралів у цих формулах згадаємо визначення гамма-функції  $\Gamma(\nu) = \int_0^{\infty} t^{\nu-1} e^{-t} dt$ . Таким чином із (9.147) маємо

$$r_j = \frac{\langle \tau_j^2 \rangle}{\langle \tau_j \rangle^2} = \frac{3\sqrt{\pi}}{4} \frac{\Gamma(2\alpha+5/2)}{\Gamma^2(\alpha+5/2)}.$$

Звідси отримуємо, що відношення  $r_p/r_n$  дорівнює одиниці. У такому випадку (9.146) переходить у (тут введено стандартне позначення для Холл-фактора  $r_H \equiv r_n$ )

$$\mathbf{J} = e(n\mu_n + p\mu_p)\mathbf{E} - r_H e(p\mu_p^2 - n\mu_n^2)(\mathbf{B} \times \mathbf{E}). \quad (9.148)$$

У розглянутій вище геометрії, де  $\mathbf{B} = (0, 0, B)$  і  $\mathbf{J} = (J, 0, 0)$ , із рівняння (9.146) маємо

$$J = e(n\mu_n + p\mu_p)E_x + r_H e(p\mu_p^2 - n\mu_n^2)BE_y, \quad (9.149)$$

$$0 = -r_H e(p\mu_p^2 - n\mu_n^2)BE_x + e(n\mu_n + p\mu_p)E_y. \quad (9.150)$$

Із (9.150) знаходимо для напруженості холлівського поля

$$E_y = RJB, \quad (9.151)$$

де коефіцієнт Холла

$$R = \frac{r_H}{e} \frac{p\mu_p^2 - n\mu_n^2}{(n\mu_n + p\mu_p)^2}. \quad (9.152)$$

У випадку власного напівпровідника, де  $n = p = n_i$ , із цієї формули маємо

$$R_i = \frac{r_H}{en_i} \frac{\mu_p - \mu_n}{\mu_n + \mu_p}. \quad (9.153)$$

Зазвичай рухливість електронів є більшою за рухливість дірок, тобто коефіцієнт Холла для власного напівпровідника  $R_i < 0$ . Із (9.153) видно також, що коефіцієнт Холла залежить від механізму розсіювання, який враховується холл-фактором  $r_H$ . Обчислимо величину холл-фактора для різних механізмів розсіювання носіїв у напівпровідниках. У випадку, коли основний внесок дає розсіювання на акустичних коливаннях ґратки,

$$\tau_l = \frac{\tau_{0l}}{(m_n^*)^{3/2} k^{1/2} T^{3/2}} t^{-1/2}. \quad (9.154)$$

Підставимо цей вираз до (9.139). Маємо

$$\langle \tau_e \rangle = \frac{4\tau_{0l}}{3\sqrt{\pi}(m_n^*)^{3/2} k^{1/2} T^{3/2}} \int_0^\infty e^{-t} t dt = \frac{4\tau_{0l}}{3\sqrt{\pi}(m_n^*)^{3/2} k^{1/2} T^{3/2}}. \quad (9.155)$$

Із (9.140) отримаємо

$$\langle \tau_e^2 \rangle = \frac{4(\tau_{0l})^2}{3\sqrt{\pi}[(m_n^*)^{3/2} k^{1/2} T^{3/2}]^2} \int_0^\infty e^{-t} t^{1/2} dt. \quad (9.156)$$

Інтеграл у (9.156)  $\Gamma(3/2) = \sqrt{\pi}/2$ . Таким чином,

$$\langle \tau_e^2 \rangle = \frac{4(\tau_{0l})^2}{3\sqrt{\pi}[(m_n^*)^{3/2} k^{1/2} T^{3/2}]^2} \frac{\sqrt{\pi}}{2}. \quad (9.157)$$

Із (9.155) та (9.157) отримуємо

$$r_n = \frac{\langle \tau_e^2 \rangle}{\langle \tau_e \rangle^2} = \frac{4}{3\sqrt{\pi}} \frac{\sqrt{\pi}}{2} \frac{(3\sqrt{\pi})^2}{4^2} = \frac{3\pi}{8}. \quad (9.158)$$

Якщо основний внесок дають процеси розсіювання на іонах домішки, то часом релаксації є

$$\tau_l = \tau_{0l} E^{3/2} = \tau_{0l} (kT)^{3/2} t^{3/2}. \quad (9.159)$$

Аналогічно до попереднього випадку обчислимо середній час релаксації за розсіювання електронів на іонах домішки

$$\langle \tau_e \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e^2 e^{-t} t^{3/2} dt = \frac{4}{3\sqrt{\pi}} \tau_{0l} (kT)^{3/2} \int_0^\infty \tau_e e^{-t} t^3 dt = \frac{4}{3\sqrt{\pi}} 6\tau_{0l} (kT)^{3/2}$$

і середній квадрат часу релаксації

$$\langle \tau_e^2 \rangle = \frac{4}{3\sqrt{\pi}} \int_0^\infty \tau_e^2 e^{-t} t^{3/2} dt = \frac{4}{3\sqrt{\pi}} [\tau_{0l} (kT)^{3/2}]^2 \int_0^\infty \tau_e e^{-t} t^{9/2} dt = \frac{4}{3\sqrt{\pi}} [\tau_{0l} (kT)^{3/2}]^2 \Gamma(11/2)$$

Оскільки  $\Gamma(n + \frac{1}{2}) = \frac{1 \cdot 3 \cdot 5 \cdot 7 \dots (2n-1)}{2^n} \Gamma(\frac{1}{2})$  і  $\Gamma(1/2) = \sqrt{\pi}$ , то  $\Gamma(11/2) = \frac{3 \cdot 5 \cdot 7 \cdot 9}{32} \sqrt{\pi}$ .

Тепер легко знаходимо для холл-фактора

$$r_H = \frac{4}{3\sqrt{\pi}} \frac{945\sqrt{\pi}}{32} \frac{(3\sqrt{\pi})^2}{4^2 6^2} = \frac{315}{512} \pi, \quad (9.160)$$

тобто за низьких температур, коли основний внесок до процесів розсіювання дає розсіювання на іонах домішки, за фактор Холла варто

брати  $r_H = 315\pi/512$ . У випадку більш високих температур, коли основну роль відіграє розсіювання на акустичних коливаннях ґратки, за фактор Холла необхідно приймати величину  $r_H = 3\pi/8$ . Для електронного напівпровідника, для якого  $n \gg p$ , із (9.152) маємо

$$R = -r_H \frac{1}{en}. \quad (9.161)$$

Зауважимо, що цей результат відрізняється від отриманого раніше саме через залежність часу релаксації від енергії. У випадку, коли час релаксації від енергії не залежить,  $r_H = 1$  і (9.161) переходить у формулу (9.129). Питома провідність такого напівпровідника визначається дрейфовою рухливістю  $\mu_d$  і дорівнює  $\sigma = en\mu_d$ . Тоді, використовуючи (9.161), маємо

$$\frac{1}{r_n} |R| \sigma = \mu_d. \quad (9.162)$$

Визначимо холлівську рухливість таким співвідношенням

$$|R| \sigma = \mu_H. \quad (9.163)$$

Тоді

$$\mu_H = r_n \mu_d. \quad (9.164)$$

У випадку, коли час релаксації не залежить від енергії,  $r_n = 1$  і холлівська та дрейфова рухливості збігаються. Подібне може спостерігатись у металів і вироджених напівпровідників. Оскільки тягнуче поле виражається через густину струмів і рухливості носіїв формулами  $E = j_n/en\mu_n$  та  $E = j_p/en\mu_p$ , а поле Холла – через густину струмів формулами  $E_H^{(p)} = (r_H/en\mu_p)j_p B$  і  $E_H^{(n)} = (r_H/en\mu_n)j_n B$ , то із геометричної побудови (див. рис. 9.4) ясно, що для малих кутів відхилення траєкторії носіїв виражається формулами

$$\varphi_p = r_H \mu_p B \quad \text{та} \quad \varphi_n = -r_H \mu_n B. \quad (9.165)$$

Тепер коротко розглянемо залежність коефіцієнта Холла від температури у зразках  $p$ -типу. Для цього проаналізуємо (9.152). Легко побачити, що при  $p = n(\mu_n/\mu_p)^2$  величина  $R_H$  змінює знак. Оскільки відношення рухливостей  $\mu_n/\mu_p$  зазвичай є великим, то у власному напівпровіднику коефіцієнт Холла виявляється від'ємним як у напівпровіднику  $n$ -типу. У напівпровіднику  $p$ -типу за низьких температур  $R_H > 0$ . За підвищення температури відбувається швидке зростання концентрації неосновних носіїв, і коефіцієнт Холла змінює знак на протилежний. У сильно компенсованих зразках  $p$ -типу також може спостерігатись зміна знаку коефіцієнта Холла з позитивного на від'ємний. Крім того, у напівпровідниках  $p$ -типу ця обставина приводить до холлівського викиду вище прямої, що відповідає власній провідності. Поведінку

коефіцієнта Холла як функції оберненої температури в антимоніді індію [4] подано на рис. 9.6. Оскільки провідність біполярного напівпровідника визначається формулою  $\sigma = e(n\mu_n + p\mu_p)$  (див. (9.145)), то питомий опір зразків  $p$ -типу матиме подібну до  $R_H(T)$  поведінку. Максимум питомого опору спостерігається не при  $p = n_i$ , а для значень концентрації основних носіїв  $p = \sqrt{\mu_n/\mu_p} n_i$ . Для InSb це в 8,94 рази більше, ніж  $n_i$ .

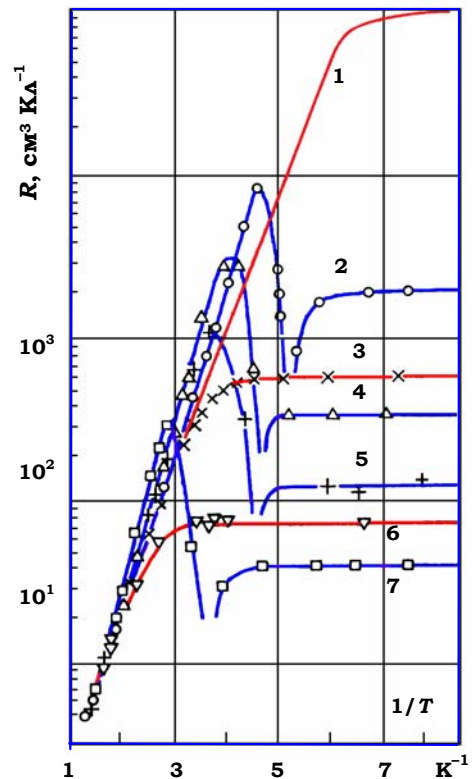
Характерними величинами ЕРС Холла за розмірів зразка порядку кількох міліметрів, концентрації носіїв  $\sim 10^{14} \text{ см}^{-3}$  і струму  $\sim 1 \text{ мА}$  в магнітному полі  $0,1 \text{ Т} - \epsilon 10^{-2} \text{ В}$ . Зауважимо, що в експериментальному дослідженні ефекту Холла може виникнути проблема, пов'язана з можливим неспіввісним розміщенням холлівських контактів. Дійсно, якщо незбіг контактів складає  $0,1 \text{ мм}$  (рис. 9.7), то різниця потенціалів на голкоподібному зразку, що визначається формулою

$$\Delta U = IR = I \frac{l}{se n \mu_H}$$

за струму  $1 \text{ мА}$  та поперечного перерізу  $1 \text{ мм}^2$ , складає величину  $\sim 10^{-2} \text{ В}$ , і має той самий порядок, що й ЕРС Холла.

Таким чином важливою експериментальною проблемою є компенсація різниці потенціалів  $\Delta U$ . Одним із можливих розв'язків цієї проблеми є конфігурація експерименту (рис. 9.8). Схема дозволяє компенсувати падіння повздовжньої напруги за вимкненого магнітного поля. При цьому потенціометр регулюється так, щоб при  $\mathbf{B} = 0$  напруга  $V$  була нульовою.

Ефект Холла є потужним засобом вивчення властивостей напівпровідників. Наприклад, визначення знаку коефіцієнта Холла дозволяє визначити тип провідності напівпровідника; якщо відомий коефіцієнт Холла, то можна визначити конче-



нтрацію носіїв заряду; ефект Холла широко використовується в приладобудуванні та техніці для вимірювання із великою точністю слабких магнітних полів; найдосконаліші системи регулювання запалення у двигунах внутрішнього згоряння базуються саме на системах визначення обертів двигуна за допомогою датчиків Холла.

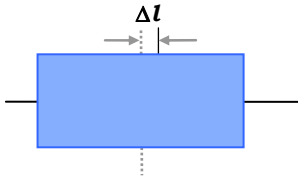


Рис. 9.7. Виникнення різниці потенціалів через незбіг положення бічних контактів при вимірюванні поля Холла

Рис. 9.6. Температурна залежність коефіцієнта Холла у зразках антимоніду індію  $n$ - та  $p$ -типів. Криві відповідають таким концентраціям донорів ( $n$ ) та акцепторів ( $p$ ):  
 1 –  $n = 4,3 \times 10^{13} \text{ см}^{-3}$ ; 2 –  $p = 4 \times 10^{15} \text{ см}^{-3}$ ;  
 3 –  $n = 1,3 \times 10^{16} \text{ см}^{-3}$ ; 4 –  $p = 2,2 \times 10^{16} \text{ см}^{-3}$ ;  
 5 –  $p = 6 \times 10^{16} \text{ см}^{-3}$ ; 6 –  $n = 10 \times 10^{17} \text{ см}^{-3}$ ;  
 7 –  $p = 2 \times 10^{17} \text{ см}^{-3}$

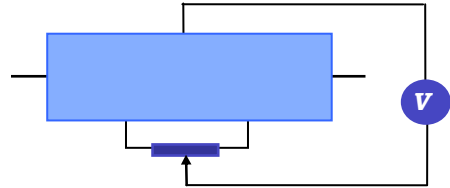


Рис. 9.8. Метод компенсування різниці потенціалів незбігу

### 9.5.3. Магніторезистивний ефект

Як ми бачили вище, рух носіїв напівпровідника, що знаходиться під дією магнітного поля, характеризується викривленням їхніх траєкторій, що спричинює зміни складової струму вздовж напрямку електричного поля, прикладеного до напівпровідника. Отже, можна стверджувати про зміну провідності напівпровідника у напрямку  $\mathbf{E}$ . Це явище називається магніторезистивним ефектом або ефектом Гаусса, його можна виміряти експериментально. Іноді кажуть, що під дією магнітного поля в напівпровіднику з'являється ефект магнітоопору, появу якого легко зрозуміти із таких простих міркувань. Завдяки дії сили Лоренца поперечне магнітне поле відхиляє носії заряду у напрямку, перпендикулярному дрейфу, що спричиняє ефективну зміну довжини вільного пробігу. Дослідимо на якісному рівні появу ефекту магнітоопору. Для цього розглянемо зміну довжини вільного пробігу носія заряду внаслідок відхилення від напрямку зовнішнього електричного поля (рис. 9.9). Рух електрона між двома перерізами показано штрихованими лініями. Нехай при вимкненому магнітному полі відстань між цими перерізами дорівнює середній довжині вільного пробігу  $l$  у напрямку  $\mathbf{E}$ . Під дією магнітного поля електрон рухатиметься викривленою траєкторією. Якщо це викривлення невелике, то припустимо, що електрон

рухається вздовж прямої, яка відхилена від  $OB$  на малий кут  $\varphi$ . Електрон, що пройшов відстань  $l$ , за руху вздовж траєкторії буде розсіюваний. Це означає, що ефективна довжина вільного пробігу вздовж напрямку поля  $\mathbf{E}$  зменшується на величину  $\Delta l = CB$ . Оскільки  $OA = OB = l$ , а кут  $\varphi$  є малим, з елементарної геометричної побудови знаходимо, що  $\Delta l = l \cdot \varphi^2$ . Після підстановки значень кутів Холла електронів і дірок (за нехтування відмінністю від одиниці Холл-фактора), отримуємо

$$\frac{\Delta l_p}{l_p} = \mu_p^2 B^2 \quad \text{та} \quad \frac{\Delta l_n}{l_n} = \mu_n^2 B^2. \quad (9.166)$$

Вважаючи, що зміни опору під дією магнітного поля відбуваються за рахунок зміни ефективної рухливості, яка викликана зменшенням ефективної (у напрямку дії електричного поля) довжини вільного пробігу, можна записати оцінку

$$\mu_{eff} = \frac{e}{m^*} \langle \tau \rangle = \frac{e}{m^*} \frac{l_{eff}}{v} \quad \text{або} \quad \mu_{eff} = \mu_0 - \Delta\mu, \quad \text{де} \quad \Delta\mu = \frac{e}{m^*} \frac{\Delta l}{v}.$$

Звідси маємо для провідності  $\sigma_{eff} = \sigma_0 - \Delta\sigma$ . Після введення параметра  $\alpha = \Delta\sigma/\sigma_0 = \Delta l/l$  знайдемо, що для  $\alpha \ll 1$   $\Delta R/R = \Delta l/l$ . Далі, використовуючи (9.166), визначимо, що наявність поперечного магнітного поля викликає збільшення електричного опору зразка пропорційно квадрату магнітного поля

$$\frac{\Delta R_n}{R_n} = \frac{\mu_n^2 B^2}{2} \quad \text{та} \quad \frac{\Delta R_p}{R_p} = \frac{\mu_p^2 B^2}{2}. \quad (9.167)$$

Таким чином, магніторезистивний ефект є ефектом другого порядку за магнітним полем. За простого розгляду ефекту, що привів до формул (9.167), дією холлівського поля знехтовано та прийнято, що всі носії рухаються з однаковою середньою швидкістю і мають однакову середню довжину вільного пробігу. Розглянемо ефект магнітоопору детальніше, базуючись на розв'язку кінетичного рівняння.

За відсутності магнітного поля напрямком густини струму в ізотропному напівпровіднику  $\mathbf{J}$  збігається з напрямком електричного поля  $\mathbf{E}$ . Тоді питомий опір (обернена провідність) визначається як

$$\rho = \frac{1}{\sigma} = \frac{E}{J}. \quad (9.168)$$

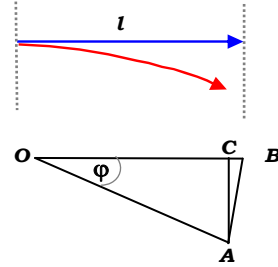


Рис. 9.9. Зміна довжини вільного пробігу носія заряду, спричинена відхиленням від напрямку  $\mathbf{E}$  під дією магнітного поля

За наявності магнітного поля напрямком вектора щільності струму не збігається з напрямком електричного поля  $\mathbf{E}$ , тобто до (9.168) ми маємо підставити замість  $E$  його проекцію на вектор струму –  $\mathbf{E} \cdot \mathbf{J}/J$ . Отже питомий опір за наявності магнітного поля запишемо як

$$\rho_B = \frac{\mathbf{E} \cdot \mathbf{J}}{J^2}. \quad (9.169)$$

Як ми бачили раніше, у слабкому магнітному полі в лінійному наближенні за  $\mathbf{V}$  густина струму має вигляд (див. (9.145))

$$\mathbf{J} = e(n\mu_n + p\mu_p)\mathbf{E} - r_n e(p\mu_p^2 - n\mu_n^2)(\mathbf{V} \times \mathbf{E}).$$

З іншого боку, питомі електронна та діркова провідності визначаються як  $\sigma_n = e\mu_n$  та  $\sigma_p = e\mu_p$ , відповідно. Тоді для густини струму можна записати

$$\mathbf{J} = \sigma\mathbf{E} - R\sigma^2(\mathbf{V} \times \mathbf{E}). \quad (9.170)$$

Зважаючи на те, що  $\mathbf{E} \cdot (\mathbf{V} \times \mathbf{E}) = 0$ , у лінійному за магнітним полем наближенні отримуємо  $J^2 = (\sigma E)^2$  і  $\mathbf{E} \cdot \mathbf{J} = \sigma E^2$ . Таким чином,

$$\rho_B = \frac{1}{\sigma} = \rho, \quad (9.171)$$

тобто, як ми це вже отримали з елементарних міркувань, у слабкому магнітному полі магніторезистивний ефект не є ефектом першого порядку. Для визначення цього ефекту необхідно врахувати доданки, квадратичні за магнітним полем. Для однорідного напівпровідника ( $\nabla E_F = 0$ ) за відсутності градієнта температури в слабкому полі ( $e\tau_e B/m_n^* \ll 1$ ) вектор  $\chi_n$ , що визначає збурену частину функції розподілу електронів, має вигляд

$$\chi_n = -\frac{e\hbar\tau_e}{m_n^*} \left\{ \frac{\mathbf{E} + \frac{e\tau_e}{m_n^*}(\mathbf{V} \times \mathbf{E}) + \left(\frac{e\tau_e}{m_n^*}\right)^2 (\mathbf{V} \cdot \mathbf{E})\mathbf{V}}{1 + \left(\frac{e\tau_e}{m_n^*} B\right)^2} \right\}. \quad (9.172)$$

Цю формулу можна спростити

$$\left[ 1 + \left(\frac{e\tau_e}{m_n^*} B\right)^2 \right]^{-1} \approx 1 - \left(\frac{e\tau_e}{m_n^*} B\right)^2.$$

Отримаємо



$$\begin{aligned}\chi_n &\equiv -\frac{e\hbar\tau_e}{m_n^*} \left\{ \mathbf{E} + \frac{e\tau_e}{m_n^*} (\mathbf{B} \times \mathbf{E}) + \left( \frac{e\tau_e}{m_n^*} \right)^2 (\mathbf{B} \cdot \mathbf{E}) \mathbf{B} \right\} \left\{ 1 - \left( \frac{e\tau_e}{m_n^*} \mathbf{B} \right)^2 \right\} = \\ &= -\frac{e\hbar\tau_e}{m_n^*} \left\{ \mathbf{E} + \frac{e\tau_e}{m_n^*} (\mathbf{B} \times \mathbf{E}) + \left( \frac{e\tau_e}{m_n^*} \right)^2 [(\mathbf{B} \cdot \mathbf{E}) \mathbf{B} - B^2 \mathbf{E}] \right\}.\end{aligned}$$

Оскільки в останньому доданку в дужках стоїть подвійний векторний добуток, запишемо остаточно

$$\chi_n \equiv -\frac{e\hbar\tau_e}{m_n^*} \left\{ \mathbf{E} + \frac{e\tau_e}{m_n^*} (\mathbf{B} \times \mathbf{E}) + \left( \frac{e\tau_e}{m_n^*} \right)^2 (\mathbf{B} \times (\mathbf{B} \times \mathbf{E})) \right\}. \quad (9.173)$$

Якщо обмежитися випадком напівпровідника  $n$ -типу із простою зонною структурою, то для густини струму можна записати

$$\mathbf{J} = -\frac{e^2 \hbar^2}{4\pi^3 (m_n^*)^2} \int_{\Omega} d\mathbf{k} \frac{\partial f_0}{\partial E} \left\{ \tau_e \mathbf{k} \cdot \mathbf{E} + \left( \frac{e\tau_e}{m_n^*} \right)^2 \mathbf{k} \cdot (\mathbf{B} \times \mathbf{E}) + \frac{e^2 \tau_e^3}{(m_n^*)^2} \mathbf{k} \cdot [\mathbf{B} \times (\mathbf{B} \times \mathbf{E})] \right\} \cdot \mathbf{k}. \quad (9.174)$$

У цій формулі наявні інтеграли трьох типів, які можна звести до обчислення інтегралів типу (див. вивід формули (9.84))

$$j_{\tau} = \int_0^{\infty} k^4 dk \tau_e^m \frac{\partial f_0}{\partial E}, \quad m=1, 2, 3. \quad (9.175)$$

Дійсно, із погляду векторної структури (9.174) та з урахування виведення (9.81), видно, що інтеграл складається з інтегралів типу

$$j = \int_{\Omega} d\mathbf{k} F(E) \mathbf{k} (\mathbf{k} \cdot \mathbf{Q}) = \frac{4\pi}{3} \mathbf{Q} \int_0^{\infty} F(E) k^4 dk. \quad (9.176)$$

Оскільки  $\frac{\partial f_0}{\partial E} = -\frac{n}{k_B T N_c} e^{-E/k_B T}$ , то (9.175) можна записати у вигляді

$$j_{\tau} = -\frac{n}{k_B T N_c} \int_0^{\infty} k^4 dk \tau_e^m e^{-E/k_B T}. \quad (9.177)$$

Перейдемо до нової змінної  $t = E/k_B T$  і як завжди припустимо, що закон дисперсії є квадратичним:  $t = (\hbar k)^2 / 2m_n^* k_B T$ ,  $k dk = (m_n^* k_B T / \hbar^2) dt$ ,  $k = (2m_n^* k_B T / \hbar^2)^{1/2} t^{1/2}$ . Таким чином, маємо

$$j_{\tau} = -\frac{n(2m_n^*)^{5/2} (k_B T)^{3/2}}{2\hbar^5 N_c} \int_0^{\infty} \tau_e^m e^{-t} t^{3/2} dt.$$

Введемо позначення  $\langle \tau_e^m \rangle = \frac{4}{3\sqrt{\pi}} \int_0^{\infty} \tau_e^m e^{-t} t^{3/2} dt$ , (9.178)

та із (9.5.58) отримаємо

$$\mathbf{J} = en\mu_n \left\{ \mathbf{E} + \mu_n \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} (\mathbf{V} \times \mathbf{E}) + \mu_n^2 \frac{\langle \tau^3 \rangle}{\langle \tau \rangle^3} [\mathbf{V} \times (\mathbf{V} \times \mathbf{E})] \right\}. \quad (9.179)$$

$\mathbf{J}$ ,  $\mathbf{V}$ ,  $\mathbf{V} \times \mathbf{J}$  є набором некопланарних векторів. Для визначення поля  $\mathbf{E}$  розкладемо його за цим набором

$$\mathbf{E} = a_1 \mathbf{J} + a_2 \mathbf{V} + a_3 (\mathbf{V} \times \mathbf{J}) \quad (9.180)$$

і підставимо розкладання до (9.179), залишаючи доданки порядку не вище, як  $V^2$

$$\begin{aligned} \mathbf{J} = en\mu_n \left\{ a_1 \mathbf{J} + a_2 \mathbf{V} + a_3 \mathbf{V} \times \mathbf{J} + \mu_n \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} a_1 \mathbf{V} \times \mathbf{J} + \right. \\ \left. + \mu_n \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} a_3 \mathbf{V} \times (\mathbf{V} \times \mathbf{J}) + \mu_n^2 \frac{\langle \tau^3 \rangle}{\langle \tau \rangle^3} a_1 \mathbf{V} \times (\mathbf{V} \times \mathbf{J}) \right\}. \end{aligned}$$

Розкриваючи подвійні векторні добутки, дорівнюємо коефіцієнти при векторах  $\mathbf{J}$ ,  $\mathbf{V}$ ,  $\mathbf{V} \times \mathbf{J}$ . Маємо такі рівняння для коефіцієнтів  $a_i$ :

$$\mathbf{J}: \quad 1 = en\mu_n a_1 - en\mu_n^2 \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} B^2 a_3 - en\mu_n^3 \frac{\langle \tau^3 \rangle}{\langle \tau \rangle^3} B^2 a_1,$$

$$\mathbf{V}: \quad en\mu_n a_2 + en\mu_n^2 \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} (\mathbf{V} \cdot \mathbf{J}) a_3 + en\mu_n^3 \frac{\langle \tau^3 \rangle}{\langle \tau \rangle^3} (\mathbf{V} \cdot \mathbf{J}) a_1 = 0,$$

$$(\mathbf{V} \cdot \mathbf{J}): \quad a_3 + \mu_n \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} a_1 = 0.$$

Звідси отримуємо

$$a_1 = \frac{1}{en\mu_n} [1 + \mu_n^2 B^2 T_M], \quad (9.181)$$

$$a_2 = -\frac{\mu_n}{en} T_M (\mathbf{V} \cdot \mathbf{J}), \quad (9.182)$$

$$a_3 \cong -\frac{1}{en} \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} (1 + \mu_n^2 B^2 T_M), \quad (9.183)$$

де  $T_M = \frac{\langle \tau_e^3 \rangle \langle \tau_e \rangle - \langle \tau_e^2 \rangle^2}{\langle \tau_e \rangle^4}$ . (9.184)

Після підстановки цих виразів до (9.180) знаходимо

$$\mathbf{E} = \frac{1}{en\mu_n}(1 + \mu_n^2 B^2 T_M) \mathbf{J} - \frac{\mu_n}{en} T_M (1 + \mu_n^2 B^2 T_M) (\mathbf{B} \cdot \mathbf{J}) \mathbf{B} - \frac{1}{en} \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} (1 + \mu_n^2 B^2 T_M) (\mathbf{B} \times \mathbf{J}). \quad (9.185)$$

Скалярно домножимо це рівняння на  $\mathbf{J}$  і поділімо на  $J^2$ . При цьому пам'ятаємо, що  $\mathbf{E} \cdot \mathbf{J} / J^2 = \rho_B$ , а при  $\mathbf{B} = 0$  маємо питомий опір напівпровідника  $\rho = E/J = 1/en\mu_n$ . Тоді відносна зміна питомого опору напівпровідника в магнітному полі

$$\frac{\rho_B - \rho}{\rho} = \mu_n^2 B^2 T_M - \mu_n^2 T_M \frac{(\mathbf{J} \cdot \mathbf{B})^2}{J^2}. \quad (9.186)$$

Якщо  $\varphi$  – кут між векторами густини струму та магнітного поля, то із цього рівняння можна записати

$$\frac{\rho_B - \rho}{\rho} = \mu_n^2 B^2 T_M \sin^2 \varphi, \quad (9.187)$$

Таким чином, для напівпровідника  $n$ -( $p$ )-типу зі сферичними ізоенергетичними поверхнями в слабких полях поперечний магнітоопір ( $\varphi = \pi/2$ ) пропорційний квадрату магнітного поля та рухливості. Поздовжній магнітоопір ( $\varphi = 0$ ) у цьому випадку є нульовим. При цьому у випадку, коли час релаксації не залежить від енергії, коефіцієнт  $T_M$  дорівнює нулю. У результаті ефект магнітоопору не спостерігатиметься.

Формулу, аналогічну (9.187), можна отримати й у загальнішому випадку, коли напівпровідник має носії обох знаків. Тоді до (9.179) необхідно додати рівняння діркової складової густини струму

$$\mathbf{J}_p = en\mu_p \left\{ \mathbf{E} - r_H \mu_p (\mathbf{B} \times \mathbf{E}) + \frac{3}{2} \gamma \mu_p^2 [\mathbf{B} \times (\mathbf{B} \times \mathbf{E})] \right\}, \quad \text{де } \gamma = \frac{\langle \tau_e^3 \rangle}{\langle \tau_e \rangle^3}.$$

Повний струм можна записати у вигляді

$$\mathbf{J} = \sigma \left\{ \mathbf{E} + r_H \frac{n\mu_n^2 - p\mu_p^2}{n\mu_n + p\mu_p} (\mathbf{B} \times \mathbf{E}) + \gamma \frac{n\mu_n^3 + p\mu_p^3}{n\mu_n + p\mu_p} [\mathbf{B} \times (\mathbf{B} \times \mathbf{E})] \right\}. \quad (9.188)$$

І, як результат, маємо для напівпровідника із простою зонною структурою, що характеризується двома типами носіїв,

$$\frac{\rho_B - \rho}{\rho} = \frac{e}{\sigma} \left[ r_H (n\mu_n^3 + p\mu_p^3) - \gamma \frac{(n\mu_n^2 - p\mu_p^2)^2}{(n\mu_n^2 + p\mu_p^2)} \right] B^2 \sin^2 \varphi. \quad (9.189)$$

Видно, що основні висновки про магнітоопір зберігаються і в цьому, загальнішому випадку. Але є одна особливість: на відміну від напівпровідника, де електричний струм забезпечується носіями одного типу, у

власного або компенсованого напівпровідника магніторезистивний ефект виявляється й у випадку, коли час релаксації не залежить від енергії. Дійсно, як видно із (9.189), при  $r_H = 1$  та  $\gamma = 1$  величина  $\Delta\rho/\rho \neq 0$ . Структура виразу (9.189) свідчить, що перший доданок у квадратних дужках правої частини визначає магнітоопір, обумовлений "закручуванням" траєкторій рухомих носіїв магнітним полем, а другий – описує компенсуючу дію поля Холла, що виникає завдяки обмеженості зразка вздовж вісі ОУ. Цей факт дозволяє дійти висновку, що явище магнітоопору сильно залежить від конфігурації зразка, оскільки поперечне поле Холла, яке виникає під дією магнітного поля, заважає змінити траєкторії носіїв і зменшує ефект магнітоопору. Дійсно, із (9.189) ми бачимо, що  $\Delta\rho/\rho$  досягає максимального значення, коли другий доданок у дужках дорівнює нулю. А це означає, що стала Холла  $R = 0$ , тобто для мінімального впливу ефекту Холла ( $R = 0$ ) варто використовувати зразки, де поле Холла є нульовим. Таким зразком є диск Корбіно (рис. 9.10), де один контакт розташований уздовж зовнішнього, а другий – уздовж внутрішнього кола отвору всередині зразка. Таким чином можна стверджувати, що короткі широкі зразки з омичними контактами на довгих боках, що шунтують ЕРС Холла, дозволяють вимірювати помітніші відносні зміни магнітоопору.

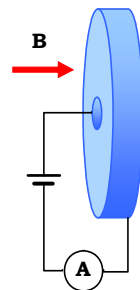


Рис. 9.10. Диск Корбіно

На рис. 9.11 подано залежність відносного магнітоопору від величини магнітного поля для зразків різної форми. Вище було розглянуто найпростіший випадок сферично симетричного напівпровідника. У напівпровідниках, що не мають сферично симетричної електронної структури, зміна опору в магнітному полі залежить від орієнтації вектора магнітного поля щодо кристалічних осей і може існувати повздовжній магніторезистивний ефект. Вимірюючи величину  $\Delta\rho/\rho$  за різних орієнтацій магнітного поля щодо осей зразка, можна отримати відомості про параметри зонної структури.

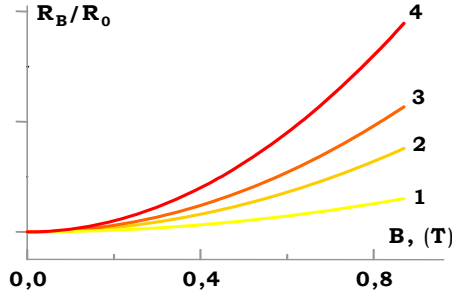


Рис. 9.11. Залежність опору зразків різної форми від магнітного поля.

Цифрами позначено зразки з відношеннями довжини до ширини:  
1 – 3; 2 – 1; 3 – 1/4; 4 – диск Корбіно

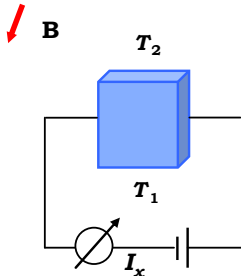
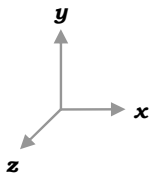
#### 9.5.4. Ефекти Еттингсгаузена та Нернста

При вимірюванні ефекту Холла іноді виникає додаткове поперечне поле за рахунок різниці температур на бічних гранях зразка. Фізична природа цього явища полягає в тому, що сила Лоренца, яка діє на електрон у магнітному полі, пропорційна його швидкості. Це означає, що електрони із більшими швидкостями відхиляються магнітним полем сильніше. Тому бік зразка, в який відхилено більше електронів із великими швидкостями, сильніше нагріватиметься (тут відносно більше швидких частинок, тобто більша температура). Ефект виникнення градієнта температури на бічних гранях зразка в магнітному полі при протіканні через нього струму називається *ефектом Еттингсгаузена*. У поданій на рис. 9.12 геометрії величина градієнта температури визначається формулою

$$\frac{\partial T}{\partial y} = P_E j_x B_z \quad (9.190)$$

Коефіцієнт Еттингсгаузена  $P_E$  вимірюється в  $\text{см}^3\text{К}\cdot\text{Вт}^{-1}\text{с}^{-1}$ . Для напівпровідника  $n$ -типу обчислення цього коефіцієнта дає

$$P_E = \frac{\mu_n}{ek} \left( \frac{\langle E\tau_n^2 \rangle}{\langle \tau_n \rangle^2} - \frac{\langle \tau_n^2 \rangle \langle E\tau_n \rangle}{\langle \tau_n \rangle^3} \right), \quad (9.191)$$



де  $k$  – теплопровідність зразка. Видно, що коефіцієнт Еттингсгаузена дорівнює нулю, якщо час релаксації не залежить від енергії. Різниця температур на бічних

**Рис. 9.12. Схема вимірювання ефекту Еттингсгаузена.**

У поданій геометрії верхня грань зразка охолоджується, а нижня – нагрівається, тобто  $T_1 > T_2$

гранях зразка призводить до виникнення термоелектричної різниці потенціалів на контактах метал–напівпровідник, що додається до різниці потенціалів, обумовленої ефектом Холла. Для зразка з лінійними розмірами  $1 \times 10^{-1} \times 10^{-2} \text{ см}^3$ , що знаходиться в магнітному полі  $B_z = 0,1 \text{ Т}$ , через який проходить струм  $10^{-3} \text{ А}$ , коефіцієнт Еттингсгаузена  $P_E = 60 \text{ см}^3 \text{ К} \cdot \text{Вт}^{-1} \text{ с}^{-1}$ . Різниця потенціалів, що визначається ефектом Еттингсгаузена  $V_E$ , при цьому становить  $3 \cdot 10^{-9} \text{ В}$ . Оскільки відношення  $V_E/V_H \cong 10^{-5}$  (де  $V_H$  – е.р.с. Холла), то зазвичай внеском цього ефекту можна знехтувати.

Ще одним гальваномагнітним ефектом є ефект Нернста: у зразку, до якого прикладене магнітне поле, перпендикулярне до напрямку електричного, виникає повздовжня різниця температури. Виникнення градієнта температури, паралельного електричному полю, пояснюється тим, що більша сила Лоренца діє на швидкіші електрони, унаслідок чого вони відокремлюються геометрично від повільних, на які діє менша сила Лоренца. У результаті вздовж грані, на яку вони відхилитимуться, виникне градієнт температури. Ефект Нерста є прикладом повздовжнього гальваномагнітного ефекту.

## ■ 9.6. Термомагнітні ефекти

Термомагнітними ефектами називається група явищ, що пов'язані із впливом магнітного поля на електричні та теплові властивості напівпровідників, в яких утворено градієнт температури. Термомагнітні явища, як і гальваномагнітні, обумовлені викривленням траєкторії носіїв заряду в магнітному полі. Іншими словами, магнітне поле викривляє електричний струм (що протікає зразком), спричинений перенесенням заряджених частинок потоком теплоти. У результаті з'являється складова електричного струму і теплового потоку, перпендикулярна магнітному полю. При цьому можуть змінюватись електронні властивості зразка вздовж градієнта температур.

До термомагнітних явищ належать повздовжній і поперечний ефекти Нернста–Еттингсгаузена та повздовжній і поперечний ефекти Риги–Ледюка. Оскільки термомагнітні ефекти не мають такого широкого застосування, як гальваномагнітні ефекти Холла та магнітоопору, зупинимось коротко на формулюванні деяких експериментальних аспектів їх визначення. Зрозуміло, що всі термомагнітні явища можна послідовно розглянути в рамках розв'язку кінетичного

рівняння, і коефіцієнти Нернста–Еттингсгаузена та Риги–Ледюка – обчислити в термінах середніх часів релаксації. Найпростіше можна виміряти поперечний ефект Нернста–Еттингсгаузена, який полягає в тому, що наявність градієнта температури вздовж вісі ОХ зразка, що є причиною теплового потоку вздовж цієї вісі, приводить до появи в магнітному полі на поперечних гранях зразка різниці потенціалів. Його фізична природа викликана тим, що потік тепла в напрямку вісі ОХ викликає результуючий струм швидких електронів від гарячого кінця зразка до холодного. Оскільки відхилення електронів у магнітному полі пропорційне їхній швидкості, то як результат у напрямку ОУ виникають одночасно і поперечна ЕРС, і поперечний градієнт температури. Для вимірювання ефекту Нернста використовують схему, аналогічну до тієї, що застосовується при вимірюванні ефекту Холла (рис. 9.13). Якщо в цьому експерименті на поперечних гранях зразка підтримувати різні температури й тим самим створювати градієнт температури  $\partial T/\partial x$ , то вздовж зразка протікатиме потік тепла, а між верхньою та нижньою гранями виникне різниця потенціалів  $V_y$ . Напруженість поля Нернста при цьому визначається формулою

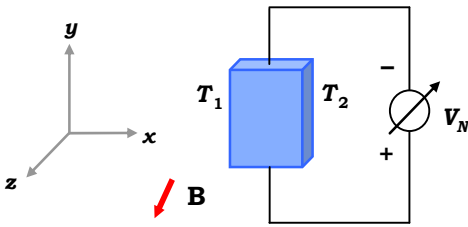


Рис. 9.13. Схема вимірювання ефекту Нернста–Еттингсгаузена.

При  $T_1 > T_2$  у представленій геометрії для зразка  $n$ -типу коефіцієнт Нернста–Еттингсгаузена є позитивним

$$E_y = Q_N \frac{\partial T}{\partial x} B_z. \quad (9.192)$$

Коефіцієнт Нернста  $Q_N$  (вимірюється в  $\text{см}^2\text{с}^{-1}\text{К}^{-1}$ ) вважається позитивним, якщо вектори  $\nabla T$ ,  $\mathbf{E}$ ,  $\mathbf{B}$  утворюють праву трійку векторів. Він виражається через середні часи релаксації і для напівпровідника  $n$ -типу

$$Q_N = \frac{\mu_n}{T} \left( \frac{\langle E\tau_n^2 \rangle}{\langle \tau_n \rangle^2} - \frac{\langle \tau_e^2 \rangle \langle E\tau_n \rangle}{\langle \tau_n \rangle^3} \right). \quad (9.193)$$

Тут і нижче використовуватимемо таке позначення

$$\langle E^p \tau_e^m \rangle = \frac{4}{3\sqrt{\pi}} \int_0^{\infty} t^p \tau_e^m e^{-t^2} dt. \quad (9.194)$$

Як видно з цієї формули, у випадку, коли час релаксації не залежить від енергії,  $Q_N = 0$ . Для зразка з лінійними розмірами  $1 \times 10^{-1} \times 10^{-2} \text{ см}^3$ , що знаходиться в магнітному полі  $B_z = 0,1 \text{ Т}$ , і градієнтом температури  $10 \text{ К см}^{-1}$

коефіцієнт Нернста–Еттингсгаузена  $Q_N \cong 10^{-1} \text{м}^2 \text{с}^{-1} \text{К}^{-1}$ . Часто цей ефект називають тепловим аналогом ефекту Холла (див., напр. [1]). Але ця аналогія не зовсім повна, оскільки за незалежного від енергії часу релаксації поперечний ефект Нернста–Еттингсгаузена відсутній (див. (9.193)), у той час, як при  $\tau(E) = \text{const}$  ефект Холла спостерігається.

Якщо в геометрії експерименту (рис. 9.14) нагрівати зразок так, щоб у напрямку ОХ утворювався градієнт температури, то за ненульового магнітного поля виникне градієнт температури також вздовж вісі ОУ. Таке явище називається ефектом Риги–Ледюка, виникнення якого можна пояснити: оскільки відхилення електронів у магнітному полі пропорційне їхнім швидкостям, то за наявності градієнта температури вздовж вісі ОХ швидкіші електрони, що знаходяться біля гарячої грані зразка, відхиляються дужче в напрямку ОУ. Це приводить до концентрації на нижній грані зразка швидкіших електронів, отже температура цієї грані буде більшою за температуру верхньої. Іншими словами, ефект Риги–Ледюка є температурним аналогом ефекту Еттингсгаузена. Величина поперечного градієнта температур у геометрії, поданій на рис. 9.14, визначається формулою

$$\frac{\partial T}{\partial y} = S_{RL} \frac{\partial T}{\partial x} B_z, \tag{9.195}$$

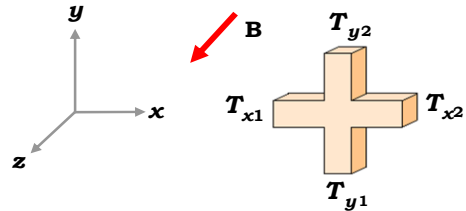
де коефіцієнт Риги–Ледюка  $S_{RL}$  (вимірюється в  $\text{см}^2 \text{В}^{-1} \text{с}^{-1}$ ) для зразка із лінійними розмірами  $1 \times 10^{-1} \times 10^{-2} \text{ см}^3$ , що знаходиться в магнітному полі,  $B_z = 0,1 \text{ Т}$  в умовах нагріву, який забезпечують градієнт температури  $\partial T / \partial x = 10 \text{ Ксм}^{-1}$ , становить величину  $2 \cdot 10^{-2} \text{ см}^2 \text{В}^{-1} \text{с}^{-1}$ . Обчислення коефіцієнта Риги–Ледюка для напівпровідника n-типу дає

$$S_{R-L} = -\frac{n\mu_n^2}{ekT} \left( \frac{\langle E^2 \tau_n^2 \rangle}{\langle \tau_n \rangle^2} + \frac{\langle E \tau_n \rangle^2}{\langle \tau_n \rangle^4} - 2 \frac{\langle E \tau_n^2 \rangle \langle E \tau_n \rangle}{\langle \tau_n \rangle^3} \right), \tag{9.196}$$

де  $\kappa$  – теплопровідність.

Якщо залежність часу релаксації від енергії є ступеневою, тобто  $\tau_n = \alpha \cdot E^{-s}$ , то вираз (9.196) перейде у

$$S_{R-L} = -nk^2 \mu_n^2 Tr \frac{5 - 4s + 2s^2}{2ek}.$$



**Рис. 9.14.** Схема вимірювання ефекту Риги–Ледюка.

У представленій геометрії за рахунок наявності градієнта температури вздовж вісі ОХ ( $T_{x1} > T_{x2}$ ) верхня грань зразка охолоджується, а нижня – нагрівається ( $T_{y1} > T_{y2}$ )



Із цієї формули видно: якщо час релаксації не залежить від енергії ( $s = 0$ ), то коефіцієнт Риги–Ледюка не є нулем. Цей факт дозволяє стверджувати, що поперечний ефект Риги–Ледюка й є тепловим аналогом ефекту Холла.

Повздовжній ефект Риги–Ледюка (зміна електронної теплопровідності) та повздовжній ефект Нернста–Еттингсгаузена (виникнення повздовжньої різниці потенціалів) можна розглянути аналогічно. У табл. 9.2 подано порівняльний опис гальваномагнітних і термомагнітних ефектів.

Таблиця 9.2

Ефекти	
гальваномагнітні	термомагнітні
<i>поперечні</i>	
<p><i>ефект Холла</i> – виникнення різниці потенціалів на бічних гранях зразка, уздовж якого протікає електричний струм у магнітному полі, перпендикулярному до електричного.</p> <p>При <math>\tau(E) = \text{const}</math> <math>R \neq 0</math></p>	<p><i>ефект Риги–Ледюка</i> – виникнення поперечної різниці температур у зразку, де утворено тепловий потік і на який діє магнітне поле, перпендикулярне потоку.</p> <p>При <math>\tau(E) = \text{const}</math> <math>S_{R-L}^{\perp} \neq 0</math></p>
<p><i>ефект Еттингсгаузена</i> – виникнення поперечної різниці температур у магнітному полі, що перпендикулярне до напрямку електричного поля.</p> <p>При <math>\tau(E) = \text{const}</math> <math>P_E = 0</math></p>	<p><i>ефект Нернста–Еттингсгаузена</i> – виникнення поперечної різниці потенціалів у магнітному полі, що перпендикулярне до напрямку теплового потоку.</p> <p>При <math>\tau(E) = \text{const}</math> <math>P_{N-E} = 0</math></p>
<i>повздовжні</i>	
<p><i>магніторезистивний ефект</i> – зміна швидкості дрейфу носіїв заряду внаслідок дії поперечного магнітного поля</p>	<p><i>повздовжній ефект Риги–Ледюка</i> – зміна електронної теплопровідності в умовах дії магнітного поля, перпендикулярного до потоку тепла</p>
<p><i>ефект Нернста</i> – виникнення повздовжньої різниці температур у магнітному полі, що перпендикулярне до напрямку електричного поля</p>	<p><i>повздовжній ефект Ернста–Еттингсгаузена</i> – виникнення повздовжньої різниці потенціалів за дії магнітного поля, перпендикулярного до потоку тепла, утвореного у напівпровіднику.</p>

## ■ 9.7. Термоелектричні явища

Ефекти Зеебека, Пельтьє та Томсона складають групу так званих термоелектричних ефектів. Перший пов'язаний з тим, що при нагріванні одного кінця зразка енергія носіїв підвищується, і вони рухаються до холодного кінця зразка, унаслідок чого порушується електронна рівновага, яка спричиняє зменшення їхньої концентрації на гарячій ділянці. У той же час гарячі носії, що дифундують до холодного

кінця зразка, приносять із собою додатковий заряд, унаслідок чого виникає різниця потенціалів або термо-ЕРС, величина якої пов'язується з різницею температур між холодним і гарячим кінцями зразка коефіцієнтом термо-ЕРС  $\alpha$

$$U_{терм.} = \alpha \cdot \Delta T . \tag{9.197}$$

Експериментально явище Зеебека можна вивчати при вимірюванні різниці потенціалів у схемі (рис. 9.15), яка складається з напівпровідникового зразка, кінці якого нагріті до різних температур. Металічними дротами зразок приєднано до вольтметра. Контакт 2 нагрітий до температури  $T_1$ , а контакт 3 – до  $T_2$ . Між точками 1–4 вимірюється різниця потенціалів, яка складається із різниць потенціалів на ділянках 1–2, 2–3 та 3–4

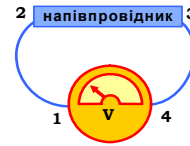


Рис. 9.15. Схема пристрою для вимірювання ефекту Зеебека

$$U_{терм.} = \int_1^2 E_{терм.}(M) dx + \int_2^3 E_{терм.}(S) dx + \int_3^4 E_{терм.}(M) dx .$$

Якщо контакти 1 та 4 мають однакову температуру  $T_0$ , то із врахуванням (9.197), і вважаючи, що  $T_2 > T_1 > T_0$ , запишемо

$$U_{терм.} = \int_{T_0}^{T_1} \alpha_M dT + \int_{T_2}^{T_0} \alpha_M dT + \int_{T_1}^{T_2} \alpha_S dT = \int_{T_1}^{T_2} (\alpha_S - \alpha_M) dT . \tag{9.198}$$

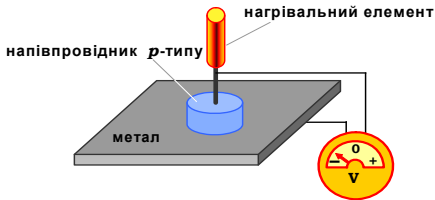
Оскільки зазвичай коефіцієнт термо-ЕРС напівпровідника набагато більший за коефіцієнт термо-ЕРС металу, то основний внесок в інтеграл у (9.198) дає перший доданок. Таким чином, вимірюючи різницю потенціалів у точках 1 та 4 за достатньо великих температур, коли залежність рівня Фермі від температури є лінійною (область власної провідності), можна отримати величину  $\alpha_S$

$$\alpha_S = \frac{U_{терм.}}{T_2 - T_1} . \tag{9.199}$$

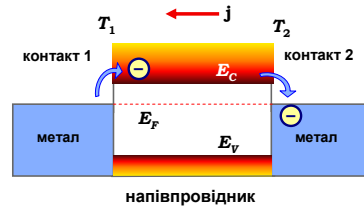
Ефект термо-ЕРС застосовують для визначення типу провідності напівпровідника. Типову схему експерименту подано на рис. 9.16. На холодній металевій підкладці розташовують зразок, до верхньої грані якого притиснутий гарячий зонд. Для напівпровідника  $p$ -типу, як показано на рисунку, потенціал зонда буде негативним, а підкладки – позитивним.

Ефект Пельтье полягає в тому, що при проходженні електричного струму через структуру **метал-напівпровідник-метал** на одному кон-

такті метал–напівпровідник буде поглинатись, а на іншому – виділяється (залежно від напрямку струму) тепло. Розглянемо контакт напівпровідника  $n$ -типу з металом (рис. 9.17). Нехай струм протікає від контакту 2 до контакту 1, тобто електрони рухаються зліва направо. На першому контакті, де енергія електронів в металі  $E_F$  менша від енергії електронів у напівпровіднику  $E_C$ , вони для того, щоб потрапити до зони провідності напівпровідника мають подолати енергетичний бар'єр  $E_C - E_F$ . Отже, до напівпровідника можуть перейти лише найбільш енергійні електрони, унаслідок чого електронний газ у металі охолоджується.



**Рис. 9.16.** Схема експерименту для термоелектричного визначення типу провідності напівпровідника за допомогою гарячого зонда



**Рис. 9.17.** Схема виникнення ефекту Пельтьє на контактах метал–напівпровідник–метал

На відновлення рівноважного розподілу витратиться енергія ґратки, унаслідок чого температура металу на контакті 1 знизиться. На контакті 2 енергія електронів, що переходять із напівпровідника до металу, істотно перевищує енергію Фермі в металі, унаслідок чого в металі поблизу контакту 2 порушується теплова рівновага, яка швидко відновлюється завдяки процесам розсіювання високоенергетичних електронів на коливаннях ґратки. У результаті електрони швидко термалізуються, віддаючи надлишкову енергію ґратці, і контакт 2 нагрівається. Тепло, виділене (або поглинене) на контакті за одиницю часу, пропорційне струму (коефіцієнт пропорційності  $\Pi$  називається коефіцієнтом Пельтьє), що протікає через контакт

$$Q = \Pi \cdot j. \quad (9.200)$$

На сучасних елементах Пельтьє за оптимальних умов можна отримати різницю температур між холодним і гарячим контактами  $\Delta T \approx 20$  К. Такі елементи Пельтьє, з'єднані послідовно, дозволяють отримати різницю температур до 200 К. Твердотільні охолоджувальні елементи Пельтьє використовуються в пристроях автоматизації та наукових застосуваннях, оскільки керування температурою в них забезпечується лише управлінням струмом, що є дуже зручним.

*Ефект Томсона:* якщо вздовж напівпровідника утворено градієнт температури, то при проходженні в об'ємі матеріалу струму, крім Джоулевого тепла, виділяється (або поглинається – залежно від на-

прямку струму) залежна від різниці температур теплота, кількість якої за одиницю часу пропорційна електричному струму

$$Q_T = j \int_{T_1}^{T_2} \tau_T dT, \quad (9.201)$$

де  $\tau_T$  – коефіцієнт Томсона. Якою є фізична причина цього ефекту? Якщо вздовж напівпровідника протікає електричний струм, направлений за градієнтом температури, то при переході електронів від більш нагрітої ділянки до холодної вони загальмовуються і передають надлишкову енергію в ґратку, завдяки чому виділяється тепло. За зворотного напрямку руху електронів вони переходять від холодної ділянки до більш нагрітої і поповнюють свою енергію за рахунок енергії ґратки, у результаті чого тепло поглинається.

## ■ 9.8. Задачі

1. Рухливість електронів у власному германії за температури  $T = 300$  К  $\mu_n = 3,8 \cdot 10^3$  см<sup>2</sup>В<sup>-1</sup>с<sup>-1</sup>. Вважаючи, що в широкому інтервалі температур ширина забороненої зони лінійно залежить від температури ( $E_g = E_g^0 - \alpha T$ ,  $E_g^0 = 0,785$  еВ,  $\alpha = 4 \cdot 10^{-4}$  еВ · К<sup>-1</sup>), знайдіть питомий опір матеріалу за температури 77 К. Візьміть до уваги, що рухливість носіїв змінюється з температурою за законом  $\mu = aT^{-3/2}$  ( $a$  – стала). Ефективна маса електронів –  $m_n^* = 0,55m_n$  і дірок –  $m_p^* = 0,36m_e$ . Вважайте відношення рухливості електронів і дірок постійним для всіх температур  $b = \mu_n/\mu_p = 2,1$ .

2. Питомий опір власного германію  $\rho_{Ge} = 47$  Ом · см, рухливостями електронів і дірок є  $\mu_n = 3,8 \cdot 10^3$  см<sup>2</sup>В<sup>-1</sup>с<sup>-1</sup>. та  $\mu_p = 1,9 \cdot 10^3$  см<sup>2</sup>В<sup>-1</sup>с<sup>-1</sup>, відповідно. Знайдіть концентрацію носіїв заряду за кімнатної температури.

3. Покажіть, що магніторезистивний ефект в основному визначається носіями із великою рухливістю.

### Список літератури

1. Смит Р. Полупроводники. – М.: Мир, 1982.
2. Зеегер К. Физика полупроводников. – М.: Мир, 1977.
3. Лифшиц Е.М., Питаевский Л.П. Физическая кинетика. – М.: Наука, 1979.
4. Characterization of Semiconductor Materials. Principles and Methods. V. 1. Ed. by McGuire G.E. – Park Ridge: Noyes Publ., 1989.
5. Kasap S.O. Principles of Electronic Materials and Devices. – N.Y.: McGraw-Hill, 2005.
6. Neamen D. Semiconductor Physics and Devices. – N.Y.: McGraw-Hill, 2003.

# Розділ 10

## ГЕНЕРАЦІЯ ТА РЕКОМБІНАЦІЯ НОСІЇВ ЗАРЯДУ У НАПІВПРОВІДНИКАХ

У попередніх розділах ми обговорювали залежність концентрації носіїв заряду від температури у напівпровідниках в умовах термодинамічної рівноваги, тобто мали справу з *рівноважними концентраціями* носіїв. При вивченні кінетичних явищ в слабких полях ми вперше вникли у тому, що концентрація носіїв також лишалась рівноважною. Хоча, зрозуміло, що в умовах, коли у напівпровіднику протікає електричний струм, стан напівпровідника є термодинамічно нерівноважним. Дійсно, під дією зовнішніх полів рівноважна функція розподілу утримує нерівноважний доданок

$$f(\mathbf{r}, \mathbf{k}) = f_0(\mathbf{r}, \mathbf{k}) + f_1(\mathbf{r}, \mathbf{k}), \quad (10.1)$$

і концентрація носіїв заряду є такою

$$n(\mathbf{r}) = \frac{1}{4\pi^3} \int d\mathbf{k} \left[ f_0(\mathbf{r}, \mathbf{k}) - \frac{\hbar}{m^*} \frac{\partial f_0}{\partial E} \boldsymbol{\chi} \cdot \mathbf{k} \right] = n_0(\mathbf{r}), \quad (10.2)$$

тобто нерівноважному стану (10.1) відповідає рівноважна концентрація носіїв заряду (10.2). При виводі формули (10.2) ми скористалися тим, що другий доданок підінтегральної функції є непарною за хвильовим вектором функцією і, як результат, інтеграл від такої функції є нулем. Умова (10.2) відповідає рівнянню  $df/dt = 0$ , звідки випливає рівняння неперервності

$$\frac{\partial \rho}{\partial t} + \text{div} \mathbf{j} = 0, \quad (10.3)$$

де повна густина заряду  $\rho = e(n - p)$ . Іншими словами, в умовах, коли повна кількість зарядів у системі є сталою (як у металів), об'ємна щільність заряду може змінюватись тільки у результаті дивергенції струму. Але в напівпровіднику зміна  $\rho$  може відбуватися ще й за рахунок процесів генерації–рекомбінації (див. нижче), коли концентрація носіїв заряду окремо може змінюватись так, щоб повна густина заряду була незмінною. Таким чином, у напівпровіднику може існувати дуже специфічний нерівноважний стан, за якого концентрація носіїв заряду є сталою (рівноважною), але парціальні концентрації електронів і дірок можуть змінюватись. При цьому відбувається локальне порушення електронейтральності, що приводить до виникнення електричних полів, які, у свою чергу, спричиняють локальні електричні струми.

У цьому розділі розглядаються процеси, пов'язані зі зміною концентрації носіїв заряду. Ці процеси викликані наявністю нерівноважних

станів у напівпровідниках, тому основними завданнями при їх вивченні буде знаходження часів життя нерівноважних носіїв заряду, що визначають час релаксації нерівноважного стану напівпровідника до рівноважного.

## 10.1. Розподіл нерівноважних носіїв заряду за енергіями

У рівноважних напівпровідниках електрони й дірки заповнюють стани поблизу дна зони провідності та стелі валентної зони й загалом характеризуються визначеною функцією розподілу. Як же розподілені за енергіями нерівноважні носії заряду в напівпровіднику? Розглянемо процес виникнення пари нерівноважних носіїв заряду за рахунок дії зовнішнього випромінювання квантами достатньо великої енергії  $\hbar\omega > E_g$ . Безпосередньо після процесу виникнення електрон-діркової пари їхня енергія може суттєво перевищувати середню енергію рівноважних носіїв у напівпровіднику, яка за порядком величини дорівнює  $kT$ . Але у подальшому за рахунок взаємодії із ґраткою кінетична енергія нерівноважних носіїв швидко зменшується до значень  $\sim kT$ : нерівноважні носії швидко набувають температури ґратки. Наскільки швидко відбуваються такі процеси термалізації? Розрахунки показують (а експерименти підтверджують), що електрон з енергією  $\sim 1$  еВ роздає надлишкову енергію фононам за  $\sim 10^3$  зіткнень із ґраткою, тобто за довжини вільного пробігу  $\sim 10^{-6}$  см і швидкості  $\sim 10^7$  см/с електрон позбудеться надлишкової енергії (що перевищує величину  $kT$ ) протягом часу  $\sim 10^{-10}$  с. Час життя носіїв у зонах (що виникли за рахунок термічних процесів) складає величини  $\sim 10^{-2} - 10^{-7}$  с. Отже більшу частину вільного існування нерівноважні носії заряду перебувають у станах з енергіями поблизу країв зон і не відрізняються за своєю середньою кінетичною енергією від рівноважних. У більшості випадків можна вважати, що розподіл за енергіями рівноважних і нерівноважних носіїв у зоні є однаковим (якщо генерація нерівноважних носіїв не значно вплинула на концентрацію носіїв, і напівпровідник зберіг свій невироджений стан).

Таким чином, генерація нерівноважних носіїв заряду в напівпровіднику викликає лише зміну концентрації вільних носіїв, не змінюючи при цьому їхнього розподілу за енергіями.

## ■ 10.2. Механізми рекомбінації носіїв заряду

На відміну від металів, у напівпровідниках існують додаткові можливості, коли концентрація носіїв заряду може змінюватись із часом, лишаючи при цьому зразок електронейтральним. Наприклад, при освітленні світлом напівпровідника можуть розриватись валентні зв'язки, що спричинить появу "вільного" електрона у зоні провідності та дірки – у валентній зоні: виникне електронно-діркова пара (рис. 10.1 а). Процеси генерації можуть відбуватись також й за рахунок домішкового поглинання (рис. 10.1 б, в)

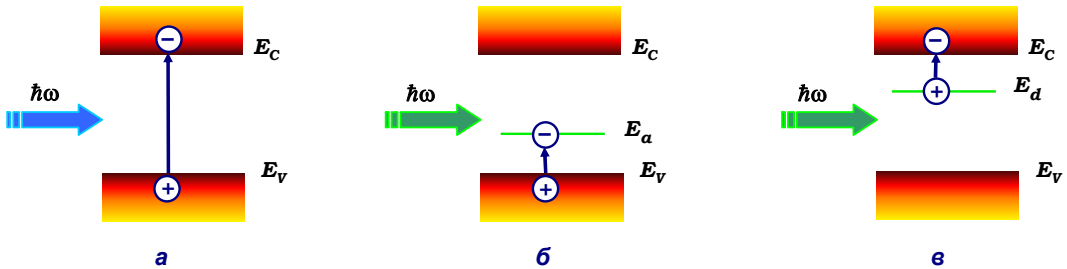


Рис. 10.1. Схема процесів генерації носіїв типу зона-зона:

а – за поглинання світла з енергією фотонів, більшу від ширини забороненої зони ( $\hbar\omega > E_g$ );

б – за домішкового поглинання світла через акцепторний рівень ( $\hbar\omega > E_a - E_v$ );

в – за домішкового поглинання світла через донорний рівень ( $\hbar\omega > E_c - E_d$ )

До такого самого результату може привести дія сильних електричних полів або інжекція носіїв та ін.

**Якщо діє додаткове до теплового джерело виникнення носіїв заряду, то кажуть, що відбувається генерація нерівноважних носіїв заряду у напівпровіднику, унаслідок якої у зразку збільшується концентрація носіїв у зоні провідності та валентній зоні.**

У той же час збільшується ймовірність зворотного процесу – їх зустріч і взаємна анігіляція (рекомбінація). Таким чином,

**Рекомбінацією є процес, унаслідок якого одночасно зникають вільний електрон із зони провідності та дірка – із валентної зони.**

Ілюстрацією цього визначення є рис. 10.2 а. За рекомбінації через домішкові рівні при переході електрона із домішкового рівня до валентної зони зникає дірка із валентної зони (рис. 10.2 б) або електрон – із зони провідності, який переходить на домішковий (донорний) рівень

(рис. 10.2 в). При цьому енергія рекомбінованих носіїв може витрачатися на випромінювання фотона (випромінювальна рекомбінація) або на народження фононів і нагрівання зразка (безвипромінювальна рекомбінація).

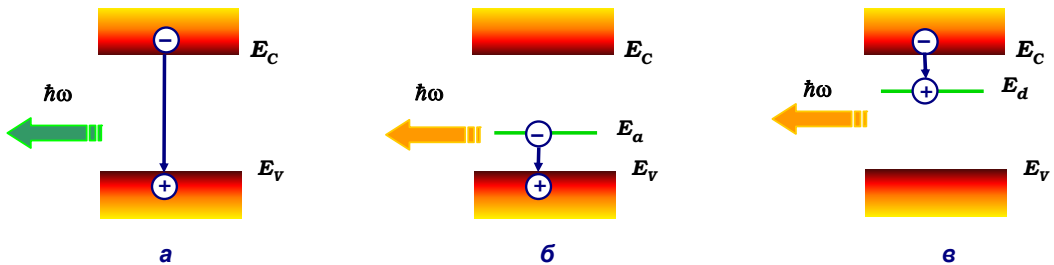


Рис. 10.2. Схема процесів випромінювальної рекомбінації носіїв типу зона-зона:  
 а – через домішковий рівень ( $\hbar\omega \leq E_g$ ); б – через акцепторний рівень ( $\hbar\omega \leq E_a - E_V$ );  
 в – через донорний рівень ( $\hbar\omega \leq E_C - E_d$ )

Процеси генерації і рекомбінації утворюють новий стаціонарний стан напівпровідника, що характеризується більшою концентрацією носіїв і не є рівноважним. Згідно із цим механізми рекомбінації зазвичай поділяють на три основних класи:

- ♦ **міжзонна рекомбінація**: вільний електрон із зони провідності рекомбінує з діркою із валентної зони за один елементарний акт;
- ♦ **рекомбінація через пастки** (домішки та дефекти): вільні електрони рекомбінують зі зв'язаними на домішках або дефектах дірками, а вільні дірки – зі зв'язаними на пастках електронами;
- ♦ **поверхнева рекомбінація** відбувається на поверхні зразка або на межі поділу гетерофазної системи. Однак поверхнева рекомбінація часто не відокремлюється в клас, оскільки також може бути як міжзонною, так і рекомбінацією через пастки.

При переході електрона із зони провідності до валентної зони внаслідок прямої рекомбінації **зона-зона** мають виконуватись закони збереження енергії та імпульсу: якщо  $E'$  та  $\mathbf{k}'$  є енергією та хвильовим вектором електрона у зоні провідності, а  $E$  та  $\mathbf{k}$  – енергією та хвильовим вектором електрона у валентній зоні, то

$$E'(\mathbf{k}') = E(\mathbf{k}) + \Delta E, \quad (10.4)$$

$$\hbar\mathbf{k}' = \hbar\mathbf{k} + \Delta\mathbf{p}. \quad (10.5)$$

У цих формулах  $\Delta E$  – енергія, що вивільняється за один акт рекомбінації, а  $\Delta\mathbf{p}$  – зміна імпульсу електрона при переході із зони провідності до валентної зони. Іншими словами, для одночасного виконання умов



(10.4) і (10.5) потрібна наявність третього агента. Якщо таким агентом є фотон, то кажуть, що відбувається випромінювальна рекомбінація. Тоді

$$\Delta E = \hbar\omega, \quad |\Delta p| = \hbar\omega/c \quad (10.6)$$

– енергія та імпульс кванта світла, що випромінюється, відповідно.

У випадку безвипромінювальної рекомбінації надлишкові енергія та імпульс можуть передаватись фононам. Існують випадки, коли дефіцит енергії та імпульсу передаються іншому вільному електрону (у напівпровіднику  $n$ -типу) або вільній дірці (у напівпровіднику  $p$ -типу). Такі процеси називаються ударною рекомбінацією (або оже-процесами). Зрозуміло, що ймовірність таких процесів значно зростає за зростання концентрації носіїв, тобто оже-рекомбінація виявляється у сильно легованих напівпровідниках. Якщо електрон і дірка утворюють як проміжний стан екситон, то кажуть що відбувається екситонна рекомбінація. Якщо ширина забороненої зони не перевищує 0,2–0,3 еВ, то вважається, що в таких речовинах основну роль відіграє міжзонна рекомбінація. Розглянемо детальніше, як виконуються закони збереження енергії та імпульсу за міжзонної рекомбінації. Як відомо, можуть існувати два типи зонної структури поблизу екстремумів енергії, які схематично подано на рис. 10.3.

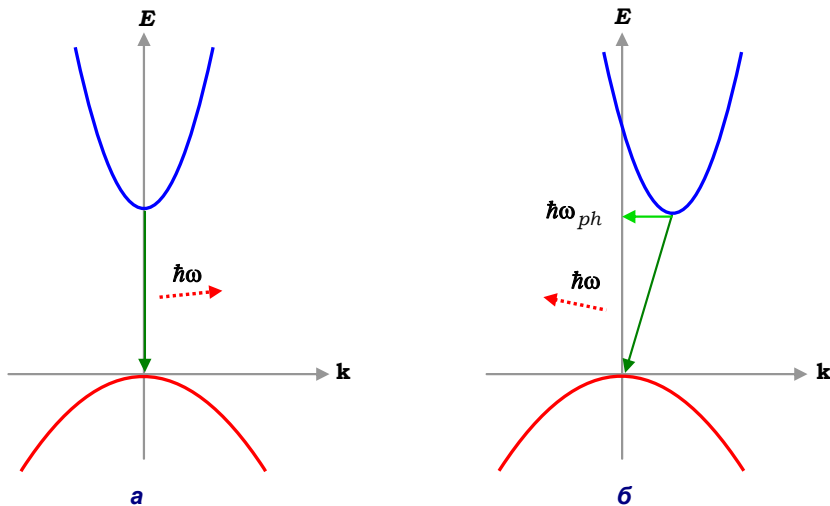


Рис. 10.3. Міжзонна випромінювальна рекомбінація у напівпровідниках: а – прямозонному; б – непрямозонному

Якщо міжзонна рекомбінація є випромінювальною, то закон збереження енергії виконується за рахунок випромінювання фотона з енергією, яка дорівнює ширині забороненої зони. Для прямозонного напівпровідника процес міжзонної рекомбінації є дуже ефективним, оскільки не виникає проблем зі збереженням імпульсу. Він виконується

автоматично, адже імпульс фотона набагато менший за імпульс електрона з відповідною енергією. Дійсно, у процесі прямої міжзонної рекомбінації у прямозонному напівпровіднику рекомбінують вільна дірка з вільним електроном, імпульси яких майже однакові (відрізняються на імпульс випроміненого фотона). У випадку непрямозонного напівпровідника для виконання законів збереження необхідний додатковий агент події рекомбінації, оскільки закон збереження імпульсу вимагає його значної зміни. Таким агентом виступає фонон. Отже, закони збереження енергії та імпульсу за міжзонної випромінювальної рекомбінації у непрямозонному напівпровіднику матимуть вигляд

$$E'(\mathbf{k}') = E(\mathbf{k}) + \Delta E, \quad \Delta E = \hbar\omega + \hbar\omega_{\text{фон.}} \quad (10.7)$$

$$\hbar\mathbf{k}' = \hbar\mathbf{k} + \Delta\mathbf{p}, \quad \Delta\mathbf{p} = \hbar\mathbf{k}_{\text{фон.}} + \hbar\mathbf{k}'_{\text{фон.}}, \quad \mathbf{k}'_{\text{фон.}} \ll \mathbf{k}'$$

Ймовірність процесу із трьома частинками є значно меншою за ймовірність двочастинкового процесу. Із цієї причини міжзонна рекомбінація проходить ефективно у прямозонних напівпровідниках і дає дуже незначний внесок до рекомбінаційних процесів у непрямозонних напівпровідниках, де важливішою є рекомбінація через домішкові центри. Якщо ширина забороненої зони перевищує 0,5 eV, то в таких речовинах рекомбінація зазвичай відбувається через локалізовані стани, які називаються рекомбінаційними пастками. Щодо рекомбінації через пастки необхідно зазначити два важливих аспекти:

- ♦ захоплюючись із зони провідності через пастку, електрон передає своєму оточенню меншу енергію, ніж якби він безпосередньо переходив до валентної зони. За таких обставин ймовірність першого процесу може бути істотно вищою, ніж другого;
- ♦ при захоплюванні носія на пастці виконання закону збереження імпульсу не є обов'язковим, оскільки носії, що потрапляють до області сильної неоднорідності поля, не знаходяться в умовах просторової однорідності, де збереження імпульсу є обов'язковим. Це значно спрощує виконання умови збереження енергії і тим самим збільшує ймовірність процесу рекомбінації через пастки.

У ролі рекомбінаційних пасток можуть виступати мілкі та глибокі рівні, при цьому процеси рекомбінації через пастки можуть відбуватися принципово по-різному. Дійсно, електронний стан на воднеподібному (мілкому) домішковому рівні характеризується досить сильною делокалізацією. Тому із принципу невизначеності імпульс такого електрона буде визначений досить точно. Цей факт на зонній діаграмі подано коротким відрізком на рівні  $E_d$  (рис. 10.4). Таким чином, процеси рекомбінації через мілкі рівні ефективно проходять у прямозонних напівпровідниках. Для глибоких рівнів ( $E_M$  на рис. 10.4), що характеризуються сильною локалізацією електронів (тобто точним визначенням координати), імпульс має широкі межі невизначеності, до яких

можуть потрапляти значення дефіциту імпульсу при переході електрона із зони провідності до валентної зони. Таким чином ймовірність процесу рекомбінації через глибокі рівні у непрямозонному напівпровіднику буде набагато більшою, ніж через мілкі. Якісні аспекти процесів генерації і рекомбінації носіїв потребують визначення поведінки кількісних характеристик таких процесів. Основними кількісними характеристиками генераційно-рекомбінаційних процесів є концентрація та час життя надлишкових нерівноважних носіїв заряду.

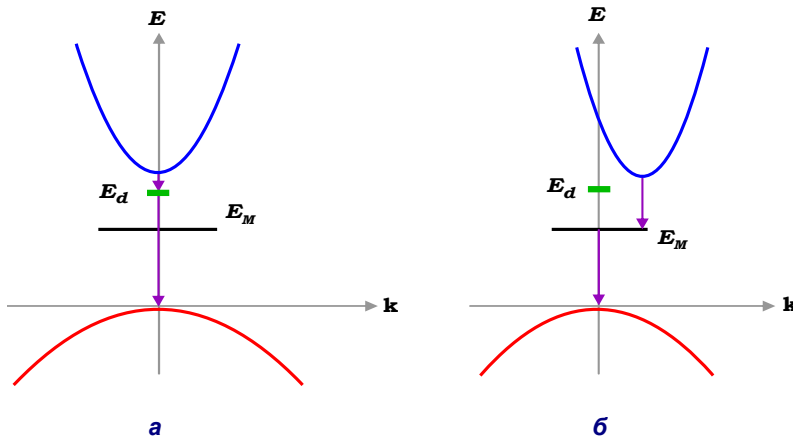


Рис. 10.4. Рекомбінація через мілкі та глибокі рівні:  
а – у прямозонному, б – у непрямозонному напівпровіднику.

Зеленим позначено делокалізовані мілкі рівні, чорним – сильно локалізовані глибокі рівні

### ■ 10.3. Кінетика генерації і рекомбінації носіїв заряду у напівпровідниках

Як зазначалося, генераційно-рекомбінаційні процеси, що можуть відбуватись у напівпровідниках, спричиняють виникнення або зникнення носіїв у валентній зоні та зоні провідності – виникає додаткова концентрація електронів і дірок, що відрізняється від рівноважної. Якщо концентрація електронів і дірок відрізняється від рівноважної, то кажуть, що в напівпровіднику є надлишкова концентрація носіїв  $\delta n(\mathbf{r}, t) = n(\mathbf{r}, t) - n_0(\mathbf{r})$ , яка може виникати, наприклад при появі у напівпровіднику додаткової пари вільних носіїв. У такому випадку відбувається *генерація носіїв заряду*, яка може бути спричинена дією зовнішніх або внутрішніх факторів. Процес генерації носіїв можна характеризувати швидкістю зміни концентрації завдяки дії зовнішніх ( $g_B$ ) і

внутрішніх ( $g_I$ ) факторів. Разом із тим у напівпровіднику можуть відбуватись і зворотні процеси зникнення носіїв заряду. У цьому випадку кажуть, що відбувається *рекомбінація носіїв заряду*. Процеси рекомбінації також можна характеризувати швидкістю зміни концентрації носіїв завдяки дії зовнішніх ( $r_E$ ) і внутрішніх ( $r_I$ ) причин. Таким чином, можна записати рівняння неперервності для електронів і дірок

$$\frac{\partial n}{\partial t} = -\operatorname{div}(\mathbf{j}_n/e_n) + G^{(n)} - R^{(n)}, \quad (10.8)$$

$$\frac{\partial p}{\partial t} = -\operatorname{div}(\mathbf{j}_p/e_p) + G^{(p)} - R^{(p)}, \quad (10.9)$$

$$\text{де} \quad G = g_E - r_E, \quad R = r_I - g_I. \quad (10.10)$$

У напівпровіднику в адіабатичних умовах, коли зовнішні фактори не впливають на генераційно-рекомбінаційні процеси (тобто коли  $g_E = r_E = 0$ ) за відсутності струму

$$\frac{\partial n}{\partial t} = -R. \quad (10.11)$$

Звідси у стаціонарних умовах, коли  $\partial n/\partial t = 0$ ,  $R = 0$  або

$$g_I = r_I, \quad (10.12)$$

тобто процеси генерації і рекомбінації взаємно компенсують один одного. Обидва процеси у цьому випадку відбуваються завдяки дії тільки внутрішніх факторів, основним з яких є взаємодія носіїв із коливаннями ґратки. Маємо рівноважні значення концентрації електронів і дірок. Якщо ж  $R \neq 0$ , рівняння (10.11) описує процес релаксації концентрації до рівноважної. Для розв'язку (10.11) припустимо, що існує деяка стала величина  $1/\tau_f$ , яка дорівнює ймовірності рекомбінації одного вільного носія за одиницю часу в одиниці об'єму, тобто за одиницю часу рекомбінують

$$R = \frac{n - n_0}{\tau_f} \quad (10.13)$$

частинок. За цього припущення (10.11) набуде вигляду

$$\frac{\partial n}{\partial t} = -\frac{n - n_0}{\tau_f}, \quad (10.14)$$

$$\text{звідки отримуємо} \quad n(t) - n_0 = [n(0) - n_0] e^{-t/\tau_f}. \quad (10.15)$$

Після припинення дії факторів, що спричиняють генерацію носіїв (при  $t = 0$ ), нерівноважний стан релаксує до рівноважного із часом  $\tau_f$ . Останній має прозорий фізичний смисл: він є часом життя нерівноважних носіїв заряду. Оскільки рівноважний стан може бути пору-

шений як для електронів, так і для дірок, то рівняння типу (10.15) мають місце як для електронів, так і для дірок. (Зрозуміло, що часи життя для електронів і дірок можуть не збігатися.)

Якщо швидкість зміни концентрації нерівноважних носіїв пропорційна до їхньої концентрації, то рекомбінація називається лінійною. За такої рекомбінації перехід носія від вільного стану до зв'язаного відбувається незалежно від наявності надлишкової концентрації носіїв протилежного знаку. Якщо ж відбувається пряме з'єднання електрона з діркою, то в одному акті рекомбінації зникають електрон разом із діркою. У такому випадку швидкості зміни концентрації електронів і дірок дорівнюють одна одній і пропорційні добутку їхніх концентрацій

$$R = \gamma(np - n_0p_0), \quad (10.16)$$

звідки 
$$\frac{\partial n}{\partial t} = \frac{\partial p}{\partial t} = -\gamma(np - n_0p_0). \quad (10.17)$$

Коефіцієнт  $\gamma$  визначає швидкість процесів генерації електрон-діркових пар і, як буде показано далі, час життя нерівноважних носіїв заряду. Вважаючи, що  $n = n_0 + \delta n$  і  $p = p_0 + \delta p$ , із (10.17) маємо

$$\frac{\partial n}{\partial t} = \frac{\partial p}{\partial t} = -\gamma(n_0\delta p + p_0\delta n + \delta n\delta p). \quad (10.18)$$

Якщо відхилення від рівноважної концентрації носіїв невелике, тобто виконується нерівність

$$\delta n\delta p \ll n_0\delta p + p_0\delta n, \quad (10.19)$$

то для електронного напівпровідника, де  $p_0 \ll n_0$ , маємо із (10.18)

$$\frac{\partial n}{\partial t} = \frac{\partial p}{\partial t} = -\gamma n_0\delta p. \quad (10.20)$$

Розв'язуючи це рівняння, отримуємо

$$\delta p(t) = \delta p(0)e^{-t/\tau_f}, \quad (10.21)$$

де 
$$\delta p(0) = p(0) - p_0, \quad \tau_f = \frac{1}{n_0\gamma}, \quad (10.22)$$

тобто надлишкова концентрація неосновних носіїв зменшується (або збільшується, якщо  $\delta p(0) < 0$ ) за експоненціальним законом із характерним часом  $\tau_f$ . Цей час є часом життя нерівноважних носіїв, визначається рівноважною концентрацією носіїв і сталою  $\gamma$ . Для визначення фізичного сенсу коефіцієнту  $\gamma$  зауважимо, що кожен нерівноважний носій заряду, наприклад електрон, що бере участь у тепловому

русі, характеризується деякою ймовірністю зустрічі із діркою та рекомбінування з нею. Очевидно, що ймовірність зустрічі електрона із діркою буде пропорційною до концентрації дірок і середньої відносної швидкості руху електронів і дірок –  $v_n$ . Якщо ввести поняття про переріз захоплювання (рекомбінації)  $q_n$  електрона, то середній час між двома можливими зустрічами електрона із діркою  $\tau_n = 1/q_n v_n p$ . Таким чином, величина  $q_n v_n p$  визначає число можливих зіткнень одного нерівноважного електрона із дірками за одиницю часу, тобто інтенсивність рекомбінації є числом актів рекомбінації (захоплювання) за одиницю часу в одиничному об'ємі напівпровідника. Домножимо це число на концентрацію електронів й усереднимо  $q_n v_n$ . У результаті отримаємо інтенсивність рекомбінації електронів  $\Delta n/\tau_n = \overline{q_n v_n} p \Delta n$ . Аналогічно можна записати ймовірність рекомбінації для дірок  $\Delta p/\tau_p = \overline{q_p v_p} n \Delta p$ . Добуток перерізу на швидкість, усереднений за всіма носіям у зоні ( $\gamma_n = \overline{q_n v_n}$  та  $\gamma_p = \overline{q_p v_p}$ ), називається коефіцієнтом рекомбінації та являє собою інтенсивність рекомбінації, коли концентрації  $n$  і  $\Delta p$  дорівнюють одиниці. Підставимо розв'язок (10.21) до (10.20). Маємо

$$\frac{\partial n}{\partial t} = -\gamma n_0 \delta p(0) e^{-t/\tau_f}. \quad (10.23)$$

Безпосереднє інтегрування цього рівняння дає

$$\delta n(t) = \gamma n_0 \delta p(0) \tau_f e^{-t/\tau_f} + C. \quad (10.24)$$

Для визначення константи інтегрування зауважимо, що у початковий момент часу

$$\delta n(0) = \delta p(0) + C. \quad (10.25)$$

З іншого боку при  $t \rightarrow \infty$   $\delta p = 0$ , тобто із (10.24) маємо  $C = 0$ . Тоді, зважаючи на визначення (10.22), отримуємо

$$\delta n(t) = \delta p(0) e^{-t/\tau_f}. \quad (10.26)$$

Таким чином, за міжзонної рекомбінації у випадку невеликих значень надлишкової концентрації носіїв відхилення від рівноважного значення концентрацій електронів і дірок однакові та змінюються згідно з експоненціальним законом з однаковим часом життя в нерівноважному стані  $\tau_f$ .

У протилежному випадку, коли надлишкова концентрація є значною, тобто

$$\delta n \gg (n_0, p_0) \quad (10.27)$$

маємо 
$$\frac{\partial n}{\partial t} = -\gamma \delta n \delta p. \quad (10.28)$$

За припущення, що  $\delta n = \delta p$ , запишемо (10.28) у вигляді

$$d\delta n \frac{1}{(\delta n)^2} = -\gamma \cdot dt. \quad (10.29)$$

Звідси отримаємо 
$$\frac{1}{\delta n(t)} = \gamma \cdot t + C. \quad (10.30)$$

При  $t = 0$   $C = 1/\delta n(0)$ . Тоді

$$\delta n(t) = \frac{\delta n(0)}{1 + \gamma \delta n(0)t}, \quad (10.31)$$

тобто за квадратичної рекомбінації велика надлишкова концентрація носіїв зменшується за гіперболічним законом. За зменшення надлишкової концентрації (коли умови  $\delta n \gg (n_0, p_0)$  порушуються) гіперболічний закон переходить в експоненціальний.

Розглянемо тепер однорідний напівпровідник, в якому існує генерація під дією зовнішнього збудження  $G = g_E - r_E$ . За відсутності струму рівняння неперервності має вигляд

$$\frac{\partial n}{\partial t} = G - R. \quad (10.32)$$

Вважатимемо, що  $G$  не залежить від часу. Тоді у стаціонарному стані

$$G = R. \quad (10.33)$$

Як було показано вище, у випадку лінійної рекомбінації  $R = (n - n_0)/\tau_f$ . Це означає, що у стаціонарному стані

$$G = \frac{n - n_0}{\tau_f}. \quad (10.34)$$

Звідси 
$$\delta n_{st} = G\tau_f, \quad (10.35)$$

і як результат знайдемо, що стаціонарний час життя нерівноважних носіїв дорівнює відношенню стаціонарного значення надлишкової концентрації до швидкості генерації:

$$\tau_f^{st} = \frac{\delta n_{st}}{G}. \quad (10.36)$$

Для квадратичної рекомбінації (при  $\delta n \gg n_0, p_0$ ) швидкість генерації носіїв у стаціонарному випадку  $G = \gamma(\delta n_{st})^2$ , звідки надлишок концентрації носіїв  $\delta n_{st} = \sqrt{G/\gamma}$ , а часом життя стаціонарної надлишкової концентрації носіїв є

$$\tau_f^{st} = \sqrt{\frac{1}{\gamma \cdot G}}. \quad (10.37)$$

У випадку, коли процес нестационарний, із (10.32) маємо

$$R = G - \frac{\partial n}{\partial t}. \quad (10.38)$$

Запишемо  $R = (n - n_0)/\tau_f$ , тоді із (10.37) отримуємо

$$\tau_f(t) = \frac{n - n_0}{G - (\partial n / \partial t)}. \quad (10.39)$$

Якщо час  $\tau_f$  не залежить від надлишкової концентрації, то у (10.39) змінні розділяються

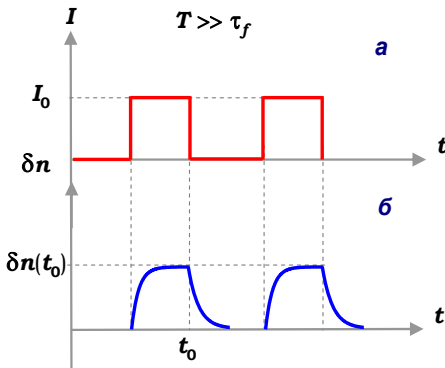
$$\frac{d(n - n_0)}{(n - n_0) - G\tau_f} = -\frac{dt}{\tau_f}, \quad (10.40)$$

і можна виконати інтегрування:  $\delta n(t) - C\tau_f = C e^{-t/\tau_f}$ . Враховуючи початкову умову  $\delta n(0) = 0$ , знайдемо константу інтегрування  $C = -G\tau_f$ . Як результат отримуємо

$$\delta n(t) = G\tau_f(1 - e^{-t/\tau_f}). \quad (10.41)$$

Нехай генерація носіїв заряду відбувається завдяки освітленню зразка світлом з енергією квантів, більшою за ширину забороненої зони напівпровідника. Тоді, якщо через деякий час після початку освітлення ( $t \gg \tau_f$ ) генерацію носіїв вимкнути, релаксація відбуватиметься за простим законом, який ми вже розглядали,

$$\delta n(t) = \delta n(t_0)e^{-t/\tau_f}, \quad (10.42)$$



**Рис. 10.3.** Релаксація надлишкової концентрації носіїв заряду  $\delta n$  (б) за генерації електронно-діркових пар прямокутними світловими імпульсами тривалістю  $T$  та інтенсивністю  $I_0$  (а)

де  $\delta n(t_0)$  просто визначається із (10.41) як насичена надлишкова концентрація носіїв. На рис. 10.3 показано, як змінюється надлишкова концентрація носіїв із часом, якщо вона утворюється прямокутними світловими імпульсами. Вважається, що тривалість імпульсів засвічування та затемнення набагато більша за час життя нерівноважних носіїв. Видно, що під дією світла відбувається інтенсивна генерація надлишкових носіїв таким чином, що до моменту вимкнення освітлення при  $t = t_0$  надлишкова концентрація носіїв за-



ряду дорівнює  $\delta n(t_0)$ . Після вимкнення освітлення процеси генерації носіїв припиняються і відбувається лише рекомбінація таким чином, що концентрація носіїв зменшується за законом (10.42) із характерним часом  $\tau_f$  (часом життя нерівноважних носіїв).

У випадку квадратичної рекомбінації рівняння неперервності дає

$$\frac{\partial n}{\partial t} = G - \gamma(\delta n)^2 \quad (10.43)$$

і допускає безпосереднє інтегрування, оскільки в ньому змінні розділяються. Маємо

$$t = \frac{1}{2\sqrt{G\gamma}} \ln \frac{\sqrt{G/\gamma} + (n - n_0)}{\sqrt{G/\gamma} - (n - n_0)} + C. \quad (10.44)$$

Для визначення константи інтегрування згадаймо, що оскільки при  $t = 0$  концентрація носіїв  $n = n_0$ , то  $C = 0$ . Таким чином із (10.44) отримуємо

$$\frac{\sqrt{G/\gamma} + (n - n_0)}{\sqrt{G/\gamma} - (n - n_0)} = e^{2\sqrt{G\gamma}t},$$

звідки

$$\delta n(t) = \sqrt{\frac{G}{\gamma}} \frac{e^{2\sqrt{G\gamma}t} - 1}{e^{2\sqrt{G\gamma}t} + 1} = \sqrt{\frac{G}{\gamma}} \cdot \text{th} \sqrt{G\gamma} \cdot t. \quad (10.45)$$

Для малих часів, коли  $t \ll 1/\sqrt{G\gamma}$ , отримуємо

$$\delta n \approx \sqrt{G/\gamma} \cdot \sqrt{G\gamma}t = Gt. \quad (10.46)$$

У протилежному випадку великих часів ( $t \gg 1/\sqrt{G\gamma}$ ) маємо

$$\delta n \approx \sqrt{G/\gamma} = \delta n_{st}. \quad (10.47)$$

Таким чином, за квадратичної рекомбінації носіїв заряду у напівпровіднику, якщо збудження, що викликає генерацію носіїв, вимкнута в деякий момент часу у стаціонарному стані, то релаксація надлишкової концентрації відбуватиметься згідно із гіперболічним законом (10.31), де початковим значенням надлишкової концентрації буде її значення у стаціонарному стані  $\delta n(0) = \sqrt{G/\gamma}$

$$\delta n(t) = \frac{\delta n_{st}}{1 + \gamma \delta n_{st} t}. \quad (10.48)$$

Поведінку  $\delta n(t)$  за квадратичної рекомбінації подано на рис. 10.6. Генерація носіїв вимкнена в момент часу  $t_0$ , коли надлишкова концентрація носіїв досягла її значення у стаціонарному стані. Таким чином бачимо, що чим меншою є величина  $\gamma$ , тим швидше зростає надлишкова концентрація носіїв і тим повільніше ця концентрація релаксує до рівноважного значення при вимиканні генерації.

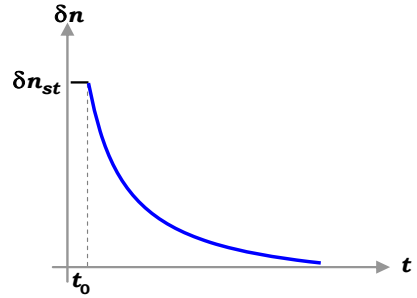


Рис. 10.6. Релаксація надлишкової концентрації носіїв до рівноважної в умовах квадратичної рекомбінації

### 10.3.1. Міжзонна випромінювальна рекомбінація

Розглянемо освітлений напівпровідник у стані термодинамічної рівноваги: є електрони із рівноважною концентрацією  $n_0$  і дірки з рівноважною концентрацією  $p_0$  та чорне (рівноважне) випромінювання. Згідно із принципом детальної рівноваги темп рівноважної випромінювальної рекомбінації  $r_0$  має дорівнювати темпу оптичної генерації пар  $g_0$ , що зумовлена поглинанням рівноважних фотонів, тобто

$$r_0 = g_0. \quad (10.49)$$

З іншого боку, оскільки процес прямої міжзонної рекомбінації є квадратичним, то можна записати

$$\gamma_e n_0 p_0 = \gamma_e n_i^2. \quad (10.50)$$

Тут коефіцієнт  $\gamma_e$  описує ймовірність захоплення носіїв із випромінюванням фотона та називається коефіцієнтом випромінювальної рекомбінації. Зрозуміло, що

$$\gamma_e = r_0 / n_i^2. \quad (10.51)$$

Оскільки у зоні провідності та валентній зоні нерівноважні й рівноважні носії не можна розрізнити, то в нерівноважних умовах випромінювальна рекомбінація визначається тим самим коефіцієнтом, що і в рівноважних. Тоді швидкістю випромінювальної рекомбінації в нерівноважних умовах є

$$r = \gamma_e n p = \frac{n p}{n_i^2} r_0. \quad (10.52)$$

Враховуючи той факт, що ці часи для електронів і дірок мають дорівнювати одне одному, введемо час життя нерівноважних носіїв як

$$\tau_n = \tau_p = \frac{\delta n}{R}, \quad R = r - g_0 \quad (10.53)$$

або, беручи до уваги (10.52),

$$\tau_n = \tau_p = \frac{\delta n n_i^2}{r_0 (np - n_i^2)}. \quad (10.54)$$

Врахуємо тепер, що  $n = n_0 + \delta n$ ,  $p = p_0 + \delta p$  та  $n_i^2 = n_0 p_0$  та із (10.54) отримаємо

$$\tau_n = \tau_p = \frac{1}{r_0} \frac{n_i^2}{n_0 (\delta p / \delta n) + p_0 + \delta p}. \quad (10.55)$$

Якщо і генерація, і рекомбінація носіїв відбуваються тільки згідно із прямим міжзонним механізмом (без участі центрів рекомбінації будь-якої природи), то  $\delta n = \delta p$  і (10.55) набуде такого вигляду

$$\tau_n = \tau_p = \frac{1}{r_0} \frac{n_i^2}{n_0 + p_0 + \delta n}. \quad (10.56)$$

У випадку слабого рівня інжекції носіїв  $\delta n \ll n_0 + p_0$ , із (10.56) знайдемо

$$\tau_n = \frac{1}{r_0} \frac{n_i^2}{n_0 + p_0}. \quad (10.57)$$

Звідси отримуємо:

- ♦ для напівпровідника *n*-типу ( $n_0 \gg p_0$ )

$$\tau_n = \frac{1}{r_0} \frac{n_i^2}{n_0} = \frac{1}{\gamma_e n_0}, \quad (10.58)$$

- ♦ для напівпровідника *p*-типу

$$\tau_p = \frac{1}{\gamma_e p_0}. \quad (10.59)$$

У випадку великого рівня інжекції ( $\delta n \gg n_0, p_0$ ) часи релаксації носіїв у напівпровідниках *n*- та *p*-типів збігаються

$$\tau_p = \tau_n = \frac{1}{\gamma_e \delta n}. \quad (10.60)$$

Таким чином, можна зробити висновки, що

- ♦ тільки за малих рівнів інжекції випромінювальний час релаксації (час життя нерівноважних носіїв) залежить від ступеня легування напівпровідника (тобто, від  $n_0$  та  $p_0$ );

- ♦ зі збільшенням ступеня легування величина випромінювального часу життя зменшується;

- ♦ у власному напівпровіднику, де  $n_0 = p_0 = n_i$ , час життя нерівноважних носіїв пропорційний до рівноважної концентрації носіїв

$$\tau_i = n_i / 2r_0. \quad (10.61)$$

Усі ці феноменологічні співвідношення набувають замкненої форми, якщо в рамках деякої моделі обчислити коефіцієнт випромінювальної рекомбінації. Для цього припустимо, що при поглинанні кожного фотона з енергією  $E_g = \hbar\omega$  у кристалі утворюється електрон-діркова пара. У такому випадку кажуть, що квантовий вихід генерації носіїв світлом дорівнює одиниці. Це означає, що швидкість поглинання фотонів (кількість фотонів, що поглинаються одиницею об'єму кристала за одиницю часу) дорівнює швидкості генерації електронно-діркових пар  $g$ . В умовах термодинамічної рівноваги швидкості генерації і рекомбінації дорівнюють одна одній. Величина  $g_0$  є числом фотонів, що поглинаються одиницею об'єму напівпровідника за одиницю часу. Вона визначається як інтеграл

$$g_0 = \int_0^{\infty} f(\omega)N(\omega)d\omega, \quad (10.62)$$

де  $N(\omega)$  – щільність фотонів в інтервалі частот  $d\omega$ , а функція  $f(\omega)$  визначає ймовірність поглинання фотона із частотою  $\omega$  за одиницю часу. Дійсно, оскільки функція розподілу частинок  $f$  визначається як ймовірність попадання однієї частинки у даний об'єм фазового простору, то для визначеного одиничного об'єму ця функція визначає ймовірність однієї частинці мати енергію, що характеризується частотою  $\omega$ . Таким чином за функцію  $f(\omega)$  братимемо модифіковану функцію розподілу Бозе–Ейнштейна

$$f(\hbar\omega) = \frac{g(\omega)}{e^{\hbar\omega/kT} - 1}, \quad (10.63)$$

де  $g(\omega)$  – ймовірність поглинання фотона даною системою, що визначається коефіцієнтом поглинання  $\alpha$

$$g(\omega) = \alpha(c/\tilde{n}). \quad (10.64)$$

Тут  $c$  – швидкість світла у вакуумі,  $\tilde{n}$  – коефіцієнт заломлення світла у даному кристалі. Знайдемо тепер функцію  $N(\omega)$ . Усі фотони в енергетичному інтервалі від  $\hbar\omega$  до  $\hbar(\omega + d\omega)$ , в імпульсному просторі знаходяться між сферами радіусів  $p$  та  $p + dp$ . Об'єм такого сферичного шару

$$V = 4\pi p^2 dp. \quad (10.65)$$

З іншого боку, імпульсом фотона є

$$p = \frac{\hbar\omega}{v} = \frac{\hbar\omega\tilde{n}}{c}. \quad (10.66)$$

Тоді

$$V = 4\pi \frac{\hbar^3 \tilde{n}^3}{c^3} \omega^2 d\omega. \quad (10.67)$$

Оскільки  $N(\omega)d\omega = V/8\pi^3\hbar^3$ , то остаточно (із врахуванням двох можливих поляризацій світла) маємо

$$N(\omega)d\omega = \frac{n^3}{\pi^2c^3}\omega^2d\omega. \quad (10.68)$$

Після підстановки (10.63), (10.64) та (10.68) до (10.62), отримаємо

$$g_0 = \frac{1}{\pi^2c^2} \int_0^\infty \frac{\tilde{n}^2\alpha}{\exp(\hbar\omega/kT)-1} \omega^2 d\omega. \quad (10.69)$$

Звідси для коефіцієнта випромінювальної рекомбінації знайдемо

$$\gamma_e = \frac{1}{\pi^2c^2n_i^2} J, \quad (10.70)$$

де

$$J = \int_0^\infty \frac{\tilde{n}^2\alpha}{\exp(\hbar\omega/kT)-1} \omega^2 d\omega.$$

Тепер можна записати вирази для часу життя нерівноважних носіїв у напівпровіднику, що визначається міжзонною випромінювальною рекомбінацією:

♦ у загальному випадку слабого рівня інжекції

$$\tau_n = \tau_p = \frac{1}{J} \cdot \frac{\pi^2c^2n_i^2}{n_0 + p_0}. \quad (10.71)$$

♦ у домішковому напівпровіднику  $n$ -типу за слабого рівня інжекції

$$\tau_n = \pi^2c^2p_0 \frac{1}{J}, \quad (10.72)$$

♦ у домішковому напівпровіднику  $p$ -типу за слабого рівня інжекції

$$\tau_p = \pi^2c^2n_0 \frac{1}{J}, \quad (10.73)$$

тобто в легovanому напівпровіднику із досить великим рівнем легування час життя нерівноважних носіїв визначається рівноважною концентрацією неосновних носіїв. У випадку великих рівнів інжекції час життя нерівноважних носіїв визначається універсальною формулою

$$\tau_n = \tau_p = \frac{1}{J} \cdot \frac{\pi^2c^2}{\delta n}. \quad (10.74)$$

Для визначення параметра  $J$  необхідно знати коефіцієнт поглинання світла напівпровідником. Для прямих міжзонних переходів він визначається щільністю станів носіїв

$$\alpha(E) = C \cdot \frac{2\pi}{\hbar^3} \left[ \frac{m_i}{m_i + m_f^{(a)}(E)} \right]^{3/2} \left[ m_f^{(a)}(E) \right]^{1/2} E^{1/2}, \quad (10.75)$$

де  $m_i$  та  $m_f^{(a)}$  – ефективні маси щільності станів вихідної та кінцевої зон, відповідно;  $C$  – деякий коефіцієнт. У випадку непрямих переходів коефіцієнт поглинання світла

$$\alpha(E) = A \left[ \frac{(E - (E_g - k\Theta))^2}{1 - \exp(-\Theta/T)} - \frac{(E - (E_g + k\Theta))^2}{1 - \exp(\Theta/T)} \right]. \quad (10.76)$$

Стала  $A$  для германію складає величину  $2600 \text{ eV}^{-2}\text{cm}^{-1}$ , для кремнію –  $3600 \text{ eV}^{-2}\text{cm}^{-1}$ ,  $\Theta$  – температура Дебая. Використовуючи ці вирази, можна отримати для коефіцієнта випромінювальної рекомбінації:

♦ **за прямих переходів**

$$\gamma_e^{(d)} = 0,58 \cdot 10^{-12} n \left( \frac{1}{m_n^* + m_p^*} \right)^{3/2} (1 + 1/m_n^* + 1/m_p^*) \left( \frac{T}{300} \right)^{3/2} E_g^2, \text{ (cm}^3\text{s}^{-1}\text{)}. \quad (10.77)$$

♦ **за непрямих переходів**

$$\gamma_e^{(i)} = 2,17 \cdot 10^{-20} A n^2 \left( \frac{1}{m_n^* m_p^*} \right) \text{cth} \frac{\Theta}{2T} E_g^2, \text{ (cm}^3\text{s}^{-1}\text{)}. \quad (10.78)$$

Як видно, завжди  $\gamma_e^{(d)} \gg \gamma_e^{(i)}$ , тобто випромінювальна міжзонна рекомбінація ефективніше відбувається у прямозонних напівпровідниках.

### 10.3.2. Міжзонна ударна рекомбінація

Процес міжзонної ударної рекомбінації вимагає участі трьох частинок: крім рекомбінуючої електронно-діркової пари потрібний ще електрон або дірка, яким передаються енергія та імпульс, вивільнені за рекомбінації. Таким чином цей процес характеризується двома пере-різами захоплювання – ймовірністю захоплювання при передачі енергії електрону  $\gamma_{nn}$  та ймовірністю захоплювання при передачі енергії дірці  $\gamma_{pp}$ . Необхідно ще ввести швидкість захоплювання  $r$ , яка тепер буде пропорційна не квадрату, а кубу концентрації носіїв. Таким чином, можна записати швидкість рекомбінації при передачі енергії рекомбінуючої пари електрону

$$r_n = \gamma_{nn} n^2 p. \quad (10.79)$$

Якщо енергія рекомбінуючої пари передається дірці, то швидкість рекомбінації

$$r_p = \gamma_{pp} n p^2. \quad (10.80)$$

Повна швидкість рекомбінації

$$r = r_n + r_p. \quad (10.81)$$

Позначимо ці швидкості процесів в умовах термодинамічної рівноваги через  $r_{n0}$  та  $r_{p0}$  і знову вважатимемо, що ймовірності захоплювання для рівноважних і нерівноважних носіїв однакові. Тоді маємо

$$\gamma_{nn} = \frac{r_{n0}}{n_0^2 p_0}, \quad \gamma_{pp} = \frac{r_{p0}}{p_0^2 n_0}. \quad (10.82)$$

Повна швидкість рекомбінації, згідно із (10.81), визначатиметься

$$r = (\gamma_{nn} n + \gamma_{pp} p) np. \quad (10.83)$$

Звідси

$$r = \left( \frac{r_{n0}}{n_0^2 p_0} n + \frac{r_{p0}}{p_0^2 n_0} p \right) np. \quad (10.84)$$

Вважатимемо, що процесу ударної рекомбінації відповідає зворотний процес ударної іонізації, який спричиняє генерацію електронно-діркових пар. Швидкість цього процесу пропорційна концентрації швидких електронів

$$g_n = \gamma'_{nn} n, \quad (10.85)$$

і, аналогічно, дірок

$$g_p = \gamma'_{pp} p. \quad (10.86)$$

Вважаючи ймовірність ударної іонізації однаковою для рівноважних і нерівноважних носіїв, можна записати

$$\gamma'_{nn} = \frac{g_{n0}}{n_0}, \quad \gamma'_{pp} = \frac{g_{p0}}{p_0}. \quad (10.87)$$

У рівноважних умовах процеси генерації і рекомбінації відбуваються з однаковою швидкістю, тобто

$$r_{n0} = g_{n0}, \quad r_{p0} = g_{p0}. \quad (10.88)$$

Тоді, зважаючи на співвідношення (10.84), (10.88) і той факт, що  $g = g_n + g_p$ , отримаємо для повної швидкості рекомбінації

$$R = r - g = \frac{np - n_i^2}{n_i^2} \left( r_{n0} \frac{n}{n_0} + r_{p0} \frac{p}{p_0} \right). \quad (10.89)$$

Оскільки час життя нерівноважних носіїв у даному процесі  $\tau = \delta n / R$ , то із (10.89) маємо

$$\tau = \frac{\delta n n_i^2}{(np - n_i^2) \cdot \left( r_{n0} \frac{n}{n_0} + r_{p0} \frac{p}{p_0} \right)}. \quad (10.90)$$

Далі перетворюємо вираз  $np - n_i^2$ , із врахуванням того, що  $n_i^2 = n_0 p_0$  і  $\delta n = \delta p$ . Тоді  $n - n_0 = p - p_0$ . Звідси  $p = n - n_0 + p_0$  і можна записати:

$np - n_i^2 = n(n - n_0 + p_0) - n_0 n_0 = n(n - n_0) + p_0(n - n_0) = \delta n(n + p_0)$ . Тепер  $n = p - p_0 + n_0$  і  $np - n_i^2 = \delta n(n + p_0) = \delta n(p - p_0 + n_0 + p_0) = \delta n(n_0 + p_0 + \delta n)$ .

У результаті (10.90) можна записати у вигляді

$$\tau = \frac{n_i^2}{(n_0 + p_0 + \delta n) \cdot (r_{n_0}(n/n_0) + r_{p_0}(p/p_0))}. \quad (10.91)$$

Введемо позначення 
$$\beta = \frac{r_{p_0} n_0}{r_{n_0} p_0}. \quad (10.92)$$

Тоді рівняння (10.91) перетворюється на

$$\tau = \frac{n_i^2}{(n_0 + p_0 + \delta n) \cdot (r_{n_0}/n_0)(n_0 + \delta n + \beta(p_0 + \delta n))}. \quad (10.93)$$

У власному напівпровіднику ( $n_0 = p_0 = n_i$ ) час життя нерівноважних носіїв за низького рівня інжекції ( $\delta n \ll p_0, n_0$ ) та умови, що за релаксації енергія в основному передається електронам (зазвичай завдяки малій їхній ефективній масі), тобто

$$\beta(p_0 + \delta n) \ll n_0 + \delta n, \quad (10.94)$$

як випливає із (10.93), визначається формулою

$$\tau_i = \frac{n_0}{2r_{n_0}}. \quad (10.95)$$

Тоді остаточно можна записати для часу життя нерівноважних носіїв за міжзонної ударної рекомбінації

$$\tau = \frac{2\tau_i n_i^2}{(n_0 + p_0 + \delta n)(n_0 + \delta n + \beta(p_0 + \delta n))}. \quad (10.96)$$

За низьких рівнів інжекції маємо

$$\tau = \frac{2\tau_i n_i^2}{(n_0 + p_0)(n_0 + \beta p_0)}, \quad (10.97)$$

звідси:

♦ **для напівпровідників *n*-типу**

$$\tau = \frac{2\tau_i n_i^2}{n_0^2} = \frac{2\tau_i p_0}{n_0} \ll \tau_i; \quad (10.98)$$

♦ **для напівпровідників *p*-типу**

$$\tau = \frac{2\tau_i n_i^2}{\beta p_0^2}. \quad (10.99)$$

На відміну від випромінювальної, за ударної рекомбінації у випадку низьких рівнів інжекції час життя нерівноважних носіїв обернено



пропорційний квадрату концентрації основних носіїв заряду. За високих рівнів інжекції ( $\delta n \gg n_0, p_0$ ) із (10.96) випливає, що час життя нерівноважних носіїв визначається формулою

$$\tau = 2\tau_i \left( \frac{n_i}{\delta n} \right)^2 \frac{1}{1+\beta}. \quad (10.100)$$

Оскільки зазвичай можна вважати, що  $\beta \ll 1$ , то із (10.100) отримуємо, що час життя нерівноважних носіїв, які виникають унаслідок процесів ударної іонізації за високих рівнів інжекції, пропорційний квадрату концентрації носіїв і обернено пропорційний рівню інжекції

$$\tau = 2\tau_i \left( \frac{n_i}{\delta n} \right)^2. \quad (10.101)$$

У всіх аналізованих в цьому підрозділі формулах характеристикою напівпровідника як матеріалу є величини швидкості захоплювання  $r_{n0}$  та  $r_{p0}$ , що визначаються через ймовірності захоплювання  $\gamma_{nn}$  та  $\gamma_{pp}$ . Мікроскопічний розгляд дає для цих величин вираз

$$\gamma_{nn} = \frac{0,2 m^* e^4}{\sqrt{\pi} \hbar^3} \left( \frac{kT}{W_i} \right)^{3/2} N_c N_v \exp\left( -\frac{E_g + W_i}{kT} \right), \quad (10.102)$$

де  $W_i$  – параметр розмірності енергії, що залежить від деталей зонної структури напівпровідника. Зазвичай приймають, що  $W_i \approx 3E_g/2$ . Таким чином, із (10.102) можна зробити висновок: у матеріалах із вузькою забороненою зоною ймовірність ударної рекомбінації може бути помітною величиною, що швидко зростає за підвищення температури.

### 10.3.3. Рекомбінація через однозарядні локальні центри (за Шоклі-Рідом)

Припустимо, що в напівпровіднику наявні дефекти, рівні енергії  $E_t$  яких лежать глибоко у забороненій зоні, а їхньою концентрацією є  $N_t$ . Розглянемо процеси захоплювання носіїв такими дефектами та спустошення дефектів за втрати ними захопленого носія:

- ♦ **нейтральний дефект захоплює вільний електрон**, і за деякий час негативно заряджений дефект віддає електрон у зону провідності. Дефект, який захоплює вільні електрони із зони провідності та згодом, за деякий час, повертає їх до зони, називається **пасткою захоплювання електрона**;

- ♦ **нейтральний дефект захоплює вільну дірку** (віддає електрон у валентну зону) і за деякий час цей позитивно заряджений дефект захоп-

лює електрон із валентної зони, тим самим вивільнюючи дірку. Такий дефект називається **пасткою захоплення дірки**;

- ♦ захопивши електрон із зони провідності, **негативно заряджений дефект захоплює дірку** (віддає захоплений електрон у валентну зону) перетворюючись на нейтральний: відбувається процес рекомбінації електрона та дірки на дефекті;

- ♦ захопивши вільну дірку, **позитивно заряджений дефект захоплює вільний електрон**, тим самим стає нейтральним: відбувається процес рекомбінації електрона та дірки на дефекті.

Характерною рисою пасток захоплення є те, що вони в основному взаємодіють тільки з однією зоною – валентною (пастка захоплення дірки) або провідності (пастка захоплення електрона). Для того, щоб дефект спричинив рекомбінацію пари вільних носіїв, він має взаємодіяти з обома зонами. Розглянемо простий випадок рекомбінації через один тип локальних центрів. Ймовірність процесу 1 (рис. 10.7) буде пропорційна кількості електронів у зоні провідності  $n$  і числу вільних від електронів рівнів на дефектах  $N_t(1-f)$ , де  $f$  – ступінь заповнення рекомбінаційних центрів електронами. З іншого боку, ймовірність оберненого процесу 1' (ге-

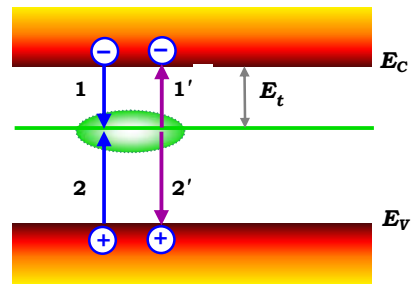


Рис. 10.7. Процес рекомбінації електрона та дірки на дефекті

нератії вільного електрона) пропорційна кількості електронів на рівні  $E_t - N_t f$ . У результаті швидкість рекомбінації електронів (тобто швидкість зміни концентрації електронів у зоні провідності)

$$R = -\frac{dn}{dt} = \gamma_n N_t n(1-f) - \beta_n N_t f, \quad (10.103)$$

де  $\gamma_n$  та  $\beta_n$  – деякі коефіцієнти. В умовах термодинамічної рівноваги у зоні провідності встановлюється стала у часі рівноважна концентрація електронів, яка описується рівноважною функцією розподілу  $f_0$ . Тоді з умови  $dn/dt = 0$  маємо

$$\beta_n = \gamma_n n_0 \frac{1-f_0}{f_0}. \quad (10.104)$$

Нехай ступінь заповнення рекомбінаційних рівнів описується рівноважною функцією розподілу Фермі  $f_0 = [\exp(-(E_t - E_f)/kT) - 1]^{-1}$ , а концентрація електронів у зоні провідності – розподілом Больцмана (напівпровідник невідроджений). Якщо енергію відраховувати від дна зони провідності, то концентрація електронів у зоні провідності описуватиметься формулою

$$n_0 = N_c e^{E_F/kT}.$$

Врахуємо тепер, що  $1 - f_0 = \exp(-(E_t - E_f)/kT)[\exp(-(E_t - E_f)/kT) + 1]^{-1}$  і знайдемо що

$$\frac{1 - f_0}{f_0} = e^{-\frac{E_t - E_F}{kT}}. \quad (10.105)$$

Підставимо (10.105) до (10.104) та отримаємо

$$\beta_n = \gamma_n N_c e^{\frac{E_F}{kT}} e^{-\frac{E_t + E_F}{kT}} = \gamma_n N_c e^{-\frac{E_t}{kT}}. \quad (10.106)$$

Для спрощення подальших записів введемо позначення

$$n_1 = N_c e^{-\frac{E_t}{kT}}, \quad (10.107)$$

що має сенс концентрації електронів у напівпровіднику, в якого рівень Фермі дорівнює енергії центрів рекомбінації  $E_F = -E_t$ . Тоді

$$\beta_n = \gamma_n n_1. \quad (10.108)$$

Оскільки повний процес рекомбінації відбувається за взаємодії локальних рівнів як із зоною провідності, так і з валентною зоною, зручно переписати рівняння балансу (10.103) для рекомбінаційних рівнів. Враховуючи (10.108), із (10.103) отримаємо для нерівноважного випадку

$$R_n = -\gamma_n N_t [n(1 - f) - n_1 f]. \quad (10.109)$$

Розглянемо тепер процес рекомбінації дірок: ймовірність захоплення дірок на рекомбінаційні центри (процес 2, рис. 10.7) пропорційна добутку кількості дірок у валентній зоні  $p$  і числу рівнів, вільних від дірок (зайнятих електронами  $N_t f$ ). Число зворотних переходів (процес 2') пропорційне кількості дірок на рекомбінаційних рівнях  $N_t (1 - f)$ .

Таким чином темп зміни концентрації дірок у валентній зоні опишуватиметься рівнянням

$$R_p = -\frac{dp}{dt} = \gamma_p N_t p f - \beta_p N_t (1 - f). \quad (10.110)$$

Аналогічно тому, як було зроблено для процесів 1 та 1', з умов рівноваги  $dp/dt = 0$ ,  $p = p_0$ ,  $f = f_0$ , отримаємо

$$\beta_p = \gamma_p p_1, \quad (10.111)$$

де

$$p_1 = N_v e^{-\frac{E_g - E_t}{kT}} \quad (10.112)$$

– концентрація дірок, за якої рівень Фермі збігається з енергетичним рівнем рекомбінаційних центрів. Тоді, аналогічно до (10.109), запишемо для швидкості зміни концентрації дірок на рекомбінаційних рівнях

$$R_p = \frac{dp}{dt} = -\gamma_p N_t [pf - p_1(1-f)]. \quad (10.113)$$

У стаціонарному випадку число електронів, що потрапили на рівень  $E_t$  і зникли звідти за одиницю часу, однакове, тобто має виконуватись рівняння балансу

$$\frac{dn}{dt} = \frac{dp}{dt}, \quad (10.114)$$

з якого отримуємо рівняння для визначення нерівноважної функції розподілу

$$\gamma_p [pf - p_1(1-f)] = \gamma_n [n(1-f) - n_1 f],$$

розв'язок якого дає

$$f = \frac{\gamma_n n + \gamma_p p_1}{\gamma_n (n + n_1) + \gamma_p (p + p_1)}. \quad (10.115)$$

Для визначення часу життя нерівноважних електронів у зоні провідності

$$\tau = \delta n / R_n^{(C)} \quad (10.116)$$

підставимо функцію розподілу (10.115) до (10.109) і замінимо  $R_n^{(C)} = -R_n$ . Маємо  $R_n^{(C)} = \gamma_n N_t [n - (n + n_1)f]$ . Зважаючи на (10.115), обчислимо тепер дужку  $n - (n + n_1)f = \gamma_p (np - n_1 p_1) / (\gamma_n (n + n_1) + \gamma_p (p + p_1))$ . Таким чином

$$R_n^{(C)} = \gamma_n N_t \frac{\gamma_p (np - n_1 p_1)}{\gamma_n (n + n_1) + \gamma_p (p + p_1)}.$$

Враховуючи, що добуток  $n_1 p_1$  записано для рівноважного випадку, тобто він дорівнює  $n_0 p_0$ , а також, що  $n = n_0 + \delta n$  і  $\delta p = \delta n$ , знайдемо для часу релаксації

$$\tau = \frac{n_0 + n_1 + \delta n}{\gamma_p N_t (n_0 + p_0 + \delta n)} + \frac{p_0 + p_1 + \delta n}{\gamma_n N_t (n_0 + p_0 + \delta n)}. \quad (10.117)$$

Для з'ясування фізичного сенсу коефіцієнтів  $\gamma_n$  та  $\gamma_p$  розглянемо донорний напівпровідник ( $n_0 \gg p_0$ ) із незначною кількістю центрів рекомбінації, рівень яких лежить у верхній частині забороненої зони ( $p_1 \ll n_1$ ) в умовах слабкої інжекції ( $\delta n \ll n_0, p_0$ ). У цих умовах (10.117) показує, що в легovanому напівпровіднику  $n$ -типу час життя нерівноважних носіїв визначається тільки концентрацією рекомбінаційних центрів та ймовірністю захоплювання дірки центром. За збільшення концентрації мілких донорів час життя зменшується та прямує до величини  $\tau_{p_0}$  – часу релаксації у сильно легovanому напівпровіднику  $n$ -типу.

$$\tau = \frac{1}{\gamma_p N_t} \left( 1 + \frac{n_1}{n_0} \right) + \frac{p_0}{\gamma_n N_t n_0} \approx \frac{1}{\gamma_p N_t} \left( 1 + \frac{n_1}{n_0} \right) = \tau_{p0}. \quad (10.118)$$

У випадку, коли енергетичний рівень пасток знаходиться в нижній частині забороненої зони ( $n_1 \ll p_1$ ), замість (10.118) маємо за зростання концентрації донорних мілких рівнів, що

$$\tau = \frac{1}{\gamma_p N_t} + \frac{p_1}{\gamma_n N_t n_0} \approx \frac{1}{\gamma_p N_t} = \tau_{p0}. \quad (10.119)$$

Аналогічно в легованому кристалі  $p$ -типу час життя нерівноважних носіїв визначається тільки концентрацією рекомбінаційних центрів та ймовірністю захоплення центром електрона

$$\tau_{n0} = \frac{1}{\gamma_n N_t}. \quad (10.120)$$

З використанням цих позначень остаточно отримуємо так звану формулу Шоклі-Ріда, яка пов'язує час життя нерівноважних носіїв з їх рівноважною концентрацією, рівнем інжекції  $\delta n$  і часом життя неосновних носіїв. Крім того, оскільки  $n_1$  та  $p_1$  визначаються енергією  $E_t$ , то ця формула визначає залежність часу життя від типу рекомбінаційних центрів

$$\tau = \tau_{p0} \frac{n_0 + n_1 + \delta n}{n_0 + p_0 + \delta n} + \tau_{n0} \frac{p_0 + p_1 + \delta n}{n_0 + p_0 + \delta n}. \quad (10.121)$$

За низького рівня інжекції формула Шоклі-Ріда переходить у

$$\tau_0 = \tau_{p0} \frac{n_0 + n_1}{n_0 + p_0} + \tau_{n0} \frac{p_0 + p_1}{n_0 + p_0}. \quad (10.122)$$

Зауважимо, що в такому наближенні час життя нерівноважних носіїв визначається винятково термодинамічно рівноважними концентраціями електронів  $n_0$  і дірок  $p_0$ . При цьому у напівпровіднику  $n$ -типу часом життя за низьких рівнів інжекції є

$$\tau_0 = \tau_{p0} \left( 1 + \frac{n_1}{n_0} \right). \quad (1.123)$$

Припустимо, що величина  $\tau_{p0}$  не залежить від температури. Тоді для аналізу залежності часу  $\tau$  від температури згадаємо, що за зміни температури в напівпровіднику відбуваються як зміна концентрації носіїв заряду, так і зміна положення рівня Фермі щодо рівнів мілких донорів (акцепторів) і глибоких рівнів рекомбінаційних центрів.

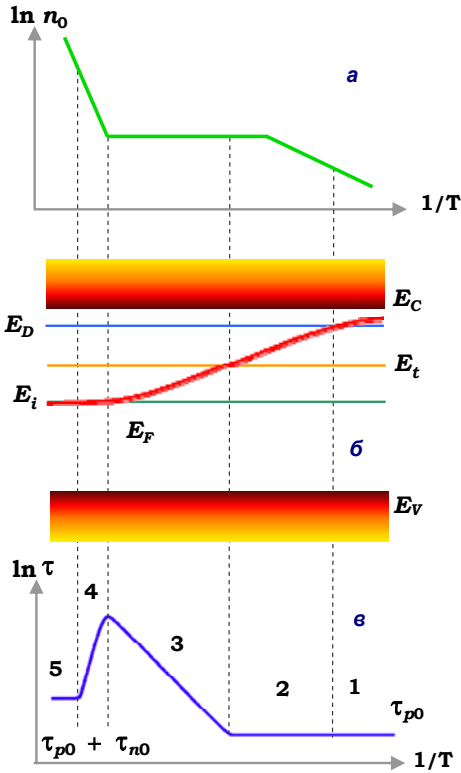


Рис. 10.8. Температурна залежність часу життя носіїв заряду за рекомбінації через пастки для напівпровідника *n*-типу (в). Для порівняння подано залежність концентрації носіїв (а) та положення рівня Фермі (б) від температури

Для порівняння подано залежність концентрації носіїв (а) та положення рівня Фермі (б) від температури

Розглянемо для визначеності напівпровідник *n*-типу. Рівень Фермі лежить між рівнем донорної домішки та дном зони провідності в області низьких температур (рис. 10.8). Оскільки тут мілкі рівні іонізуються, а глибокі (пастки з енергією  $E_t$ ) – не дають помітної залежності від температури, то  $n_0 \gg n_1$ ,  $n_0 \gg p_0$  і  $n_0 \gg p_1$ . У такому разі із (10.123) випливає, що

$$\tau = \tau_{p0}, \quad (10.124)$$

тобто за низьких температур час життя нерівноважних носіїв визначається часом життя неосновних носіїв і не залежить від температури (область 1 на рисунку). За подальшого зростання температури рівень Фермі переходить в область нижче рівня мілких донорів і вище глибоких рекомбінаційних центрів, тобто знаходиться в області повної іонізації донорних домішок і може знизитись до рівня глибоких пасток. Нехтуючи залежністю  $n_1$  від температури, із (10.123) маємо

$$\Delta\tau = \frac{n_1}{N_C} e^{-\frac{E_F}{kT}}, \quad (10.125)$$

де  $\Delta\tau = (\tau - \tau_{p0})/\tau_{p0}$ . Оскільки за низьких температур у напівпровіднику рівень Фермі збігається з енергією домішки, то із (10.125) випливає

$$\ln \Delta\tau = -\frac{1}{kT} E_F - \ln \frac{N_C}{n_1} = \frac{E_D}{k} \cdot \frac{1}{T} - \ln \frac{N_C}{n_1}. \quad (10.126)$$

При виведенні цієї формули ми врахували, що за відліку енергії від рівня  $E_C$  енергія  $E_D$  буде негативною. У цьому температурному інтервалі час життя нерівноважних носіїв визначається винятково числом і властивостями глибоких центрів та  $\ln \Delta\tau$  лінійно зменшується із тангенсом нахилу  $\text{tg}\phi = E_D/k$  (область 2 на рисунку). За подальшого підвищення температури всі донори стають іонізованими, і рівноважна концентрація електронів визначається концентрацією донорів  $N_D$ .

Тепер можна вважати  $n_0(T) = \text{const}$ , а час релаксації визначатиметься залежністю від температури  $n_1(T)$ . Тоді із (10.123) маємо

$$\Delta\tau = \frac{N_C}{N_D} e^{-\frac{E_t}{kT}}, \quad (10.127)$$

звідки легко знаходимо 
$$\ln\Delta\tau = -\frac{E_t}{k} \cdot \frac{1}{T} + \ln \frac{N_C}{N_D}. \quad (10.128)$$

Це означає, що зростання температури, за якого рівень Фермі зсувається в область нижче рівня глибоких пасток, приводить до зростання часу життя нерівноважних носіїв таким чином, що  $\ln\Delta\tau$  лінійно збільшується із тангенсом нахилу  $\text{tg}\phi = -E_t/k$ . При цьому час збільшується аж до температури, за якої настає власна провідність (область 3 на рисунку). У цій області відбувається зростання часу життя нерівноважних носіїв температур, оскільки зростання температури приводить до інтенсивної емісії електронів із рівня пасток. У результаті зменшується заповнення пасток електронами, і як наслідок – збільшується час життя неосновних носіїв заряду (дірок), який і визначає час життя нерівноважних носіїв. За подальшого зростання температури всі центри іонізовані, і температурну залежність часу релаксації знов визначає поведінка концентрації основних носіїв  $n_0(T)$ , яка тепер поводитьсь із температурою як у власному напівпровіднику. Підставимо до (10.126) значення рівня Фермі власного напівпровідника, що можна зробити за високих температур в умовах виснаження домішок  $E_F = -E_g/2 + (3/4)\ln(m_p^*/m_n^*)$ , отримаємо

$$\ln\Delta\tau = \frac{E_g}{2k} \cdot \frac{1}{T} = \ln \frac{N_C}{n_1} - \frac{3}{4} \ln \frac{m_p^*}{m_n^*}. \quad (10.129)$$

Час життя нерівноважних носіїв знову зменшується таким чином, що  $\ln\Delta\tau$  лінійно зменшується із тангенсом нахилу  $\text{tg}\phi = E_g/2k$  (область 4 на рисунку). Необхідно зазначити, що в області температур, де рівень Фермі виходить на плато  $E_F = E_i$ , концентрації електронів і дірок майже однакові

$$n_0 \approx p_0 \approx n_i = (N_C N_V)^{1/2} e^{-\frac{E_g}{2kT}}. \quad (10.130)$$

У цьому випадку рівняння Шоклі–Ріда (10.122) має простий вигляд

$$\tau = \frac{\tau_{p0}}{2n_i} (n_i + n_1) + \frac{\tau_{n0}}{2n_i} (n_i + p_1). \quad (10.131)$$

При використанні (10.107), (10.112) та (10.130) вираз (10.131) набуває вигляду

$$\tau = \frac{\tau_{p0}}{2} \left[ 1 + (N_C/N_V)^{1/2} e^{-\frac{(E_C+E_V)/2-E_t}{kT}} \right] + \frac{\tau_{n0}}{2} \left[ 1 + (N_V/N_C)^{1/2} e^{-\frac{(E_C+E_V)/2-E_t}{kT}} \right]. \quad (10.132)$$

В області високих температур, коли

$$kT \gg \frac{E_C + E_V}{2} - E_t, \quad (10.133)$$

(10.132) спрощується до

$$\tau = \frac{\tau_{p0}}{2} (1 + \sqrt{N_C/N_V}) + \frac{\tau_{n0}}{2} (1 + \sqrt{N_V/N_C}). \quad (10.134)$$

Отже у цьому інтервалі температур час життя нерівноважних носіїв є майже сталою величиною, для оцінки якої можна вважати, що  $N_C/N_V \approx 1$ , що приводить до такої величини часу  $\tau$  (область 5 на рисунку)

$$\tau = \tau_{p0} + \tau_{n0}. \quad (10.135)$$

За великого рівня інжекції, коли  $\delta n \gg n_0$ ,  $p_0$ ,  $n_1$ ,  $p_1$  із (10.121) безпосередньо випливає, що час життя нерівноважних носіїв

$$\tau_\infty = \tau_{0n} + \tau_{0p}, \quad (10.136)$$

тобто у випадку великих рівнів інжекції є сталим і збігається із часом життя за високих температур.

У теорії Шоклі–Рида було введено перерізи захоплення електронів і дірок  $\gamma_n$  і  $\gamma_p$ , що й несуть інформацію про матеріал, для якого обчислюється час релаксації. Обчислення цих параметрів є складною мікроскопічною задачею, а послідовних теорій для таких обчислень немає. Для оцінок стають у нагоді прості моделі, що дозволяють оцінити ці перерізи. Наприклад, за випромінювальної рекомбінації на локальних центрах, коли при переході носіїв із зони на локальний центр випромінюється квант світла, на основі воднеподібної моделі можна отримати оцінку

$$\gamma = 0,069 \cdot 10^{-12} \frac{\tilde{n}}{Z^2} \left( \frac{1}{m^*} \right)^{3/2} \left( \frac{300}{T} \right)^{3/2} E_t, \text{ см}^3 \text{с}^{-1}, \quad (10.137)$$

де  $\tilde{n}$  – показник заломлення світла у кристалі,  $Z$  – зарядовий стан іонізованого локального центру. Наприклад, для іонів  $B$  у кремнії, де  $E_t = 0,046$  еВ, розрахунок за (10.137) дає величину  $\gamma_p = 3,8 \cdot 10^{-14} \text{ см}^3 \text{с}^{-1}$ , у той час як експериментальне значення цієї самої величини складає  $\gamma_p = 2,8 \cdot 10^{-14} \text{ см}^3 \text{с}^{-1}$ .

Оскільки енергія іонізації пастки (зазвичай локальними центрами рекомбінації є глибокі рівні) складає кілька десятків електрон-вольта, а енергія фонона на порядок менша, то для захоплення носія на такий центр необхідна участь десятків фононів. Ця обставина значно усклад-



нює розрахунки  $\gamma$  за фононної рекомбінації на локальних центрах. Придатною виявилась модель Лекса, де запропоновано каскадний механізм рекомбінації: вільний носій спочатку захоплюється на збуджений рівень пастки – емітується один фонон, носій переходить на нижчий рівень – емітується інший фонон і т.д. Але цей механізм не прийнятний для розрахунків рекомбінації на нейтральних центрах.

За ударної рекомбінації на локальних центрах вивільнена енергія передається електрону (або дірці), що безпосередньо знаходяться поблизу центра рекомбінації. У цьому випадку

$$\gamma = AE_g^{-3} \left( \frac{E_t}{E_g} \right) \left( \frac{E_t}{E_g - E_t} \right)^{3/2} \frac{n_0}{(1 + bn_0)^2}, \quad \text{см}^3 \text{с}^{-1}, \quad (10.138)$$

де  $A$  та  $b$  – сталі, залежні від властивостей напівпровідника. При  $n_0 < b$   $\gamma$  зростає зі збільшенням концентрації носіїв. При  $n_0 > b$   $\gamma$  зменшується зі збільшенням концентрації носіїв. Зауважимо, що залежність  $\gamma$  від рівноважної концентрації носіїв визначається участю у такому механізмі рекомбінації трьох частинок.

#### 10.3.4. Спін-залежна рекомбінація в напівпровідниках

Явище спін-залежної рекомбінації полягає у взаємозалежності між інтенсивністю процесів рекомбінації нерівноважних носіїв та їх спіновим станом. Аналогічно до механізмів рекомбінації, які ми розглядали раніше, спін-залежна рекомбінація може відбуватись як за участі, так і без участі центрів рекомбінації. Спочатку розглянемо спін-залежну рекомбінацію, що відбувається в електронно-дірковому газі, де за нульових температур і магнітних полів енергетичні стани як електронів, так і дірок є дворазово виродженими за проекціями спіну. Розглянемо випромінювальну рекомбінацію, коли третім агентом у процесах рекомбінації виступають фотони із визначеною поляризацією. Із загальних міркувань ясно: якщо під дією зовнішнього збурення до зони провідності потрапляє додатковий (нерівноважний) електрон із проекцією спіну  $+1/2$ , то сумарне значення проекції спіну електронів у заповненій зоні зменшується на величину  $1/2$ , тобто виникає дірка із проекцією спіну  $-1/2$ . І навпаки, електрон із проекцією спіну  $-1/2$ , виходячи із валентної зони, залишає по собі дірку із проекцією спіну  $+1/2$ . Найважливішими з практичного погляду чинниками, що ведуть до спінової поляризації нерівноважних носіїв струму у напівпровідниках, є поляризоване лазерне опромінювання та магнітне поле.

У разі опромінювання напівпровідника поляризованим світлом ймовірності генерації електронів (дірок) із проекціями спіну  $+1/2$  та  $-1/2$  не дорівнюють одна одній. У найпростішому випадку стаціонарного опромінювання для врахування цього факту слід ввести дві різні ймовірності  $g_+ dt$  та  $g_- dt$  – генерації нерівноважного електрона із проекцією спіну  $+1/2$  та  $-1/2$ . Рекомбінаційні електронні переходи супроводжуються випромінюванням фотонів. Оскільки кожен фотон має ненульову проекцію спіну  $s_z \pm 1$  на напрямок хвильового вектора, то ймовірність випромінювальної рекомбінації електрона із заданою проекцією спіну у першому наближенні можна вважати пропорційною до концентрації дірок із таким самим знаком проекції спіну, тобто  $r_{\pm} = \alpha \cdot p_{\pm}$  (припущення справджується за збереження в кожному акті рекомбінації сумарного значення проекцій спіну нерівноважних носіїв і фотона). Можна вважати, що так само відбувається процес випромінювальної рекомбінації й у дірок. Тоді кінетичні рівняння для парціальних концентрацій нерівноважних носіїв можна записати у вигляді

$$\frac{dn_{\pm}}{dt} = g_{\pm} - \alpha \cdot n_{\pm} p_{\pm} - \frac{n_{\pm} - n_{\pm}}{2\tau_{sn}}, \quad (10.139)$$

$$\frac{dp_{\pm}}{dt} = g_{\pm} - \alpha \cdot n_{\pm} p_{\pm} - \frac{p_{\pm} - p_{\pm}}{2\tau_{sp}}. \quad (10.140)$$

Останні доданки у правих частинах рівнянь (10.139)–(10.140) ураховують, що тепловий рух може спричинити стохастизацію напрямків спінів, тобто концентрацію носіїв, які відрізняються лише знаком проекції спіну. При цьому характерні часи спінової стохастизації електронів і дірок  $\tau_{sn}$  та  $\tau_{sp}$  можна вважати рівними до часів поперечної спінової релаксації електронів провідності й валентних електронів, відповідно. Введемо швидкість рекомбінації нерівноважних носіїв як загальну кількість актів рекомбінації, що відбуваються в одиниці об'єму напівпровідника за одиницю часу. Очевидно, що швидкість має бути пропорційна добутку концентрацій електронів і дірок. Крім того, варто врахувати вимоги принципу Паулі. Таким чином, можна записати

$$R = \alpha \cdot (n_+ p_+ + n_- p_-). \quad (10.141)$$

Характеристикою спінового стану системи ферміонів є спінова поляризація, яку для електронів ( $P_n$ ) і дірок ( $P_p$ ) у напівпровіднику можна визначити так

$$P_n = \frac{n_+ - n_-}{n}, \quad P_p = \frac{p_+ - p_-}{p}, \quad (10.142)$$

де  $n = n_+ + n_-$  і  $p = p_+ + p_-$  – повні концентрації нерівноважних електронів і дірок, відповідно. Швидкість рекомбінації (10.141) тепер неважко подати у термінах спінової поляризації

$$R(P_n, P_p) = \alpha \cdot (1 + P_n P_p)(np)/2. \quad (10.143)$$

Звідси випливає, що відносна зміна швидкості випромінювальної рекомбінації, обумовлена спіноюю поляризацією нерівноважних носіїв, визначається добутком спінової поляризації електронів і дірок

$$\frac{\Delta R}{R} = \frac{R(P_n, P_p) - R(0, 0)}{R(0, 0)} = P_n P_p. \quad (10.144)$$

Для оцінки величин спінової поляризації носіїв у напівпровіднику звернемося до стаціонарного випадку генераційно-рекомбінаційних процесів, що описується рівняннями (10.139)–(10.140), де похідні за часом вважаються нульовими

$$g_{\pm} - \frac{n_{\pm}}{\tau_{\pm}} - \frac{n_{\pm} - n_{\pm}}{2\tau_{sn}} = 0, \quad (10.145)$$

$$g_{\pm} - \frac{p_{\pm}}{\tau_{\pm}} - \frac{n_{\pm} - n_{\pm}}{2\tau_{sn}} = 0. \quad (10.146)$$

Тут введено позначення для середніх часів життя електронів  $\tau_{\pm}$  і дірок  $\tau'_{\pm}$  зі спіноюю проекцією  $\pm 1/2$ . Звідси для концентрації нерівноважних носіїв  $n_{\pm}$  та  $p_{\pm}$  маємо

$$n_{+} - n_{-} = (g_{+}\tau_{+} - g_{-}\tau_{-}) \cdot \frac{\tau_{sn}}{\tau_{sn} + \tau}, \quad n_{+} + n_{-} = (g_{+}\tau_{+} + g_{-}\tau_{-}) \cdot \left[ 1 + \frac{\tau_{-} - \tau_{+}}{2(\tau_{sn} + \tau)} \right],$$

$$p_{+} - p_{-} = (g'_{+}\tau'_{+} - g'_{-}\tau'_{-}) \cdot \frac{\tau_{pn}}{\tau_{pn} + \tau'}, \quad p_{+} + p_{-} = (g'_{+}\tau'_{+} + g'_{-}\tau'_{-}) \cdot \left[ 1 + \frac{\tau'_{-} - \tau'_{+}}{2(\tau_{pn} + \tau')} \right].$$

У наближенні  $2(\tau_{sn(p)} + \tau) \gg \tau_{-} - \tau_{+}$ , звідси отримуємо

$$P_n = P_n^{(\max)} \frac{\tau_{sn}}{\tau_{sn} + \tau}, \quad P_p = P_p^{(\max)} \frac{\tau_{pn}}{\tau_{pn} + \tau}. \quad (10.147)$$

При виведенні цих формул ми припускали, що  $g_{+} > g_{-}$  і  $g'_{+} > g'_{-}$ . Крім того, тут використано такі позначення для середніх часів життя нерівноважних електронів  $\tau = (\tau_{+} + \tau_{-})/2$  і дірок  $\tau' = (\tau'_{+} + \tau'_{-})/2$ , відповідно. Через

$$P_n^{(\max)} = \frac{g_{+}\tau_{+} - g_{-}\tau_{-}}{g_{+}\tau_{+} + g_{-}\tau_{-}} \quad (10.148)$$

та

$$P_p^{(\max)} = \frac{g'_{+}\tau'_{+} - g'_{-}\tau'_{-}}{g'_{+}\tau'_{+} + g'_{-}\tau'_{-}} \quad (10.149)$$

позначено максимальні значення спінової поляризації, що відповідають нескінченно великим часам спінової стохастизації. Підставимо

(10.148)–(10.149) до (10.143) і знайдемо, що відносна зміна часу життя нерівноважних носіїв визначається добутком максимальних значень спінової поляризації електронної і діркової підсистем напівпровідника

$$\frac{\Delta R}{R} = P_n^{(\max)} P_p^{(\max)} \frac{\tau_{sn} \tau_{sp}}{(\tau_{sn} + \tau)(\tau_{sn} + \tau)}, \quad (10.150)$$

де вважалось, що  $\tau = (\tau_+ + \tau_-)/2 = (\tau'_+ + \tau'_-)/2$  – середній час життя електрон-діркової пари. Оскільки величина  $P_n^{(\max)} P_p^{(\max)}$  є додатною, то із (10.150) випливає важливий факт: наявність спінової поляризації приводить до прискорення процесу випромінювальної рекомбінації. Для грубої оцінки можна покласти, що  $P_n^{(\max)} = P_p^{(\max)} \approx 1/2$ , та  $\tau \sim \tau_{sp} \ll \tau_{sn}$  (такі припущення непогано справджуються для GaAs). Тоді

$$\frac{\Delta R}{R} \approx \frac{1}{4} \frac{\tau_{sp}}{\tau_{sp} + \tau}. \quad (10.151)$$

Таким чином можна стверджувати, що за досить великих часів спінової релаксації у системі дірок відносно прискорення швидкості рекомбінації за рахунок спінової поляризації може наближатись до 0,25. В експериментах спостерігались часи релаксації  $\tau_{sp} \approx 2,8 \times 10^{-10}$  с. Отже, унаслідок опромінювання напівпровідника повністю поляризованим світлом швидкість електрон-діркової рекомбінації може змінюватись майже на 10 %.

Вираз (10.150) можна використати для оцінки впливу зовнішнього магнітного поля на швидкість спін-залежної рекомбінації. Дійсно, у випадку неvirодженого напівпровідника  $n_{\pm} \sim \exp(\mu_B H / kT)$ , де  $\mu_B$  – магнетон Бора, і, як випливає з (10.142)  $P_n = P_p = \mu_B H / kT$ , тобто

$$\frac{\Delta R}{R} = \left( \frac{\mu_B H}{kT} \right)^2. \quad (10.152)$$

Цей вираз дозволяє стверджувати, що для полів, які не перевищують 10 кЕ, за кімнатних температур відносна зміна швидкості рекомбінації буде меншою від  $5 \cdot 10^{-6}$ . Експериментальні дослідження показали, що величина  $\Delta R/R$  значно перевищує цю оцінку. Головною причиною такої розбіжності є те, що рекомбінація носіїв головним чином пов'язана не з прямими переходами електронів між валентною зоною та зоною провідності, а з переходами через локалізовані електронні стани, обумовлені наявністю домішок і дефектів. Інтенсивність рекомбінаційних переходів через електронні центри залежить від спінового стану електронно-діркової системи напівпровідника. Розглянемо прості електронні центри донорного та акцепторного типів. Неіонізо-

ваний донорний центр несе на собі "зайвий" електрон провідності, а незаряджений акцепторний центр – дірку. В обох випадках на центрі існує некомпенсований спін  $s = 1/2$ . Такий центр у магнітному полі поводить себе як парамагнітний атом, і сам може бути названий парамагнітним. Процес іонізації як донорного, так і акцепторного центрів супроводжується його переходом із парамагнітного стану до стану зі спіном  $s = 0$ . Спінова залежність процесу рекомбінації, таким чином, пов'язана саме з наявністю у напівпровіднику парамагнітних електронних центрів. Дуже важливо, що основним легуючим домішкам напівпровідників зазвичай відповідають мілкі енергетичні рівні у забороненій зоні, тому за не надто низьких температур ці домішки практично повністю іонізовані, отже непарамагнітні. У такій ситуації спін-залежним каналом рекомбінації виявляються електронні переходи через глибокі рівні у забороненій зоні.

Розглянемо схему процесу рекомбінації через глибокий рівень. До початку рекомбінації електронний центр знаходиться у парамагнітному стані із проекцією спіну  $-1/2$  (на рис. 10.9 а позначено стрілочкою вниз). На першому етапі центр захоплює електрон із проекцією спіну  $+1/2$ , унаслідок чого стає діамагнітним (рис. 10.9 б). На другому етапі діамагнітний центр захоплює дірку із зони провідності й знов стає парамагнітним (рис. 10.9 в). На всіх етапах процесу рекомбінації повний електричний заряд і сумарна проекція спіну всіх носіїв

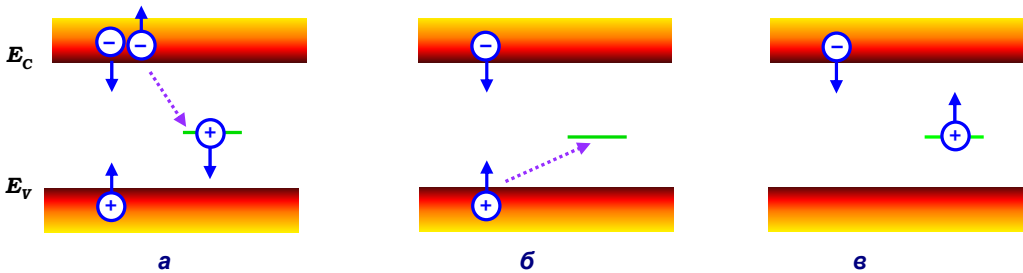


Рис. 10.9. Спін-залежна рекомбінація через парамагнітні центри (глибокі рівні)

дорівнюють нулю, згідно із принципом електронейтральності та тенденцією до збереження проекції спіну. Початковий і кінцевий стани системи відрізняються лише кількістю електронно-діркових пар, і рекомбінаційний перехід може повторюватись у стаціонарному режимі з огляду на наявність теплової генерації носіїв. Зауважимо, що через пастку із заданим спіновим станом може рекомбінувати лише електрон із тим самим напрямком проекції спіну. Однак ймовірність такої події є відносно малою, оскільки при захоплюванні електрона на центр спінова проекція електрона або центра має змінитися на протилежну – захоплювання без зміни проекції спіну принципом Паулі заборонено. Таким чином, якщо спін-орбітальна взаємодія незначно впливає на

рекомбінаційні переходи через глибокі рівні, має відчуватися дія "закону збереження проекції спіну". На відміну від прямих міжзонних переходів, рекомбінація через глибокі рівні не супроводжується випромінюванням фотонів, а її швидкість має прямо залежати від концентрації парамагнітних центрів із додатною  $N_+$  та від'ємною  $N_-$  проекціями спіну

$$R = \alpha(N_+n_- + N_-n_+). \quad (10.153)$$

Це рівняння можна звести до вигляду, аналогічному (10.143),

$$R(P_n, P_p) = \alpha(1 + P_n P_p)(nN)/2, \quad (10.154)$$

де  $N = N_+ + N_-$  – концентрація глибоких парамагнітних центрів. Тоді відносна зміна швидкості рекомбінації на парамагнітних центрах

$$\frac{\Delta R}{R} = \frac{R(P_n, P_p) - R(0, 0)}{R(0, 0)} = -P_n P_p. \quad (10.155)$$

Відмінність знаку у правій частині від (10.144) пов'язана з існуванням спінової проекції  $\pm 1$  у фотона, що випромінюється внаслідок прямої міжзонної рекомбінації. У випадку рекомбінації через парамагнітні центри у зовнішньому магнітному полі можна записати формулу, аналогічну (10.155). Оскільки відносна зміна швидкості рекомбінації пропорційна відносній зміні концентрації нерівноважних носіїв або відносній зміні провідності, то можна записати

$$\frac{\delta(\Delta\sigma)}{\Delta\sigma} = \left( \frac{\mu_B H}{kT} \right)^2. \quad (10.156)$$

Оцінки показують, що за кімнатної температури в полях  $\sim 10^3$  Ерстед  $\delta(\Delta\sigma)/\Delta\sigma \approx 10^{-6}$ . Зрозуміло, що експериментально вимірювати такі ефекти – непроста задача. Дійсно, під дією магнітного поля провідність може змінюватись і завдяки ефекту магнітоопору (що майже на п'ять порядків сильніший) та на його фоні ефект спин-залежної рекомбінації стає надзвичайно слабким. Необхідно створити умови для вимірювання лише різниці у провідності за рахунок спінової поляризації. Цього можна досягнути, використовуючи умови спінового резонансу. Як відомо, спіновий резонанс настає, коли частота електромагнітного поля дорівнює енергії зеєманівського розщеплення рівня. Експериментальні спостереження спин-залежної рекомбінації через парамагнітні рівні проводились на зразках кремнію  $n$ -типу, де глибокі електронні рівні походили від дефектів, розташованих поблизу поверхні зразків. Унаслідок опромінювання кремнію світлом у ньому виникали нерівноважні електрони, а докладання до зразка різниці потенціалів обумовлювало нерівноважний електричний струм. Спінова поляриза-

ція електронів провідності та парамагнітних центрів утворювалась стаціонарним магнітним полем напруженістю 3000 Ерстед, що дозволяла спостерігати спіновий резонанс системи парамагнітних центрів у надвисокочастотному магнітному полі із довжиною хвилі 3,0 см. У момент виникнення спінового резонансу спостерігалася зміна нерівноважної складової  $\delta(\Delta\sigma)/\Delta\sigma \approx 10^{-5-4}$ . Було помічено, що об'єкти, в яких спостерігаються ефекти спін-залежної рекомбінації, є суттєво неупорядкованими системами. Це, насамперед поверхня кремнію, аморфний кремній, пластично деформований і полікристалічний кремній. У таких дефектних структурах може існувати велика кількість різноманітних центрів рекомбінації. Таким чином спін-залежна рекомбінація може відбуватись за переходу носіїв заряду між близькими парами локалізованих станів на центрах рекомбінації.

Механізм рекомбінації на парах просторово близьких центрів було запропоновано Капланом, Соломоном і Моттом. За механізмом Каплана–Соломона–Мотта (КСМ) розглядається пара просторово близьких центрів ( $a$  та  $b$ ) локалізації електронів. Припускається, що кожному з них відповідає один електронний рівень у забороненій зоні (рис. 10.10). Між центрами локалізації електронів (або, що те ж саме, між їхніми енергетичними рівнями  $E_a$  та  $E_b$ ) можливі електронні переходи.

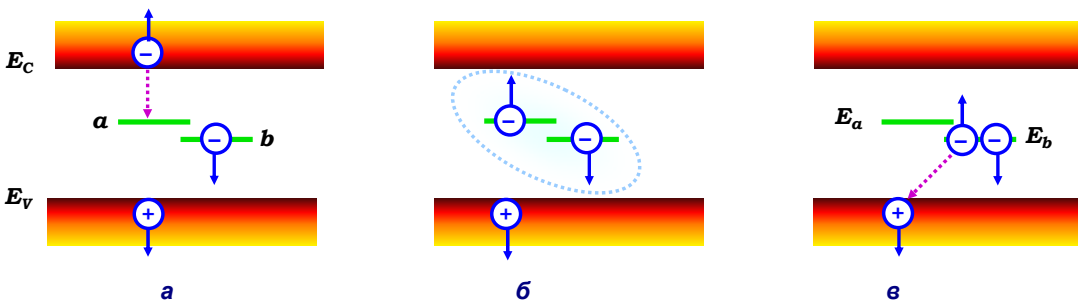


Рис. 10.10. Спін-залежна рекомбінація за механізмом КСМ

Нехай на початку акту рекомбінації центр  $a$  незаповнений, а на центрі  $b$  знаходиться один електрон, тобто центр  $b$  є парамагнітним. Акт рекомбінації за механізмом КСМ відбувається у три етапи: на першому – шляхом переходу електрона із зони провідності на центр  $a$  утворюється двоелектронний стан на парному центрі (обидва рівня заповнені електронами – рис. 10.10 а); на другому – відбувається перехід електрона із центра  $a$  на центр  $b$  (рис. 10.10 б); на третьому – електрон переходить із центра  $b$  до валентної зони, рекомбінуючи із діркою (рис. 10.10 в). Принциповим у механізмі КСМ є припущення, що акт рекомбінації відбувається з утворенням проміжного двоелектронного стану. Взагалі кажучи, ймовірності всіх трьох процесів (рис. 10.10) є спін-залежними. Але найвідчутнішою є спінова залежність імовірності

переходу між локалізованими рівнями, оскільки тут важлива не поляризація всієї спінової системи, а лише значна кореляція спінів двох спарених електронів, обумовлена принципом Паулі. Послідовна теорія механізму КСМ показує, що відносна зміна провідності  $\delta\Delta\sigma/\Delta\sigma \sim H^0$  узагалі не залежить від магнітного поля. У цій моделі магнітне поле необхідне тільки для створення умов спінового резонансу. Спінова поляризація при цьому не має значення. Експериментальні дослідження показали, що модель КСМ досить добре описує спін-залежну рекомбінацію у неупорядкованому кремнії.

Фундаментальні дослідження явищ спін-залежної рекомбінації дозволяють не тільки правильно ідентифікувати технологічні домішки та дефекти у напівпровідникових структурах, але й сподіватись на можливість розробки принципово нових типів електронних приладів, заснованих на спін-залежному транспорті (т. зв. приладів спітроніки). У цьому зв'язку велике значення має перехід до нанометрових розмірів конструктивних елементів електронних пристрів у разі поєднання в них напівпровідникових і феромагнітних властивостей. Дійсно, кожна нанометрова феромагнітна частинка є монодоменною, тобто в ній не можуть утворюватись домени із протилежними напрямками намагніченості. Завдяки цьому навіть за відсутності зовнішнього магнітного поля маленька феромагнітна частинка має ненульовий магнітний момент і може розглядатися як елемент магнітного носія інформації. З іншого боку, навколо феромагнітної частинки, імплантованої до напівпровідникового пристрою, утворюється сильне внутрішнє поле, яке має ті ж самі фізичну природу й порядок величини, що й молекулярне поле Вейса (аналог т. зв. ближнього поля в оптиці). Оскільки останнє принаймні на два порядки перевищує типові значення полів, що використовуються в експериментах з електронного парамагнітного резонансу, зміна магнітного стану феромагнітного елемента за певних умов істотно впливає на спін-залежний компонент рекомбінації електрон-діркових пар, що відбивається на електропровідності всього електронного пристрою.

## 10.4. Задачі

1. Визначте час життя нерівноважних носіїв за рекомбінації через центр захоплення, що утворює два енергетичних рівня.

**Розв'язок.** Нехай двозарядова рекомбінаційна пастка характеризується двома енергетичними рівнями  $E_1$  та  $E_2$ . Таким чином, аналогічно до (10.109) можна записати, що сумарним темпом захоплення електронів на рівень  $E_1$  багатозарядної пастки є

$$R_{n1} = -\gamma_{n1} N_t (f_0 n - f_1 n_1). \quad (10.157)$$



Сумарний темп захоплювання електронів на рівень  $E_2$

$$R_{n2} = -\gamma_{n2} N_t (f_1 n - f_2 n_2). \quad (10.158)$$

У цих формулах введено такі позначення:  $\gamma_{n1}$  та  $\gamma_{n2}$  – коефіцієнти захоплювання електрона на рівні  $E_1$  (коли на пастці не було захоплено електронів) та  $E_2$  (якщо на пастці вже був один електрон), відповідно;  $N_t$  – концентрація рекомбінаційних пасток;  $N_t f_i$  – нерівноважна концентрація пасток в  $i$ -му зарядовому стані;  $n_1$  та  $n_2$  – рівноважна концентрація електронів у зоні за збігу рівня Фермі із рівнями  $E_1$  та  $E_2$ , відповідно.

Аналогічно можна записати сумарний темп захоплювання дірок:

♦ на рівень  $E_1$  
$$R_{p1} = -\gamma_{p1} N_t (f_1 p - f_0 p_1), \quad (10.159)$$

♦ на рівень  $E_2$  
$$R_{p2} = -\gamma_{p2} N_t (f_2 p - f_1 p_2), \quad (10.160)$$

де  $\gamma_{p1}$  та  $\gamma_{p2}$  – коефіцієнти захоплювання дірки на рівень  $E_1$  (коли на пастці не було захоплено електронів) та рівень  $E_2$  (якщо на пастці вже був один електрон), відповідно;  $p_i$  – рівноважна концентрація дірок у зоні за збігу рівня Фермі з енергетичним рівнем  $E_i$  пастки.

Таким чином, повні темпи захоплювання електронів і дірок на пастки у різних зарядових станах є сумами

$$R_n = R_{n1} + R_{n2} \quad \text{та} \quad R_p = R_{p1} + R_{p2}. \quad (10.161)$$

У стаціонарному стані концентрація пасток у кожному зарядовому стані не залежить від часу. Це означає, зокрема, що

$$R_{n1} = R_{p1}, \quad R_{n2} = R_{p2}. \quad (10.162)$$

Крім того, умова сталості повної концентрації пасток дає

$$f_0 + f_1 + f_2 = 1. \quad (10.163)$$

Три рівняння (10.163)–(10.164) дають повну систему для визначення всіх нерівноважних чисел заповнення  $f_0, f_1, f_2$ . Зокрема, легко отримати, що

$$f_1 = \frac{\gamma_{n1} n + \gamma_{p1} n p_1}{\gamma_{n1} n_1 + \gamma_{p1} p} f_0, \quad (10.164)$$

$$f_2 = \frac{\gamma_{n2} n + \gamma_{p2} n p_2}{\gamma_{n2} n_2 + \gamma_{p2} p} f_1. \quad (10.165)$$

Вважаючи напівпровідник невідродженим і використовуючи закон діючих мас  $np = n_i^2$ , отримуємо із (10.157)–(10.165)

$$R_{n1} = N_t \frac{(f_1 + f_0)(n_i^2 - np)}{(n + n_1)/\gamma_{p1} + (p + p_1)/\gamma_{n1}}, \quad (10.166)$$

$$R_{n2} + N_t \frac{(f_2 + f_1)(n_i^2 - np)}{(n + n_2)/\gamma_{p2} + (p + p_2)/\gamma_{n2}}. \quad (10.167)$$

За малих концентрацій пасток ( $\delta p \cong \delta n$ ) і невеликого відхилення від рівноваги ( $\delta n \ll n_0, p_0$ ) у (10.166)–(10.167) можна покласти, що  $n = n_0 + \delta n$ ,  $p = p_0 + \delta p$ ,  $f_i = f_i^0 + \delta f_i$ . Зрозуміло, що  $\delta f_1 \sim \delta n$ . Тоді, із точністю до лінійних доданків за  $\delta n$  отримуємо

$$R_{nj} = \frac{N_t(n_0 + p_0)(f_j^0 + f_{j-1}^0)}{(n_0 + n_j)/\gamma_{pj} + (p_0 + p_j)/\gamma_{nj}} \delta n. \quad (10.168)$$

Звідси, зважаючи на те, що

$$\frac{1}{\tau} = \frac{R_1 + R_2}{\delta n}, \quad (10.169)$$

отримаємо, що час життя нерівноважних носіїв за рекомбінації через двозарядові пастки виражається формулою

$$\frac{1}{\tau} = \frac{N_t(n_0 + p_0)}{p_0 + p_1 + p_1 p_2 / p_0} \left\{ \frac{p_0}{p_0/\gamma_{n1} + n_1/\gamma_{p1}} + \frac{p_1}{p_0/\gamma_{n2} + n_2/\gamma_{p2}} \right\}. \quad (10.170)$$

**2.** Час життя нерівноважних носіїв у германії  $p$ -типу в інтервалі температур від  $T_1$  до  $T_2$  залежить від температури, згідно з емпіричною формулою

$$\tau = (A + B \cdot \text{th}(a - b/T)) \cdot 10^{-5} \text{ с}, \quad (10.171)$$

де  $A$ ,  $B$ ,  $a$  та  $b$  – константи, що визначаються з експерименту.

Відомо, що рекомбінація відбувається за участю рекомбінаційних пасток, які утворюють два рівні –  $E_1$  у нижній частині забороненої зони, та  $E_2$  – у верхній. Концентрація пасток –  $N_t$ . Вважаючи, що в даному інтервалі температур  $N_V$  є сталою, визначте коефіцієнти захоплення електронів  $\gamma_{n1}$  та  $\gamma_{n2}$  і концентрацію основних носіїв  $p_0$ .

**Розв'язок.** Оскільки рівень  $E_2$  знаходиться поблизу дна зони провідності, а  $E_1$  – поблизу стелі валентної зони, то рівень  $E_2$  буде в основному заповнюватись електронами, а  $E_1$  – дірками. Це означає, що помітними величинами будуть лише концентрації  $n_2$  та  $p_1$ , тобто можна вважати, що  $n_0 \ll p_0$ ,  $n_1 \ll p_0$ ,  $p_2 \ll p_0$ . У такому випадку (10.170) спрощується до вигляду

$$\tau^{-1} = \frac{N_t}{1 + p_1/p_0} \left( \gamma_{n1} + \frac{\gamma_{n2}(p_1/p_0)}{1 + (\gamma_{n2}n_2)/(\gamma_{p2}p_0)} \right). \quad (10.172)$$

Емпірична формула (10.171) показує, що в поведінці  $\tau(T)$  існує високотемпературна область, де час життя нерівноважних носіїв майже не залежить від температури. Згідно із (10.172) таке високотемпературне плато може існувати лише у випадку, коли другий доданок у знаменнику в дужках буде істотно менший за одиницю. У такому випадку маємо

$$\tau^{-1} = \frac{N_t}{p_0 + p_1} (\gamma_{n1}p_0 + \gamma_{n2}p_1). \quad (10.173)$$

Врахуємо, що  $p_1 = N_V e^\zeta$ , де  $\zeta = (E_V - E_1)/kT$ . Тоді (10.173) можна переписати так

$$N_t \tau = \frac{\frac{P_0}{N_V} e^{-\zeta} + 1}{\gamma_{n2} \left( \frac{\gamma_{n1} P_0}{\gamma_{n2} N_V} e^{-\zeta} + 1 \right)}. \quad (10.174)$$

Беручи до уваги, що  $\frac{\gamma_{n1} P_0}{\gamma_{n2} N_V} e^{-\zeta} = \exp \left[ -\zeta + \ln \left( \frac{\gamma_{n1} P_0}{\gamma_{n2} N_V} \right) \right]$ , із (10.174) отримаємо

$$\tau = \frac{1}{2N_t} \left( \frac{1}{\gamma_{n1}} + \frac{1}{\gamma_{n2}} \right) - \frac{1}{2N_t} \left( \frac{1}{\gamma_{n2}} - \frac{1}{\gamma_{n1}} \right) \cdot \text{th} \left[ (1/2) \ln \left( \frac{\gamma_{n1} P_0}{\gamma_{n2} N_V} \right) - \frac{E_V - E_1}{2kT} \right]. \quad (10.175)$$

Порівнюючи цей вираз з емпіричною формулою (10.171), отримаємо рівняння

$$\frac{1}{2N_t} \left( \frac{1}{\gamma_{n1}} + \frac{1}{\gamma_{n2}} \right) = A \cdot 10^{-5}, \quad (10.176)$$

$$\frac{1}{2N_t} \left( \frac{1}{\gamma_{n2}} - \frac{1}{\gamma_{n1}} \right) = -B \cdot 10^{-5}, \quad (10.177)$$

$$(1/2) \ln \left( \frac{\gamma_{n1} P_0}{\gamma_{n2} N_V} \right) = a, \quad (10.178)$$

$$E_V - E_1 = 2kb, \quad (10.179)$$

звідси маємо

$$E_1 = E_V - 2kb, \quad \gamma_{n2} = \frac{1}{N_t(A+B)} 10^5,$$

$$\gamma_{n1} = \frac{1}{N_t(A-B)} 10^5, \quad p_0 = N_V e^{2a} \frac{A-B}{A+B}.$$

**3.** У германії  $n$ -типу концентрація рекомбінаційних центрів  $N_t = 5 \cdot 10^{12}$  см. Вони характеризуються енергетичним рівнем  $E_1 = E_g/2$ . За кімнатної температури перерізи захоплювання електронів і дірок однакові за невеликих відхилень від рівноваги. При цьому  $\tau = 10^{-4}$  с,  $\rho = 5$  Ом·см. Визначте переріз захоплювання.

**4.** У германії  $n$ -типу концентрація основних носіїв  $n_0 = 10^{14}$  см<sup>-3</sup>. Концентрація рекомбінаційних центрів  $N_t = 5 \cdot 10^{12}$  см<sup>-3</sup>. Їхній енергетичний рівень  $E_t$  розташований у верхній половині забороненої зони. За кімнатної температури час життя нерівноважних носіїв  $\tau = 17 \cdot 10^{-6}$  с, при  $T = 200$  К час життя  $\tau = 2 \cdot 10^{-6}$  с. За більш низьких температур час життя залежить від температури як  $\tau = C/\sqrt{T}$ ,  $C$  – константа. Визначте  $E_t$  та переріз захоплювання дірок, вважаючи їх сталими.

### Список літератури

1. *Бон-Бруевич В.Л., Калашиников С.Г.* Физика полупроводников. – М.: Наука, 1990.
2. *Смит Р.* Полупроводники. – М.: Мир, 1982.
3. *Neamen D.* Semiconductor Physics and Devices. – N.Y.: McGraw-Hill, 2003.
4. *Зеегер К.* Физика полупроводников. – М.: Мир, 1977.
5. *Третьяк О.В., Львов В.А., Барабанов О.В.* Фізичні основи спінової електроніки. – К.: ВПЦ "Київський університет", 2002.

# Розділ 11

## ДИФУЗІЯ ТА ДРЕЙФ НОСІІВ ЗАРЯДУ

Сучасна напівпровідникова електроніка зазвичай має справу не з однорідними напівпровідниками, а з напівпровідниковими структурами, що базуються на використанні неоднорідних матеріалів. Найпростішим прикладом таких матеріалів є  $p$ - $n$ -перехід, що утворюється на межі розподілу напівпровідника із різним типом провідності. Якщо такий перехід утворено з одного напівпровідника, легovanого різного типу домішками (по один бік переходу – донорами, а по інший – акцепторами), то він називається гомопереходом; якщо такий перехід утворено межею розподілу двох різних напівпровідників, наприклад Si-Ge, – він називається гетеропереходом. Іншим прикладом неоднорідних напівпровідникових матеріалів є так звані квантові ями та надґратки, що являють собою один або систему тонких ( $\sim 10$  нм) шарів напівпровідника, перемежованих тонкими шарами діелектрика або широкозонного напівпровідника. Усі ці структури характеризуються сильною залежністю концентрації носіїв від координат, тобто наявністю ненульових градієнтів концентрації, яку викликає додатковий тип перенесення зарядів – дифузійний струм, який нарівні із дрейфовим струмом визначає характерні для напівпровідників риси електронного транспорту. Однак градієнти концентрації носіїв у напівпровідниках можуть виникати і в макроскопічно однорідних системах. Які ж властивості таких систем?

### ■ 11.1. Дифузійні та дрейфові струми

Для прикладу розглянемо неоднорідний напівпровідник, в якого концентрації електронів  $n(\mathbf{r})$  і дірок  $p(\mathbf{r})$  є функціями координат, що змінюються від точки до точки. Із цієї причини поряд із дрейфовим струмом, що викликаний електричним полем, виникне дифузійний струм, обумовлений дифузією носіїв заряду з областей, де їхня концентрація є великою в області з меншою концентрацією, тобто дифузійний струм, викликаний наявністю градієнтів концентрації носіїв. Це стає зрозумілим, якщо розглянути хаотичний рух частинок в області напівпровідника, де концентрація носіїв зростає вздовж одного із напрямків. Нехай для визначеності це буде напрямком уздовж вісі  $Ox$  (рис. 11.1). Розглянемо рух носіїв заряду через шари 1 та 2 товщиною  $dx$ . Завдяки хаотичному руху носіїв заряду залишатимуть шар 1. А оскільки електрони з рівною ймовірністю можуть рухатись як уздовж, так і проти напрямку вісі  $Ox$ , половина їх піде до шару 2. Разом із тим за

цей час до шару 1 перейдуть електрони із шару 2. А оскільки в шарі 2 електронів більше за їхню кількість в шарі 1, то зворотний потік частинок ( $2 \rightarrow 1$ ) буде більшим за прямий ( $1 \rightarrow 2$ ). Нехай концентрація електронів у шарі 1 дорівнює  $n(x - dx/2)$ , а їхня концентрація в шарі 2 –  $n(x + dx/2)$ . Потік частинок пропорційний різниці цих концентрацій

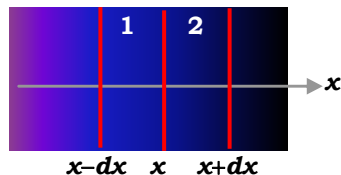


Рис. 11.1. Формування дифузійного потоку

$$I \sim [n(x - dx/2) - n(x + dx/2)] = -\frac{dn}{dx} dx, \quad (11.1)$$

тобто частинки перетікають з області великих до області малих концентрацій. Іншими словами, дифузійний рух частинок відбувається в напрямку, протилежному градієнту їхньої концентрації. Таким чином, можна записати, що дифузійний потік електронів

$$I_n = -D_n \frac{dn}{dx}. \quad (11.2)$$

Напрямок електричного струму електронів визначається як протилежний рух заряджених частинок, тому можна записати для густини дифузійного струму електронів

$$j_n = eD_n \frac{dn}{dx}. \quad (11.3)$$

Оскільки напрямок електричного струму дірок збігається із напрямком руху частинок, то густина дифузійного струму дірок

$$j_p = -eD_p \frac{dp}{dx}. \quad (11.4)$$

Якщо напівпровідник знаходиться в зовнішньому електричному полі  $E$ , то під його дією електрони та дірки набувають направлено руху, й з'являється електронний і дірковий компоненти дрейфового струму

$$j_n^{(dr)} = en\mu_n E, \quad j_p^{(dr)} = ep\mu_p E. \quad (11.5)$$

Густина повного електричного струму в неоднорідному напівпровіднику таким чином визначатиметься рівнянням

$$j = \sigma \cdot E + e \left( D_n \frac{dn}{dx} - D_p \frac{dp}{dx} \right), \quad (11.6)$$

де електропровідність  $\sigma = e(n\mu_n + p\mu_p)$ . У неоднорідному напівпровіднику з анізотропним законом дисперсії, де транспортні властивості визначаються тензорами провідності  $\sigma_{il}$  і дифузії  $D_{il}^{(n,p)}$ , повний струм можна описати тензорним рівнянням

$$j_k = \sigma_{kl} E_l + e \left( D_{il}^{(n)} \frac{dn}{dx_l} - D_{il}^{(p)} \frac{dp}{dx_l} \right). \quad (11.7)$$

## ■ 11.2. Нерівноважні напівпровідники. Квазірівні Фермі

Перед детальним обговоренням дифузійного та дрейфового рухів носіїв заряду зауважимо, що дифузія та дрейф є важливими характеристиками транспорту носіїв у системах із надлишковою концентрацією носіїв, тобто у нерівноважних системах. Для опису таких систем за допомогою знайомих методів необхідно припустити можливість використання функції розподілу Фермі, але з іншими значеннями параметрів.

Розглянемо нерівноважні системи з нормальним розподілом частинок за квантовими станами. Зрозуміло, що в цьому випадку функція розподілу електронів (і дірок) відрізняється від рівноважної функції розподілу. Але якщо система знаходиться в майже рівноважному стані, то функція розподілу незначно відрізняється від рівноважної функції розподілу Фермі. Припустимо, що в такому випадку систему можна описати функцією розподілу Фермі, де рівень Фермі замінюється деякою іншою енергією  $F^*$ , що залежить від віддаленості системи від рівноважних умов. Керуючись цими міркуваннями, Шоклі припустив, що рівноважне співвідношення статистики можна розповсюдити на нерівноважні системи за допомогою формального введення нового параметра – квазірівня Фермі: вважається, що нерівноважну концентрацію електронів можна описати формулою

$$n = N_C \Phi_{1/2}(\xi^*), \quad (11.8)$$

де

$$\xi^* = \frac{F_n^* - E_C}{kT} \quad (11.9)$$

– приведений квазірівень Фермі для електронів. І аналогічно для дірок –

$$p = N_V \Phi_{1/2}(-\tilde{\eta}^* - \varepsilon_i) = N_V \Phi_{1/2}(\eta^*) \quad (11.10)$$

із

$$\tilde{\eta}^* = \frac{F_p^* - E_C}{kT}, \quad \eta^* = \frac{E_V - F_p^*}{kT} \quad (11.11)$$

приведеним квазірівнем Фермі  $\eta^*$  для дірок ( $\varepsilon_i = E_g/kT$ ). Добуток концентрацій електронів і дірок для нерівноважного стану відрізняється від добутку для рівноважного стану. Дійсно, у неvirодженому напівпровіднику маємо

$$np = n_0 p_0 e^{\xi^* - \eta^*}. \quad (11.12)$$

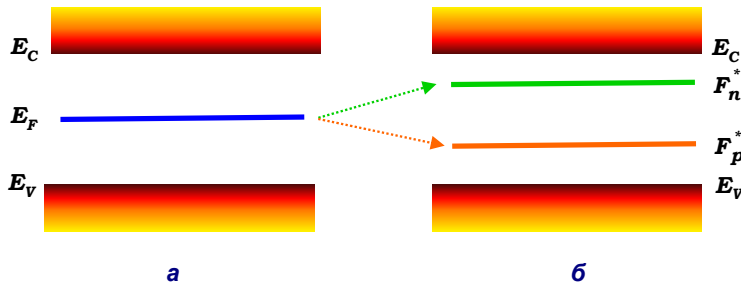
**Відстань між квазірівнями Фермі ( $\xi^* - \eta^*$ ) характеризує відхилення системи від рівноваги, тобто наявність у валентній зоні та зоні провідності нерівноважних носіїв можна описати розщепленням рівня Фермі рівноважного стану та появою нових квазірівнів Фермі для електронів і дірок окремо.**

При цьому чим далі систему виведено від рівноважного стану, тим більшою буде різниця між  $F_n^*$  та  $F_p^*$  (рис. 11.2.). Розглянемо тепер ситуацію, за якої у напівпровіднику протікає електричний струм та отримуємо зв'язок між густиною струму й значенням рівня Фермі. Якщо концентрація носіїв змінюється у просторі, то густина струму визначається не тільки дрейфом частинок, а й дифузією. Нехай коефіцієнтом дифузії електронів є  $D_n$ . Тоді густина струму

$$\mathbf{j}_n = en\mu_n \mathbf{E} + eD_n \nabla n. \quad (11.13)$$

Аналогічно до формули (11.2.6) для діркового компоненту струму маємо

$$\mathbf{j}_p = en\mu_p \mathbf{E} - eD_p \nabla p. \quad (11.14)$$



**Рис. 11.2. Виникнення квазірівнів Фермі за освітлювання напівпровідника  $n$ -типу: а – рівноважного; б – нерівноважного**

Дифузійні струми (як і сам коефіцієнт дифузії) мають сенс тільки у випадку, якщо зміна концентрації частинок на довжині вільного пробігу є малою, тобто за умов виконання нерівності

$$|\nabla n| l \ll n. \quad (11.15)$$

Для невиродженого напівпровідника, що знаходиться в полі потенціалу  $\varphi$ , концентрація електронів визначається формулою

$$n = n_0 e^{\frac{e\varphi}{kT}}. \quad (11.16)$$

Звідси маємо для градієнта концентрації

$$\nabla n = n \frac{e}{kT} \nabla \varphi = -n \frac{e}{kT} \mathbf{E}. \quad (11.17)$$

Тоді умова (11.15) запишеться як



$$\frac{e|\mathbf{E}|l}{kT} \ll 1, \quad (11.18)$$

тобто введення коефіцієнту дифузії припустимо за умов, що енергія, яку отримує носій від зовнішнього поля на довжині вільного пробігу, є набагато меншою за характерну теплову енергію кристала. Рухливість і коефіцієнт дифузії не є незалежними величинами, оскільки для даного типу частинок із заданою ефективною масою рухливість залежить тільки від середнього часу вільного пробігу (часу релаксації). Тією самою величиною визначається й коефіцієнт дифузії. Тепер зосередимося на зв'язку між рухливістю та коефіцієнтом дифузії носіїв у напівпровіднику.

### 11.2.1. Співвідношення Ейнштейна

Розглянемо електрони у невідродженому напівпровіднику за наявності градієнта концентрації в умовах термодинамічної рівноваги (тобто за відсутності струму). Тоді, підставляючи (11.17) до (11.14) за умови  $\mathbf{j} = 0$ , отримуємо відоме співвідношення Ейнштейна

$$\frac{\mu_n}{D_n} = \frac{e}{kT}. \quad (11.19)$$

Дану формулу можна застосувати до будь-якої системи частинок, що утворюють невідроджений газ. Її універсальність полягає в тому, що за відомої рухливості частинок можна обчислити коефіцієнт дифузії, безпосередні виміри якого є непростою експериментальною проблемою. Зрозуміло, що (11.19) можна отримати й для дірок

$$\frac{\mu_p}{D_p} = \frac{e}{kT}. \quad (11.20)$$

Для узагальнення формули Ейнштейна на випадок доволно виродженого газу необхідно згадати, що концентрація електронів є функцією безрозмірного хімічного потенціалу

$$\xi^* = \xi_0^* + \frac{e\varphi}{kT} = \frac{E_F - E_C}{kT} + \frac{e\varphi}{kT}, \quad (11.21)$$

яка визначається формулою

$$n = N_C \Phi_{1/2}(\xi^*). \quad (11.22)$$

Тоді

$$\nabla n = \frac{dn}{d\xi^*} \frac{e}{kT} \nabla \varphi = -\frac{dn}{d\xi^*} \frac{e}{kT} \mathbf{E}. \quad (11.23)$$

Підставимо цей вираз до (11.13), і за умови сумарного струму, рівного нулю, отримаємо

$$n\mu_n - \frac{e}{kT} D_n \frac{dn}{d\xi^*} = 0. \quad (11.24)$$

Враховуючи, що

$$\frac{1}{n} \frac{dn}{d\xi^*} = \frac{d \ln n}{d\xi^*},$$

із (11.24) отримаємо остаточно

$$\frac{\mu_n}{D_n} = \frac{e}{kT} \frac{d \ln n}{d\xi^*}. \quad (11.25)$$

Зауважимо, що у випадку невідродженого електронного газу

$$n = N_c e^{\xi^*}, \quad \frac{d \ln n}{d\xi^*} = 1$$

(11.25) переходить у  $\mu_n/D_n = e/kT$ . Для дірок співвідношення, аналогічне (11.25), має вигляд

$$\frac{\mu_p}{D_p} = \frac{e}{kT} \frac{d \ln p}{d\eta^*}, \quad \eta^* = \frac{E_V - E_F}{kT} - \frac{e\varphi}{kT}. \quad (11.26)$$

У випадку анізотропних систем необхідно враховувати, що рухливість і коефіцієнт дифузії є тензорами. Тоді замість (11.25) запишемо

$$\mu_{ij}^{(n)} = \frac{e}{kT} \frac{d \ln n}{d\xi^*} D_{ij}^{(n)}. \quad (11.27)$$

Як відомо із загальних принципів статистичної фізики, у стані термодинамічної рівноваги електрохімічний потенціал (тобто рівень Фермі) має однакове значення у всіх частинах системи

$$E_F = \text{const}. \quad (11.28)$$

З'ясуємо сенс цього факту щодо напівпровідника. Припустимо, що рівень Фермі змінюється у просторі, і на напівпровідник діє електричне поле так, що потенціал і концентрація електронів залежать від координат. Тоді енергетичні зони будуть викривлені, і хімічний потенціал змінюватиметься у просторі

$$E_F - E_C = \xi(\mathbf{r}). \quad (11.29)$$

Знайдемо тепер повну густину струму. Оскільки концентрація  $n$  є функцією безрозмірного хімічного потенціалу  $\xi^* = \xi/kT$ , то для градієнта концентрації можна записати

$$\nabla n = \frac{dn}{d\xi^*} \nabla \xi^* = \frac{1}{kT} \frac{dn}{d\xi^*} \nabla \xi. \quad (11.30)$$

Враховуючи цю формулу, запишемо густину електронного струму

$$\mathbf{j}_n = \mu_n n \left[ -e \nabla \varphi + \frac{e}{kT} \frac{D_n}{\mu_n} \frac{d \ln n}{d\xi^*} \nabla \xi \right]. \quad (11.31)$$

Тепер використаємо узагальнене співвідношення Ейнштейна (11.25) та отримаємо густину електронного струму

$$\mathbf{j}_n = \mu_n n \left[ -e \nabla \varphi + \frac{e}{kT} \left( \frac{e}{kT} \frac{d \ln n}{d \xi^*} \right)^{-1} \frac{d \ln n}{d \xi^*} \nabla \xi \right] = \mu_n n \nabla (\xi - e \varphi). \quad (11.32)$$

Згадаємо, що  $\xi - e \varphi = E_F$ . Тоді замість (11.32) отримаємо

$$\mathbf{j}_n = \mu_n n \nabla E_F, \quad (11.33)$$

і аналогічно для дірок

$$\mathbf{j}_p = \mu_p p \nabla E_F. \quad (11.34)$$

Таким чином, умови рівноваги (11.28) означають відсутність струму, тобто струм є порушенням термодинамічної рівноваги, і його повна густина має визначатися градієнтом квазірівня Фермі. Для біполярного напівпровідника у нерівноважних умовах, виходячи із (11.33) та (11.34), можна записати

$$\mathbf{j}_n = \mu_n n(\mathbf{r}) \nabla F_n^*, \quad (11.35)$$

$$\mathbf{j}_p = \mu_p p(\mathbf{r}) \nabla F_p^*. \quad (11.36)$$

Іншими словами, за наявності струму квазірівні Фермі змінюються у просторі тим швидше, чим меншими є локальні значення концентрації носіїв.

### 11.2.2. Експериментальне визначення коефіцієнта дифузії за Хайнсом–Шоклі

Для вивчення процесів дифузії і дрейфу у напівпровідниках Хайнс і Шоклі запропонували оригінальну схему експерименту (рис. 11.3): до однорідного видовженого германієвого зразка  $n$ -типу прикладалось статне електричне поле. До поверхні кристала приєднувалися два контакти для створення надлишкової концентрації неосновних носіїв (емітер) і реєстрації "хмари" неосновних носіїв, що дрейфувала під дією електричного поля (колектор). Під дією напруги, що подавалась на контакт метал-напівпровідник, відбувалась інжекція неосновних носіїв заряду<sup>3</sup>. До імпульсного генератора, що видає прямокутні імпульси, приєднувалися емітер та осцилограф для контролю форми розподілу інжекттованих дірок. Короткий імпульс струму через контакт спрямлення створював у кристалі різко обмежену область надлишкової концентрації нерівноважних носіїв (дірок) – "хмару", що зсувалась у напрямку поля. Протягом часу  $t$ , що відповідає часу переміщення "хмари" дірок від емітера до колектора внаслідок невпорядкованого дифузійного руху носіїв, ширина імпульсу зростатиме (рис. 11.4).

<sup>3</sup> Явище інжекції неосновних носіїв заряду детально розглядається у п. 11.3 та 14.4.

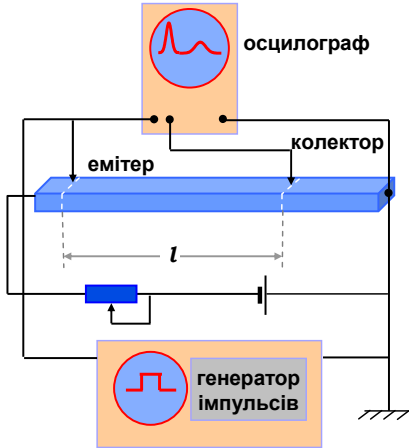


Рис. 11.3. Схема експерименту Хайнса-Шоклі

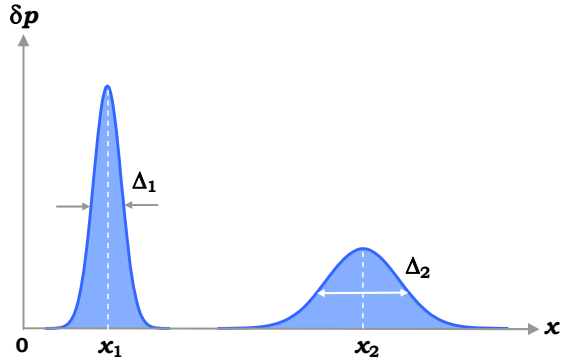


Рис. 11.4. Зміна профілю "хмари" неосновних носіїв за дрейфу "хмари" від емітера ( $x_1$ ) до колектора ( $x_2$ ).

$\Delta_1$  та  $\Delta_2$  – напівширини профілів розподілу носіїв на емітері та колекторі, відповідно

Якщо знехтувати рекомбінацією, то профіль розподілу концентрації дірок зі зміною часу за відсутності поля описується рівнянням<sup>4</sup>

$$\frac{\partial \delta p(x, t)}{\partial t} = D_p \frac{\partial^2 \delta p(x, t)}{\partial x^2}, \quad (11.37)$$

розв'язком якого є

$$\delta p(x, t) = \frac{S_p}{2\sqrt{\pi t D_p}} e^{-x^2/(4D_p t)}, \quad (11.38)$$

де  $S_p$  – поверхнева концентрація дірок, що інжектують у зразок у початковий момент часу  $t = 0$ . У формулі (11.38) експонента описує розширення пакету із часом завдяки дифузії. Передекспоненційний фактор указує на зменшення амплітуди імпульсу зі зростанням часу. Прикладення слабкого електричного поля  $E \ll (D_p/\mu_p) \cdot (1/p) \cdot (dp/dx)$  змушує "хмару" неосновних носіїв заряду рухатися в напрямку поля зі швидкістю  $v_{dr} = \mu_p E$ , тобто максимум розподілу у момент часу  $t_{dr}$  буде зсунутий уздовж напрямку поля на відстань  $t_{dr} / v_{dr}$ . Тоді профіль сигналу, що знімається на колекторі, розташованому у точці  $l$ , описується формулою

$$\delta p(x, t) = \frac{S_p}{2\sqrt{\pi t D_p}} \exp \left[ -\frac{(1-t/t_{dr})^2 l^2}{4D_p t} \right]. \quad (11.39)$$

<sup>4</sup> Детальніше вивід рівняння типу (11.37) та аналіз його розв'язків див. у 11.7.

Нехай амплітудою імпульсу в момент часу  $t_{dr}$  є  $\delta p_m$ . Перейдемо до системи координат із центром у точці  $x_{dr} = t_{dr}/v_{dr}$  та отримаємо миттєвий розподіл концентрації дірок, що описується (11.38), де  $x$  – відхилення від координати максимуму. Якщо позначити через  $\Delta x$  напівширину імпульсу, тобто  $\delta p(\Delta x/2) = \delta p_m/2$ , то

$$\frac{\delta p(\Delta x/2)}{\delta p_m} = \exp\left[-\frac{\Delta x^2}{4} \cdot \frac{1}{4D_p t_{dr}}\right]. \quad (11.40)$$

Якщо вважати, що протягом часу  $\Delta t$  "хмара" інжектованих дірок рухається зі сталою швидкістю  $v_{dr} = l/t_{dr}$ , то величину  $\Delta x$  можна визначити як

$$\Delta x = \Delta t l / t_{dr}. \quad (11.41)$$

Підставимо  $\Delta x$  до (11.40) та отримаємо

$$D_p = \frac{(l\Delta t)^2}{16 \ln 2 \cdot t_{dr}^3}. \quad (11.42)$$

Схема експерименту Хайнса–Шоклі дозволяє також експериментально перевірити співвідношення Ейнштейна. Дійсно, підставляючи до (11.41) величину  $\Delta t$  із (11.42), отримуємо, що напівширина імпульсу в момент часу  $t_{dr}$  визначається співвідношенням

$$\Delta x = 4\sqrt{\ln 2 D_p t_{dr}}. \quad (11.43)$$

Довжина дрейфу в полі  $E$

$$l = \mu_p E t_{dr}. \quad (11.44)$$

Оскільки напруженість поля для видовженого зразка дорівнює напрузі між емітером і колектором, що поділена на  $l$ , то

$$l = \sqrt{\mu_p U t_{dr}}. \quad (11.45)$$

Із (11.43) та (11.44) отримуємо

$$\frac{D_p}{\mu_p} = \frac{U}{16 \ln 2} \left(\frac{\Delta x}{l}\right)^2. \quad (11.46)$$

Оскільки величини у правій частині співвідношення (11.46) вимірюються експериментально, то воно дозволяє в принципі експериментально перевірити співвідношення Ейнштейна. Однак на практиці краще оперувати не величинами розмірності довжини, а часовими проміжками, адже вони легко вимірюються за допомогою осцилографа. У такому випадку перевірка співвідношення Ейнштейна зводиться до експериментальної перевірки співвідношення (див. додаток F)

$$\frac{D_p}{\mu_p} = \frac{U}{16 \ln 2} \frac{\Delta t_1^2 - \Delta t_2^2}{t_1(t_1 - t_2)}, \quad (11.47)$$

де  $\Delta t_1$  – напівширина імпульсу, що відповідає половині амплітуди за час дрейфу  $t_1$ ,  $\Delta t_2$  – напівширина імпульсу за час дрейфу  $t_2$ . Варто зауважити: оскільки відношення  $D_p/\mu_p$  залежить від різниці квадратів вимірюваних величин, то одиничні вимірювання не дають достатньо надійних результатів. Для забезпечення більшої точності експерименту застосовують багаторазові вимірювання, результати яких у подальшому підлягають статистичній обробці. Численні вимірювання показали, що із точністю, кращою за 5 %, співвідношення Ейнштейна виконуються.

### ■ 11.3. Дифузія та дрейф нерівноважних носіїв заряду в монополярному напівпровіднику

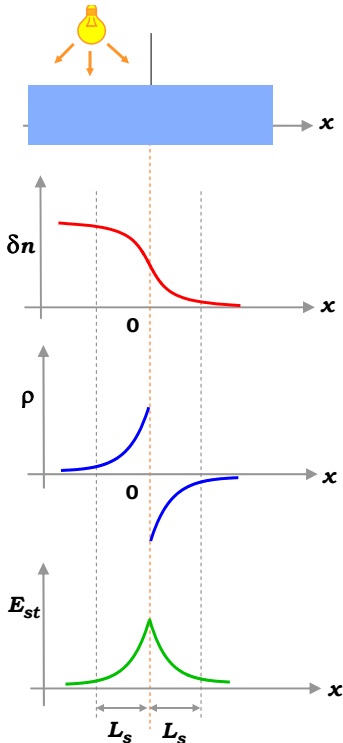


Рис. 11.5. Розподіл надлишкових носіїв, густина заряду та електричне поле на межі освітленої та неосвітленої областей напівпровідника

Розглянемо дифузію та дрейф нерівноважних носіїв заряду в монополярному напівпровіднику, коли вільні носії заряду виникають тільки завдяки іонізації домішок. Для визначеності візьмемо довгий однорідний зразок напівпровідника  $n$ -типу, частина якого ( $x < 0$ ) освітлена (рис. 11.5). У цій частині напівпровідника внаслідок іонізації донорів і переходу електронів із донорних рівнів до зони провідності відбувається генерація електронів. Концентрація нерівноважних електронів  $n$  визначається концентрацією рівноважних електронів  $n_0$  і надлишкових електронів  $\delta n$ , які емітували до зони провідності з донорних рівнів. За відсутності в напівпровіднику центрів захоплювання рівняння електронейтральності матиме вигляд

$$\delta n = \Delta N_d^+. \quad (11.48)$$

Концентрація електронів в освітленій частині зразка перевищує їх концентрацію у неосвітленій. У результаті нерівноважні електрони дифундують з

освітленої частини напівпровідника до неосвітленої, відбувається локальне порушення електронейтральності та виникає об'ємний заряд, що спричиняє виникнення електричного поля. Оскільки в освітленій частині зразка концентрація електронів зменшується, а в неосвітленій – збільшується, то освітлена частина буде заряджена позитивно, а неосвітлена – негативно. Ці заряди індукують електричне поле  $E_{st}$ , напрямок дії якого заважатиме дифузії нерівноважних електронів. У результаті встановиться стаціонарний стан, за якого сумарний струм електронів – його дрейфова (під дією електричного поля  $E_{st}$ ) та дифузійна складові – дорівнюватимуть нулю. Таким чином, виникнення дифузного струму  $j_{diff} = eD_n \nabla_x n$  викликає появу електростатичного поля, яке, у свою чергу, спричиняє появу дрейфового струму  $j_{dr} = en\mu_n E_{st}$ . Обидва процеси приводять до встановлення стаціонарного стану, який характеризується відсутністю електричного струму. Таким чином можна стверджувати, що повний струм  $j_{tot} = j_{diff} + j_{dr}$  має дорівнювати нулю. Звідси маємо рівняння

$$eD_n \frac{dn}{dx} = -en\mu_n E_{st}. \quad (11.49)$$

Використовуючи співвідношення Ейнштейна, отримуємо

$$E_{st} = -\frac{kT}{en} \frac{dn}{dx}. \quad (11.50)$$

Таким чином, на межі тінь–світло виникає електрорушійна сила Дембера, яку можна визначити через поле  $E_{st}$  за допомогою стандартної формули

$$U_D = \oint_C \mathbf{E}_{st} d\mathbf{l}, \quad (11.51)$$

де інтегрування ведеться за деяким замкненим контуром, що включає межу розподілу. Зважаючи на те, що  $n = n_0 + \delta n$ , і у випадку малих рівнів оптичного збудження, коли  $n_0 \gg \delta n$ , із (11.50) маємо

$$E_{st} = -\frac{kT}{en_0} \frac{d\delta n}{dx}. \quad (11.52)$$

Після диференціювання (11.52) за координатою скористаємось рівнянням Пуассона  $dE_{st}/dx = -(4\pi e/\epsilon)\delta n$ , де  $-e\delta n$  – об'ємний заряд у неосвітленій частині напівпровідника. Знайдемо

$$\frac{d^2\delta n}{dx^2} - \frac{4\pi e^2 n_0}{\epsilon kT} \delta n = 0. \quad (11.53)$$

Введемо параметр розмірності довжини

$$L_s = \sqrt{\epsilon kT / 4\pi e^2 n_0}. \quad (11.54)$$

Тоді (11.53) перепишеться у вигляді диференційного рівняння

$$\frac{d^2\delta n}{dx^2} - \frac{1}{L_s}\delta n = 0, \quad (11.55)$$

яке має загальний розв'язок

$$\delta n = C_1 e^{x/L_s} + C_2 e^{-x/L_s}. \quad (11.56)$$

Для неосвітленої області зразка, де при  $x \rightarrow +\infty$ ,  $\delta n = 0$  розв'язком буде

$$\delta n = \delta n(0) e^{-x/L_s}. \quad (11.57)$$

Це означає, що для напівпровідника з монополярним типом провідності концентрація надлишкових носіїв заряду в неосвітленій частині зразка зменшується експоненційно зі сталою спадання  $L_s$  при заглибленні до неосвітленої частини. Величина  $L_s$  називається довжиною екранування (або дебаєвським радіусом екранування) і, як видно із (11.54), визначається концентрацією основних носіїв. Довжина екранування може варіюватись у широкому інтервалі значень залежно від питомої провідності матеріалу та складати величину  $\sim 10^{-4} - 10^{-6}$  см.

## ■ 11.4. Час релаксації Максвелла

Вище було введено параметр  $L_s$ , що має сенс характерної довжини, на якій надлишкова концентрація носіїв заряду зменшується при заглибленні до неосвітленої частини зразка. Використовуючи співвідношення Ейнштейна, перепишемо цей вираз у вигляді

$$L_s = \sqrt{\frac{\epsilon k T}{4\pi e^2 n_0}} = \sqrt{\frac{D_n}{\mu_n} \frac{\epsilon}{4\pi e n_0}} = \sqrt{D_n \tau_M}, \quad (11.58)$$

де  $\tau_M = \epsilon/4\pi\sigma_0$  – максвеллівський час релаксації. Згадаємо, що при розгляді релаксації надлишкової концентрації нерівноважних носіїв після вимкнення оптичного збудження ми з'ясували: вона відбувається із характерним часом  $\tau_f$  (див. (11.42)) і після вимкнення збуджуючого світла зменшується в часі за законом

$$\delta n = \delta n(0) e^{-t/\tau_f}. \quad (11.59)$$

Отже:

**Розповсюдження носіїв заряду в монополярному напівпровіднику на відстань довжини екранування відбувається протягом максвеллівського часу релаксації, що відіграє роль ефективного часу встановлення дифузійно-дрейфової рівноваги.**



Можна також розглянути прямокутний зразок із матеріалу, що характеризується питомою провідністю  $\sigma_0$ , електричний опір якого можна записати у вигляді

$$R = \frac{1}{\sigma_0} \frac{l}{S}, \quad (11.60)$$

де  $l$  – довжина зразка,  $S$  – його поперечний переріз. Ємність такого зразка можна записати як  $C = \epsilon S / 4\pi l$ . Тоді резонансною частотою коливань електромагнітного поля у зразку (частота власної моди коливань електромагнітного поля) буде  $\Omega_r = 1/RC = 4\pi\sigma_0/\epsilon = 1/\tau_M$ , тобто  $\tau_M$  є сталою часу даного матеріалу. А оскільки коливання електромагнітного поля в розглянутій моделі викликано наявністю носіїв заряду, то характерні періоди коливань поля є ні чим іншим, як часом встановлення дифузійно-дрейфової рівноваги в матеріалі. Характерні значення максвеллівського часу релаксації – від  $10^{-7}$  до  $10^{-12}$  сек.

### ■ 11.5. Дифузія та дрейф неосновних надлишкових носіїв заряду в домішковому напівпровіднику

Для з'ясування особливостей дифузійного та дрейфового руху неосновних носіїв заряду в домішковому напівпровіднику розглянемо видовжений зразок напівпровідника  $n$ -типу ( $n_0 \gg p_0$ ), розташований у зовнішньому електричному полі напруженості  $\mathbf{E}_{ext}$  (рис. 11.6). Вважатимемо при цьому, що вузька область напівпровідника ( $-l \leq x \leq 0$ ) опромінюється світлом із довжиною хвилі  $\lambda \leq 2\pi\hbar c/E_{ext}$  таким чином, що по всьому об'єму освітленої області зразка відбувається рівномірна генерація електронно-діркових пар за рахунок переходів **зона-зона**. Прийнемо інтенсивність освітлювання не дуже великою, а концентрацію основних носіїв набагато більшою за надлишкову концентрацію неосновних носіїв заряду  $\delta n$ . В освітленій частині, де відбувається генерація електронно-діркових пар ( $\delta n = \delta p$ ), концентрація нерівноважних дірок  $p = p_0 + \delta p$  буде істотно більшою за їхню концентрацію в неосвітленій частині напівпровідника, завдяки чому виникне дифузія дірок з освітленої області до неосвітленої.

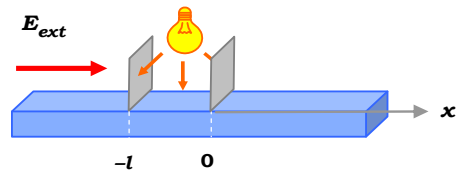


Рис. 11.6. Визначення дифузійного та дрейфового руху неосновних носіїв заряду

У неосвітленій частині зразка виникне об'ємний позитивний заряд, який утворить статичне електричне поле, що викличе перерозподіл носіїв заряду: до області з освітленої частини зразка протягом часу життя надлишкових нерівноважних дірок перейдуть електрони, які компенсують позитивний заряд надлишкових дірок. Таким чином, завдяки дифузії надлишкових неосновних носіїв заряду (дірок) углиб електронного напівпровідника виникне рух рівної кількості основних носіїв заряду (електронів): об'ємний заряд не створюється, і дифузія зарядів відбувається як дифузія квазінейтрального утворення нерівноважних носіїв із однакової кількості основних і неосновних носіїв заряду. За мірою просування цього нейтрального утворення вглиб зразка електрони та дірки рекомбінуватимуть, і їхня концентрація за зростання відстані від області освітлення зменшуватиметься.

Припустимо, що на такий зразок напівпровідника діє зовнішнє однорідне електричне поле  $E_{ext}$ . Тоді дірки, що прискорюються зовнішнім полем, затягують за собою й електрони. При цьому концентрація електронів дорівнюватиме концентрації дірок. Якщо ж поле направлене так, що загальмовуватиметься дифузійний рух дірок, разом із дірками гальмуватимуться й електрони. Таким чином, аналогічно із попереднім випадком, у зовнішньому полі в напівпровіднику електронного типу квазінейтральне утворення нерівноважних носіїв заряду рухатиметься в напрямку руху надлишкових дірок зі швидкістю дрейфу неосновних носіїв. Нехай зовнішнє поле створює електричний струм  $J$ . Тоді зміна концентрації дірок у довільній точці зразка задовольняє рівнянню неперервності

$$\frac{\partial p}{\partial t} = G - \frac{1}{e} \frac{\partial J_p}{\partial x} - \frac{p - p_0}{\tau_p}, \quad (11.61)$$

де  $G$  – швидкість генерації електронно-діркових пар,  $\tau_p$  – час життя нерівноважних надлишкових дірок. Як зміниться концентрація неосновних носіїв у неосвітленій області зразка у стаціонарному стані? У цьому випадку в (11.61) необхідно покласти  $\partial p / \partial t = 0$  та  $G = 0$ . У результаті отримаємо

$$-\frac{1}{e} \frac{\partial J_p}{\partial x} - \frac{\delta p}{\tau_p} = 0. \quad (11.62)$$

Обчислимо тепер похідну від діркового струму, вважаючи, що діюче у зразку поле майже не відрізняється від зовнішнього

$$\frac{\partial}{\partial x} \left( e p \mu_p E_{ext} - e D_p \frac{\partial p}{\partial x} \right) = e \mu_p E_{ext} \frac{\partial p}{\partial x} - e D_p \frac{\partial^2 p}{\partial x^2}. \quad (11.63)$$

Після підстановки отриманого виразу до (11.62) і введення позначення довжини, на яку дифундує неосновний носій заряду за час максвеллівської релаксації

$$L_D = \sqrt{\tau_p D_p}, \quad (11.64)$$

а також довжини дрейфу неосновного носія в полі  $E_{ext}$  за час  $\tau_p$

$$L_E = \tau_p \mu_p E_{ext}, \quad (11.65)$$

отримаємо рівняння

$$\frac{d^2 \delta p}{dx^2} - \frac{L_E}{L_D^2} \frac{d\delta p}{dx} - \frac{\delta p}{L_D^2} = 0. \quad (11.66)$$

Розв'яжемо його окремо для областей  $x \leq -l$  і  $x \geq 0$ . Розв'язки в цих областях шукаємо у вигляді

$$\delta p = C e^{\alpha x}, \quad (11.67)$$

де  $\alpha$  – розв'язок характеристичного рівняння

$$\alpha^2 - \frac{L_E}{L_D^2} \alpha - \frac{1}{L_D^2} = 0, \quad (11.68)$$

а коефіцієнт  $C$  визначається із граничних умов. Розв'язками характеристичного рівняння є

$$\alpha_1 = 1/L_1, \quad L_1 = \frac{2L_D^2}{L_E + \sqrt{L_E^2 + 4L_D^2}}, \quad (11.69)$$

$$\alpha_2 = -1/L_2, \quad L_2 = \frac{2L_D^2}{\sqrt{L_E^2 + 4L_D^2} - L_E}.$$

Враховуючи, що в неосвітленій частині зразка концентрація носіїв може лише зменшуватись щодо стаціонарної концентрації, сформованої в освітленій області, в області  $x \leq -l$  граничні умови запишемо у вигляді

$$\delta p|_{x=-l} = \delta p(0), \quad \delta p|_{x \rightarrow -\infty} = 0. \quad (11.70)$$

Звідси отримаємо

$$\delta p = \delta p(0) e^{(x+l)/L_1}. \quad (11.71)$$

Аналогічно для області  $x \geq 0$  маємо

$$\delta p|_{x=0} = \delta p(0), \quad \delta p|_{x \rightarrow +\infty} = 0 \quad (11.72)$$

і

$$\delta p = \delta p(0) e^{-x/L_2}. \quad (11.73)$$

У результаті з'ясуємо, що по обидва боки від освітленої частини зразка концентрація неосновних носіїв заряду зменшується за експоненціальним законом зі сталою затухання  $L_1$  – для області  $x \leq -l$ , і  $L_2$  – для області  $x \geq 0$  (рис. 11.7 а).

Якщо зовнішнє електричне поле відсутнє, то струм забезпечується тільки дифузією надлишкових дірок, і концентрація неосновних носіїв заряду змінюватиметься з відстанню за законом

$$\delta p = \delta p(0) \cdot \begin{cases} e^{(x+l)/L_D}, & x \leq -l \\ e^{-x/L_D}, & x \geq 0 \end{cases}, \quad (11.74)$$

тобто експоненціальний спад концентрації неосновних носіїв заряду буде симетричним з обох боків освітленої області та визначатиметься величиною (рис. 11.7 б)

$$L_D^p = \sqrt{\tau_p D_p}. \quad (11.75)$$

Для напівпровідника *p*-типу неосновними носіями будуть електрони, спад концентрації яких визначатиметься дифузійною довжиною

$$L_D^e = \sqrt{\tau_e D_e}, \quad (11.76)$$

яка може досягати значень  $\sim 10^{-1}$  см, що набагато перевищує довжину екранування основних носіїв.

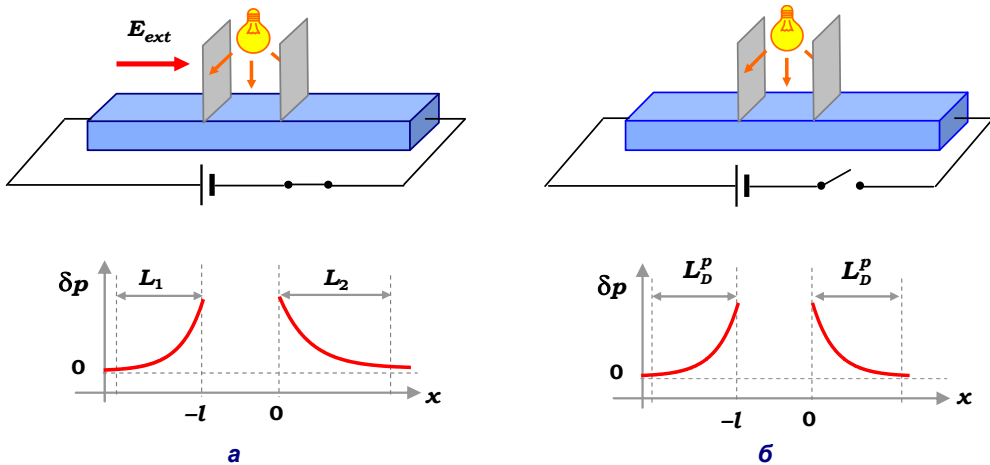


Рис. 11.7. Розподіл концентрації надлишкових носіїв уздовж зразка: а – під дією зовнішнього електричного поля; б – за його відсутності

## ■ 11.6. Дифузія неосновних носіїв заряду в електричному полі

Раніше ми з'ясували, що за відсутності зовнішнього електричного поля концентрація надлишкових неосновних носіїв заряду зменшується симетрично по обидва боки від освітленої області зразка. Зро-

зуміло, що наявність зовнішнього поля призведе до порушення такої симетрії. При визначенні поведінки концентрації надлишкових неосновних носіїв заряду ми припускали, що локальне поле, яке діє на носії, не відрізняється від зовнішнього. У реальних ситуаціях локальне поле відрізняється від зовнішнього. Особливо ця різниця виявляється у випадку, коли лінійні розміри частинки в напрямку зміни поля порівнянні із характерними масштабами зміни локального поля. Оскільки зміни локального поля у просторі зумовлені дрейфом надлишкових носіїв заряду і відбуваються на масштабах порядку  $L_D$ , то для нехтування дією поля, обумовленого наявністю неоднорідного розподілу носіїв, лінійний розмір зразка вздовж напрямку зовнішнього поля має бути набагато більшим за  $L_D$ . Ясно, що використане припущення також виконуватиметься у випадку сильного зовнішнього поля. Припустимо: зовнішнє поле настільки велике, що  $|L_E| > 2L_D$ . Тоді, якщо вектор  $\mathbf{E}$  направлений вздовж вісі  $OX$ ,

$$L_1 = \frac{2L_D^2}{\sqrt{L_E^2 + 4L_D^2} + L_E} < L_D < L_2 = \frac{2L_D^2}{\sqrt{L_E^2 + 4L_D^2} - L_E}. \quad (11.77)$$

Це означає, що наявність зовнішнього поля приводить до різної величини довжини затягування носіїв до неосвітлених областей – довжина затягування в напрямку вздовж поля  $L_2$  (до неосвітленої області  $x > 0$ ) є більшою за довжину затягування у протилежному напрямку  $L_1$  (до неосвітленої області  $x < -l$ ). Ясно, що зміна напрямку вектора зовнішнього поля надасть протилежний результат – величина  $L_1$  буде більшою за  $L_2$ . Визначимо довжини затягування  $L_1$  та  $L_2$  у випадку сильного зовнішнього поля, яке направлено вздовж вісі  $OX$ . В області  $x > 0$  за умови  $L_E \gg L_D$  маємо ( $\beta = 2L_D/L_E \ll 1$ )

$$L_2 = \frac{L_D\beta}{\sqrt{1+\beta^2} - 1} \approx \frac{L_D\beta}{2\beta^2} = L_E \quad (11.78)$$

– довжину затягування, що дорівнює дрейфовій довжині. Тоді розподіл надлишкової концентрації неосновних носіїв

$$\delta p = \delta p(0)e^{-x/(\tau p \mu_p E_{ext})}. \quad (11.79)$$

Тобто так само, як й у випадку винятково дифузійного руху носіїв надлишкова концентрація неосновних носіїв спадає з відстанню від освітленої області експоненційно, але вже зі сталою спадання  $L_E$ . Порівнювання (11.79) та (11.74) свідчить, що в неосвітленій області  $x > 0$  при  $L_E^2 \gg 4L_D^2$  (в умовах дії на напівпровідник сильного зовнішнього електричного поля) надлишкова концентрація неосновних носіїв є істотно більшою, ніж за відсутності зовнішнього поля. Таким чином, при про-

ходженні струму, обумовленого сильним електричним полем, надлишкові дірки в напівпровіднику  $n$ -типу зтягуються до області  $x > 0$ , і напівпровідник локально збагачується неосновними носіями заряду в істотно більшій кількості, ніж за наявності тільки дифузії, коли  $E_{ext} = 0$  (рис. 11.8 а). У такому випадку кажуть, що відбувається *інжекція неосновних носіїв заряду*, підкреслюючи при цьому збільшення концентрації неосновних носіїв під дією електричного поля. Зрозуміло, що для напівпровідникового зразка  $p$ -типу інжекція в область  $x > 0$  відбуватиметься при  $E_{ext} < 0$ .

В області  $x < -l$  в умовах дії на зразок сильного електричного поля, що направлене вздовж вісі ОХ ( $E_{ext} > 0$ ), довжина зтягування

$$L_1 = \frac{L_D \beta}{\sqrt{1 + \beta^2} + 1} \approx \frac{L_D \beta}{2 + \beta^2/2} = \frac{L_D^2}{L_E}, \quad (11.80)$$

а концентрація неосновних носіїв заряду визначається формулою

$$\delta p = \delta p(0) e^{(x+l)L_D^2/L_E}, \quad (11.81)$$

тобто дія зовнішнього електричного поля викликає збіднення неосновними носіями напівпровідника в області  $x < -l$ . Це явище називається *ексклюзією неосновних носіїв заряду*. Для напівпровідника  $p$ -типу таке явище в області  $x < -l$  спостерігатиметься за умови, що поле направлене проти напрямку вісі ОХ, тобто при  $E_{ext} < 0$ .

Зміна напрямку дії зовнішнього поля приведе до того, що в неосвітленій області зразка  $x > 0$  спостерігатиметься зменшення концентрації надлишкових носіїв (*явище екстракції*), а в області  $x < -l$  – навпаки, збільшення концентрації надлишкових носіїв заряду, тобто *аккумуляція* (рис. 11.8 б).

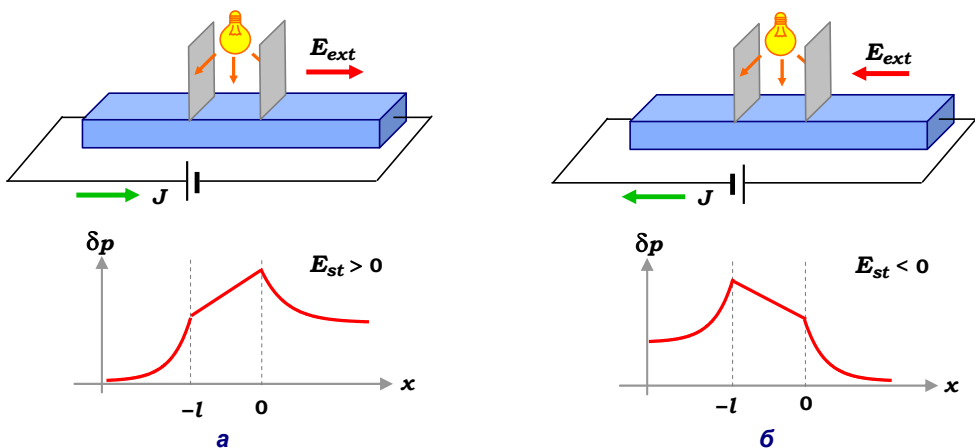


Рис. 11.8. Розподіл концентрації надлишкових носіїв заряду у напівпровіднику  $n$ -типу під дією зовнішнього електричного поля: а – явище інжекції; б – явище екстракції та аккумуляції

## ■ 11.7. Дрейф неосновних надлишкових носіїв заряду

Раніше ми розглядали схему експерименту Хайнса–Шоклі, в якому вивчались особливості дифузійно-дрейфового руху неосновних носіїв за їхньої інжекції через випрямляючий контакт. Тепер розглянемо особливості дифузійно-дрейфового руху неосновних носіїв за їхньої інжекції за допомогою освітлювання зразка, наприклад рух хмари неосновних надлишкових носіїв заряду (для напівпровідника  $n$ -типу це – дірки) під дією зовнішнього електричного поля. Для цього припустимо, що видовжений зразок напівпровідника освітлюється так, що коли в момент часу  $t = 0$  освітлення вимикається, у вузькій області із центром у точці  $x = 0$  хмара неосновних носіїв має гауссову форму розподілу

$$\delta p = C e^{-x^2/x_0^2} \quad (11.82)$$

із напівшириною  $x_0$ . У момент вимкнення освітлення до зразка було прикладене електричне поле  $E_{ext}$ . Тоді завдяки двом типам руху (дифузії і дрейфу) первісна діркова хмара розпливатиметься, одночасно зсуваючись уздовж зразка в напрямку дії електричного поля (рис. 11.9). Рівняння руху розподілу дірок у хмарі має такий вигляд (вимкнене освітлення означає відсутність процесів генерації,  $G = 0$ )

$$\frac{\partial \delta p}{\partial t} - D_p \frac{\partial^2 \delta p}{\partial x^2} + \frac{\delta p}{\tau_p} + \mu_p E_{ext} \frac{\partial \delta p}{\partial x} = 0. \quad (11.83)$$

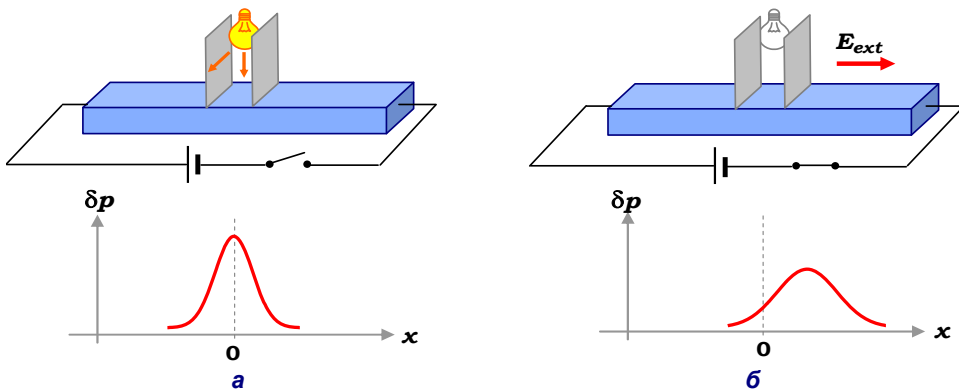


Рис. 11.9. Дифузія і дрейф хмари дірок. Розподіл концентрації надлишкових дірок: а – у момент часу  $t = 0$  (вимкнено освітлення та ввімкнено електричне поле); б – у момент часу  $t \neq 0$

Введемо позначення

$$t_0 = x_0^2/4D_p, \quad (11.84)$$

і перейдемо до нової змінної  $t' = t + t_0$ . (11.85)

При цьому величина  $\delta p$  визначатиметься рівнянням, що формально збігається із (11.83) із заміною  $t$  на  $t'$ . Розв'язок отриманого рівняння шукаємо у вигляді

$$\delta p = \varphi(x, t') e^{-t'/\tau_p}. \quad (11.86)$$

Підставимо цей розв'язок до рівняння та отримуємо

$$\frac{\partial \varphi(x, t')}{\partial t'} + \mu_p E_{ext} \frac{\partial \varphi(x, t')}{\partial x} - D_p \frac{\partial^2 \varphi(x, t')}{\partial x^2} = 0. \quad (11.87)$$

Наступним кроком пошуку залежності щільності розподілу дірок від координати та часу буде введення нової змінної  $x' = x - \mu_p E_{ext} t'$ . При цьому варто пам'ятати, що диференціювання за  $t'$  приведе до появи похідної за координатою

$$\frac{\partial \varphi(x', t')}{\partial t'} \rightarrow \frac{\partial \varphi(x', t')}{\partial t'} - \mu_p E_{ext} \frac{\partial \varphi(x', t')}{\partial x'}. \quad (11.88)$$

Підставляючи похідну (11.88) до (11.87), отримаємо рівняння

$$\frac{\partial \varphi(x', t')}{\partial t'} - D_p \frac{\partial^2 \varphi(x', t')}{\partial x'^2} = 0. \quad (11.89)$$

Із курсу математичної фізики відомо, що подібне рівняння зустрічається в теорії теплопровідності й має розв'язок у вигляді

$$\varphi(x', t') = \frac{C}{\sqrt{t'}} e^{-x'^2/(4D_p t')}. \quad (11.90)$$

Таким чином, у змінних  $x' - t'$  розв'язок (11.7.2) можна записати як

$$\delta p = \frac{C}{\sqrt{t'}} e^{-\frac{x^2}{4D_p t'} - \frac{t'}{\tau_p}}. \quad (11.91)$$

Для визначення нормуючого коефіцієнта  $C$  вважатимемо відомою повну кількість неосновних носіїв заряду у хмарі в момент вимкнення освітлення: нехай вона дорівнює  $P_c$ . Тоді, інтегруючи концентрацію (11.91) за  $x$  від  $-\infty$  до  $+\infty$  за відсутності процесів рекомбінації (що можна робити) та враховуючи великі розміри зразка вздовж всі ОХ і швидке спадання  $\delta p$  зі зростанням  $|x|$ , маємо

$$\frac{C}{\sqrt{t'}} \int_{-\infty}^{\infty} d\zeta \cdot e^{-\zeta^2/(4D_p t')} = P_c. \quad (11.92)$$

Звідси, беручи до уваги, що  $\int_{-\infty}^{\infty} d\zeta \cdot e^{-\zeta^2/(4D_p t')} = \sqrt{4\pi D_p t'}$ , отримуємо



$$C = \frac{P_c}{\sqrt{4\pi D_p}}. \quad (11.93)$$

Таким чином, профіль концентрації хмари неосновних носіїв в умовах дії зовнішнього поля змінюватиметься за законом

$$\delta p(x, t) = \frac{P_c}{\sqrt{4\pi D_p(t+t_0)}} \cdot \exp\left\{-\frac{(x - \mu_p E_{ext}(t+t_0))^2}{4D_p(t+t_0)} - \frac{t+t_0}{\tau_p}\right\}, \quad (11.94)$$

тобто хмара рухається у напрямку електричного поля зі швидкістю  $\mu_p E_{ext}$ , поширюючись при цьому пропорційно кореню квадратному із часу, а її щільність зменшується як  $\exp(-t/\tau_p)$ . За допомогою вимірювання зміни параметрів профілю хмари можна отримати величину рухливості  $\mu_p$ , яка називається амбіполярною дрейфовою рухливістю. Для електронного ( $n \gg p$ ) або діркового ( $n \ll p$ ) напівпровідника ця рухливість збігається із рухливістю дірок або електронів, відповідно, тобто із рухливістю неосновних носіїв заряду. Але для власного або добре компенсованого напівпровідника (коли  $n \approx p$ ) виникає інша ситуація, яку розглянемо далі.

### ■ 11.8. Дифузія та дрейф носіїв заряду в напівпровідниках із провідністю, близькою до власної

Обговорюючи особливості процесу руху носіїв заряду в неоднорідних напівпровідниках, ми свідомо обмежувались розглядом домішкових напівпровідників, де для концентрації носіїв заряду виконувались умови  $n \gg p$  або  $p \gg n$ . При цьому надлишкові концентрації як основних, так і неосновних носіїв вважались однаковими, тобто всюди використовувалася умова  $\delta n = \delta p$ . Крім того, у рівняннях нехтувалось доданком, пропорційним  $\partial E / \partial x$ . Для з'ясування фізичного сенсу цих умов вважатимемо, що генерація носіїв заряду відбувається так, що  $\delta n - \delta p \neq 0$ . Тоді різниця концентрації нерівноважних надлишкових носіїв пропорційна локальній густині заряду, що утворився завдяки процесам дифузії, а рівняння Пуассона запишеться як

$$\operatorname{div} \mathbf{E} = \frac{4\pi e}{\varepsilon} (\delta n - \delta p). \quad (11.95)$$

Тепер запишемо одновимірні рівняння неперервності, зберігаючи доданки, пропорційні похідній електричного поля за координатою,

$$\frac{\partial \delta p}{\partial t} = -\frac{\delta p}{\tau_p} + D_p \frac{\partial^2 \delta p}{\partial x^2} - \mu_p p \frac{\partial E}{\partial x} - \mu_p E \frac{\partial \delta p}{\partial x}, \quad (11.96)$$

$$\frac{\partial \delta n}{\partial t} = -\frac{\delta n}{\tau_n} + D_n \frac{\partial^2 \delta n}{\partial x^2} - \mu_n p \frac{\partial E}{\partial x} - \mu_n E \frac{\partial \delta n}{\partial x}. \quad (11.97)$$

Зважаючи на (11.95), перепишемо рівняння неперервності у вигляді

$$\frac{\partial \delta p}{\partial t} = -\frac{\delta p}{\tau_p} + D_p \frac{\partial^2 \delta p}{\partial x^2} + \frac{4\pi\sigma_p}{\varepsilon} (\delta n - \delta p) - \mu_p E \frac{\partial \delta p}{\partial x}, \quad (11.98)$$

$$\frac{\partial \delta n}{\partial t} = -\frac{\delta n}{\tau_n} + D_n \frac{\partial^2 \delta n}{\partial x^2} + \frac{4\pi\sigma_n}{\varepsilon} (\delta n - \delta p) - \mu_n E \frac{\partial \delta n}{\partial x}. \quad (11.99)$$

У цих рівняннях різницю  $|\delta n - \delta p|$  можна вважати незначною. Але оскільки коефіцієнти перед такими дужками є великими, то доданками, пропорційними  $(\delta n - \delta p)$ , не можна нехтувати. Цей факт і означає, що нехтування роллю об'ємного заряду (тобто використання рівності  $\delta n = \delta p$ ) стає неможливим. Зокрема, це може відбуватися у випадку, коли концентрація електронів і дірок в напівпровіднику складають величини одного порядку. Тоді як змінюється надлишкова концентрація нерівноважних носіїв у часі та просторі (розв'язок рівнянь (11.98) та (11.99)) в умовах, коли концентрації носіїв приблизно однакові? Бажано виключити з них доданки, пропорційні  $(\delta n - \delta p)$ , за допомогою множення першого з рівнянь на  $\sigma_n$ , а другого – на  $\sigma_p$  і додавання одного отриманого рівняння до другого. Унаслідок матимемо

$$\begin{aligned} \sigma_n \frac{\partial \delta p}{\partial t} + \sigma_p \frac{\partial \delta n}{\partial t} = & - \left( \sigma_n \frac{\delta p}{\tau_p} + \sigma_p \frac{\delta n}{\tau_n} \right) + \left( \sigma_n D_p \frac{\partial^2 \delta p}{\partial x^2} + \sigma_p D_n \frac{\partial^2 \delta n}{\partial x^2} \right) + \\ & + \sigma_p \mu_n E \frac{\partial \delta n}{\partial x} - \sigma_n \mu_p E \frac{\partial \delta p}{\partial x}. \end{aligned} \quad (11.100)$$

За подальшого аналізу рівняння вже можна використати той факт, що надлишкові концентрації нерівноважних електронів і дірок майже дорівнюють одна одній, тобто  $\delta n = \delta p$ . Крім того, вважатимемо, що й похідні  $\partial \delta p / \partial t$  і  $\partial \delta n / \partial t$  дорівнюють одна одній. Останнє твердження тягне за собою дуже важливий наслідок: рівність часів рекомбінації електронів і дірок. Дійсно, розглянемо процеси генерації і рекомбінації носіїв за відсутності пасток і невеликих концентрацій носіїв. Це означає, що генераційно-рекомбінаційні процеси відбуваються тільки за механізмом **зона-зона**, тобто електрони та дірки генеруються й рекомбінують парами. Тоді за відсутності зовнішніх впливів кінетику генераційно-рекомбінаційних процесів можна описати рівняннями

$$\frac{d\delta p}{dt} = G_0 - \frac{p}{\tau'_p} = -\frac{\delta p}{\tau_p}, \quad (11.101)$$

$$\frac{d\delta n}{dt} = G_0 - \frac{n}{\tau'_n} = -\frac{\delta n}{\tau_n}. \quad (11.102)$$

Звідси за умови  $\delta\delta p/dt = \delta\delta n/dt$  і  $\delta p = \delta n$  отримуємо  $\tau_p = \tau_n$ . Зважаючи на ці міркування, отримаємо із (11.100)

$$(\sigma_n + \sigma_p) \frac{\partial\delta n}{\partial t} = -(\sigma_n + \sigma_p) \frac{\delta p}{\tau_p} + (\sigma_n D_p + \sigma_p D_n) \frac{\partial^2\delta p}{\partial x^2} - (\sigma_n \mu - \sigma_p \mu_n) E \frac{\partial\delta p}{\partial x}. \quad (11.103)$$

Введемо два коефіцієнти:

♦ **амбіполярної дифузії** 
$$D = \frac{\sigma_n D_p + \sigma_p D_n}{\sigma_n + \sigma_p}; \quad (11.104)$$

♦ **амбіполярної дрейфової рухливості**

$$\mu = \frac{\sigma_n \mu_p - \sigma_p \mu_n}{\sigma_n + \sigma_p}. \quad (11.105)$$

Поділимо обидві частини (11.103) на  $\sigma_n + \sigma_p$  і, використовуючи визначення (11.8.10–10.8.11), знайдемо рівняння, якому підпорядкований рух хмари нерівноважних дірок у сильно компенсованому напівпровіднику

$$\frac{\partial\delta n}{\partial t} = -\frac{\delta p}{\tau_p} + D \frac{\partial^2\delta p}{\partial x^2} - \mu E \frac{\partial\delta p}{\partial x}. \quad (11.106)$$

Порівнювання отриманого рівняння із (11.83) переконує, що обидва вони збігаються із точністю до позначень. Таким чином, можна стверджувати, що в умовах, коли  $n \approx p$ , рух нерівноважних носіїв заряду відбувається за повною аналогією із випадком для напівпровідника  $n$ -типу: під дією зовнішнього електричного поля хмара нерівноважних носіїв заряду рухається вздовж напрямку дії поля зі швидкістю  $v_a = \mu E$ , розпливаючись у просторі пропорційно кореню квадратному з часу, а її щільність зменшується як  $\exp(-t/\tau_p)$ .

Оскільки використання сильно компенсованих напівпровідників і напівпровідників із майже власною провідністю має практичне застосування, розглянемо структуру коефіцієнта амбіполярної дифузії детальніше. По-перше, (11.104) зручно записувати через рухливості носіїв

$$D = \frac{n\mu_n D_p + p\mu_p D_n}{n\mu_n + p\mu_p}, \quad (11.107)$$

або, скориставшись співвідношенням Ейнштейна,

$$D = (n + p) \left( \frac{n}{D_n} + \frac{p}{D_p} \right)^{-1}. \quad (11.108)$$

Звідси видно, що у випадку, коли  $n \approx p$ , коефіцієнт амбіполярної дифузії визначається дифузією тих носіїв, коефіцієнт дифузії яких менший. Оскільки зазвичай у напівпровідниках рухливість дірок менша за рухливість електронів, то  $D$  визначатиметься коефіцієнтом дифузії дірок.

Для визначення фізичного сенсу амбіполярної рухливості (11.105) припустимо, що маємо справу із власним напівпровідником, в якому  $n = p = n_i$ . Видно, що для власного напівпровідника  $\mu = 0$ . Однак, це не означає, що носії у власному напівпровіднику не рухаються, просто  $\mu$  визначається не швидкістю дрейфу частинок у полі, а швидкістю дрейфу збурення концентрації носіїв – квазінейтральної хмари нерівноважних носіїв заряду. У випадку напівпровідника  $n$ -типу  $\mu > 0$ , що означає: хмара нерівноважних носіїв рухається у напрямку, в якому би рухались під дією зовнішнього поля дірки. Якщо ж напівпровідник являє собою матеріал  $p$ -типу, то  $\mu < 0$ , тобто хмара нерівноважних носіїв рухається в напрямку проти поля (як би рухались електрони).

## ■ 11.9. Задачі

1. Доведіть, що співвідношення Ейнштейна є прямим наслідком больцманівського розподілу носіїв.

**Розв'язок.** Припустимо, що система електронів знаходиться в рівновазі у зовнішньому електричному полі  $E$

$$E = -\frac{\partial\varphi}{\partial x}. \quad (11.109)$$

Згідно із законом розподілу невироджених електронів число частинок в одиниці об'єму в деякому перерізі  $x$

$$n(x) = n_0 e^{-\frac{e\varphi}{kT}}. \quad (11.110)$$

З іншого боку, концентрацію частинок на перерізі  $x$  можна визначити з умови балансу дрейфового та дифузійного струмів

$$\mu_n e n E - e D_n \frac{dn}{dx} = 0, \quad (11.111)$$

яке після інтегрування дає  $n(x) = n_0 e^{-\frac{\mu_n \varphi}{D_n}}$ . (11.112)

Порівнюючи (11.110) та (11.112), знаходимо  $D_n/\mu_n = kT/e$ , що й підтверджує: співвідношення Ейнштейна є прямим наслідком больцманівського розподілу носіїв.

2. Нехай закон дисперсії електронів напівпровідника визначається діагональним тензором ефективних мас. Визначте залежність коефіцієнта дифузії від  $m_x$ ,  $m_y$ ,  $m_z$ , вважаючи напівпровідник виродженим.

**Розв'язок.** Розглянемо для визначеності електронний напівпровідник. Використовуючи (11.24), отримаємо

$$D_n = n\mu_n kT \frac{1}{e(dn/d\xi^*)}. \quad (11.113)$$

Концентрація електронів визначається функцією розподілу

$$n = \frac{M}{4\pi^3} \int dk \{1 + \exp((E(k) - E_F)/kT)\}^{-1}, \quad (11.114)$$

де  $M$  – число еквівалентних долин у законі дисперсії електронів напівпровідника. Перейдемо до нових змінних  $t = \hbar/\sqrt{2m_x kT}$ ,  $u = \hbar/\sqrt{2m_y kT}$ ,  $v = \hbar/\sqrt{2m_z kT}$ . Маємо

$$\begin{aligned} n &= \frac{M}{4\pi^3} \frac{\sqrt{2m_x m_y m_z}}{\hbar^3} (kT)^{3/2} \int dt du dv \left\{ 1 + \frac{E_C - E_F}{kT} + t^2 + u^2 + v^2 \right\}^{-1} = \\ &= \frac{M}{\pi^2} \frac{\sqrt{2m_x m_y m_z}}{\hbar^3} (kT)^{3/2} \frac{\sqrt{\pi}}{2} \Phi_{1/2}(\xi), \end{aligned} \quad (11.115)$$

де  $\xi = (E_C - E_F/kT)$ . Для виродженого напівпровідника застосуємо таке наближення для оцінки інтеграла Фермі

$$\Phi_{1,2}(\xi) \approx \xi^{3/2} / \Gamma(3/2) = 2\xi^{3/2} / \sqrt{\pi}. \quad (11.116)$$

Використовуючи цю формулу, із (11.115) отримуємо

$$n = \frac{M}{\pi^2} \frac{\sqrt{2m_x m_y m_z}}{\hbar^3} (kT)^{3/2} \xi^{3/2}. \quad (11.117)$$

Звідси знайдемо

$$\frac{dn}{d\xi} = \frac{3M}{2\pi^2} \frac{\sqrt{2m_x m_y m_z}}{\hbar^3} (kT)^{3/2} \xi^{1/2},$$

або після визначення величини  $\xi$  із (11.117),

$$\frac{dn}{d\xi} = \frac{3}{2} \left( \frac{M}{\pi^2} \right)^{2/3} \frac{(2m_x m_y m_z)^{1/6} n^{1/3}}{\hbar^2} kT. \quad (11.118)$$

Підставимо вираз (11.118) до (11.113) та остаточно знайдемо

$$D_n = \mu_n \frac{2}{3} \left( \frac{\pi^2}{M} \right)^{2/3} \frac{\hbar^2 n^{2/3}}{(2m_x m_y m_z)^{1/6}}. \quad (11.119)$$

**3.** Оцініть максимальні концентрації рівноважних носіїв для Ge, Si та GaAs, за яких співвідношення Ейнштейна є справедливим за кімнатної температури та температури рідкого азоту.

**4.** Зразок власного кремнію неоднорідно легований бором так, що концентрація домішки має таку залежність від координати  $N_B = N_B^0 e^{-ax}$ . Знайдіть величину та

напрямок вбудованого електричного поля у стані термодинамічної рівноваги в області  $N_B > n$ .

5. Експеримент у схемі Хайнса–Шоклі проводиться за досить низького значення поля у зразку та за малої амплітуди імпульсу. Поясніть необхідність виконання цих умов.

6. За допомогою (11.38) покажіть, що повна концентрація інжекттованих неосновних носіїв, зареєстрована на колекторі, не залежить від відстані між емітером і колектором (тобто від часу дрейфу), якщо знехтувати рекомбінацією.

### ***Список літератури***

1. *Зеєгер К.* Физика полупроводников. – М.: Мир, 1977.
2. *Бонч-Бруевич В.Л., Калашиников С.Г.* Физика полупроводников. – М.: Наука, 1990.
3. *Смит Р.* Полупроводники. – М.: Мир, 1982.
4. *Sze S.M.* Semiconductor devices: Physics and Technology. – J.Wiley & Sonc, 2002.
5. *Chin-Tang San.* Fundamentals of Solid State Electronics. – World Scientific, 1991.

## Розділ 12

# НЕСТІЙКОСТІ У НАПІВПРОВІДНИКАХ

При розгляді дії електричного поля на напівпровідник ми обмежувались випадком слабких полів, коли зв'язок між струмом і напругою описувався законом Ома. Головною умовою виконання закону Ома була дія несильного електричного поля, за якої під час між двома зіткненнями електрон встигав би втратити енергію, набуту від зовнішнього поля. Дія сильного поля відбувається навпаки: на довжині вільного пробігу електрон набуває від поля енергії, яку не встигає передавати ґратці за рахунок зіткнень. У таких умовах вольт-амперні характеристики стають істотно нелінійними. Крім того, енергія електронного газу збільшується (електронна температура підвищується і відрізняється від температури ґратки): відбувається ефект розігріву електронів. Наявність електронів із високими значеннями енергії спричиняє виникнення суттєво нерівноважних станів, які можуть бути нестійкими. У цьому розділі ми розглянемо характерні риси поведінки напівпровідників і напівпровідникових структур за умов виникнення в них нестійкостей.

### ■ 12.1. Електропровідність напівпровідників у сильних електричних полях

При розгляді кінетичного рівняння ми вважали, що час релаксації не залежить від електричного поля. У цьому випадку для невідродженого напівпровідника зі сферичними ізоенергетичними поверхнями густина струму підлягає закону Ома, тобто

$$\mathbf{J} = \sigma \mathbf{E}, \quad (12.1)$$

де  $\sigma$  – провідність у слабкому полі. Для аналізу можливого впливу зовнішнього електричного поля на електропровідність напівпровідника згадаємо, що

$$\sigma = \frac{e^2 n \langle \tau \rangle}{m^*}. \quad (12.2)$$

Із (12.2) випливає, що:

**Порушення закону Ома можуть викликати дві причини: зі зміною поля змінюється час релаксації  $\langle \tau \rangle$  або концентрація носіїв.**

Поміркуємо про зміну часу релаксації у сильних полях. За ізотропного розсіювання час релаксації практично дорівнює часу вільного пробігу, оскільки швидкість направленої руху практично повністю зникає за

одне зіткнення. Але швидкість направленого руху залежить тільки від поля. Енергія, що передається електроном за одне зіткнення, складає малу частину повної енергії кристала

$$\langle \Delta E \rangle \approx \frac{m^*}{M} E. \quad (12.3)$$

Дійсно, закон збереження імпульсу при зіткненні електрона із ґраткою дає  $m^* v = M v_{ph}$ . Звідси швидкість руху іона після зіткнення  $v_{ph} = (m^*/M)v$ . Тоді енергія, передана ґратці за одноразового зіткнення електрона з іоном,  $\Delta E = M v_{ph}^2 / 2 = (m^*/M) m^* v^2 / 2 = (m^*/M) E$ .

Нехай відносний дефіцит енергії при зіткненні носія із ґраткою

$$\beta = \frac{\Delta E}{kT}, \quad (12.4)$$

тобто при зіткненні електрона із фононом енергія електрона змінюється на величину

$$\Delta E = \beta kT. \quad (12.5)$$

При розгляді процесів електропровідності у слабких полях ми вважали, що в процесі розсіювання електронів за їхнього руху в кристалі під дією електричного поля енергія електрона між двома актами зіткнень несильно змінюється (строго кажучи, не змінюється). Вважалось, що окремий акт взаємодії електрона з іоном при зіткненні є пружним. Процес розсіювання електронів за їхнього руху в кристалі під дією електричного поля буде пружним, якщо енергія, придбана електроном в електричному полі на довжині вільного пробігу, буде істотно меншою від енергії, яку втратив електрон у середньому за один акт зіткнення

$$eEl \ll \beta kT. \quad (12.6)$$

Це означає, що закон Ома є справедливим за умови, що електричне поле задовольняє нерівності

$$E \ll \beta \frac{kT}{el}. \quad (12.7)$$

У таких слабких полях розподіл швидкостей носіїв заряду залишається майже максвеллівським із температурою, що дорівнює температурі ґратки. За дії слабого поля на хаотичний рух носіїв заряду накладається зумовлений полем їхній повільний дрейф, швидкість якого в цих умовах є значно меншою за їхню теплову швидкість

$$v_d \ll v_T. \quad (12.8)$$

Вище зазначалося, що в сильних електричних полях закон Ома порушується за рахунок того, що:

- ♦ швидкість дрейфу носіїв стає порівнюваною з їхньою тепловою швидкістю, яка відбиватиметься на процесах розсіювання та викличе



зміни часу релаксації, а це в свою чергу – зміни рухливості. Із цим механізмом дії сильних електричних полів пов'язані *ефекти розігріву електронно-діркового газу та ефект Ганна*;

♦ сильні електричні поля спроможні викликати зміни концентрації носіїв, що є причиною *ефектів ударної іонізації, Зінера та електро-статичної іонізації*.

**У сильних електричних полях, коли нерівність (12.7) не виконується, носії заряду при зіткненні із фононами не встигають віддавати кристалічній ґратці енергію, набуту ними в полі за час між зіткненнями. У результаті середня енергія носіїв збільшується та електронно-дірковий газ розігрівається.**

З іншого боку, зростання енергії носіїв у першу чергу означає зростання їх середньої швидкості, що зменшує час вільного пробігу, тобто збільшується ймовірність розсіювання носіїв на теплових коливаннях ґратки. Завдяки цьому зростання енергії носіїв зі збільшенням напруженості поля буде уповільнюватись, доки електрон-фононна взаємодія не підсилиться до такого ступеня, що в кристалі встановиться стаціонарний стан балансу енергії, яку отримують носії у полі, та енергії, якої вони позбавляються внаслідок зіткнень із ґраткою. Енергія, яку за одиницю часу отримує носій від поля, дорівнює  $eEv_d$ . Вона у рівновазі дорівнюватиме енергії, яку електрон віддає ґратці за одиницю часу через зіткнення із фононами  $\beta(kT)/\tau$ , тобто

$$eEv_d = \beta \frac{kT}{\tau}. \quad (12.9)$$

За розсіювання на теплових коливаннях ґратки довжина вільного пробігу  $l$  не залежить від енергії. Отже часом вільного пробігу є  $\tau = l/v_T$ . У такому разі із (12.9) випливає, що дрейфовою швидкістю носіїв є

$$v_d = \beta \frac{kTv_T}{eEl}. \quad (12.10)$$

З іншого боку, дрейфова швидкість пов'язана із полем співвідношенням

$$v_d = \frac{e\tau}{m^*} E = \frac{el}{m^* v_T} E. \quad (12.11)$$

Ці рівняння дозволяють записати

$$(eEl)^2 = 2kT\beta m^* v_T^2 / 2. \quad (12.12)$$

Оскільки кінетична енергія  $m^* v_T^2 / 2$  у середньому дорівнює  $kT$ , то із (12.12) знаходимо, що критична напруженість поля, за якої починається розігрів електронного газу, дорівнює

$$E_{cr} = \frac{kT}{el} \sqrt{2\beta}. \quad (12.13)$$

Після підстановки напруженості критичного поля до (12.10) отримаємо, що розігрів електронного газу відбуватиметься тоді, коли швидкість, якої носії заряду набувають у зовнішньому полі,

$$v_d = \sqrt{\beta/2} \cdot v_T, \quad (12.14)$$

тобто стає порівнюваною із тепловою швидкістю носіїв. Із (12.12) можна визначити енергію носіїв заряду через теплову швидкість

$$\left( \frac{m^* v_T^2}{2} \right)^2 = \frac{(eEl)^2}{2\beta}, \quad (12.15)$$

звідки можна знайти

$$v_T = \sqrt{\frac{2eEl}{m^*(2\beta)^{1/2}}}. \quad (12.16)$$

Підставляючи це значення до (12.14), маємо

$$v_d = A\sqrt{E}, \quad (12.17)$$

де константа  $A = \sqrt{(el/m^*)\sqrt{\beta/2}}$  залежить як від ефективної маси, так і від параметрів процесу розсіювання: довжини вільного пробігу та енергії, що передається під час одиничного акту зіткнення. Таким чином, отримуємо, що в сильних електричних полях, коли

$$E > E_{cr} = \frac{kT}{el} \sqrt{2\beta}, \quad (12.18)$$

дрейфова швидкість носіїв збільшується пропорційно кореню квадратному з електричного поля

$$v_d \sim \sqrt{E}. \quad (12.19)$$

Відповідно й час релаксації зменшуватиметься (оскільки, згідно із (12.16),  $v_T \sim \sqrt{E}$ )

$$\tau = \frac{1}{v_d + v_T} \sim 1/\sqrt{E}. \quad (12.20)$$

Разом із ним зменшуватимуться рухливість і провідність

$$\mu_n = \frac{e\tau}{m^*} = \mu_0/\sqrt{E}, \quad \sigma = en\mu_n = \sigma_0/\sqrt{E}. \quad (12.21)$$

Обчислення залежності рухливості від поля за розсіювання носіїв на коливаннях ґратки виконав Шоклі та в результаті отримав формулу

$$\mu(E) = \mu_0 \frac{\sqrt{2}}{\sqrt{1 + \sqrt{1 + (3\pi/8)(\mu_0 E/c_s)^2}}}, \quad (12.21)$$

де  $c_s$  – швидкість звука,  $\mu_0$  – рухливість у слабкому полі. Для кремнію  $c_s = 7 \cdot 10^5$  см/с. У полях порядку  $5 \cdot 10^3$  В/см електрони набувають швидкості  $\sim c_s$ . За зростання поля до величин  $E \gg c_s/\mu_0$  (12.21) дає

$$\mu \cong \mu_0 \frac{4}{\sqrt{3\pi}} \left( \frac{c_s}{\mu_0 E} \right)^{1/2}.$$

Таким чином, за напруженості зовнішнього поля, більшої від критичної, закон Ома не виконується, а зв'язок між густиною струму та полем має вигляд

$$J = \sigma_0 \sqrt{E}. \quad (12.22)$$

Наприклад, у германії відхилення від закону Ома спостерігається за кімнатних температур уже в полях із напруженістю  $\sim 700$  В/см.

За подальшого зростання напруженості поля енергія носіїв збільшується настільки, що стає можливою їхня взаємодія з оптичними коливаннями ґратки. Оптичні фонони характеризуються більшою енергією за акустичні, і процеси розсіювання на оптичних фононах є для електронів більш енергоємними. Іншими словами, при зіткненні з оптичним фононом енергія носія може змінюватись на величину  $\hbar\omega_0$  ( $\omega_0$  – частота оптичного фонона), і таким чином швидкість втрат енергії збільшується. У результаті через деякий час встановлюється новий стаціонарний стан, за якого енергія  $eEv_d$ , набута носієм у полі за одиницю часу, дорівнює енергії  $\Delta W/\tau$ , яку він втрачає за той самий час при розсіюванні на оптичному фононі

$$eEv_d = \frac{\Delta W}{\tau}. \quad (12.23)$$

З іншого боку, енергією, яку втрачає носій за одиничний акт зіткнення з оптичним фононом, є (див. додаток E)

$$\Delta W = \hbar\omega_0 \text{th}(\hbar\omega_0/2kT), \quad (12.24)$$

звідки 
$$eEv_d = \frac{v_0}{1} \hbar\omega_0 \text{th}(\hbar\omega_0/2kT). \quad (12.25)$$

Оскільки із співвідношень  $v_d = \mu E = (e\tau/m^*)E = (el/m^*v_0)E$  випливає, що  $E = (m^*v_0v_d/el)$ , то після підстановки цього значення поля до (12.25) отримаємо для дрейфової швидкості носіїв заряду

$$v_d = \sqrt{\frac{\hbar\omega_0}{m^*} \text{th} \frac{\hbar\omega_0}{2kT}}. \quad (12.26)$$

Оскільки величина  $\text{th}(\hbar\omega_0/2kT)$  при зростанні  $\omega_0$  швидко наближається до одиниці, то можна визначити граничне значення дрейфової швидкості як

$$v_c = \sqrt{\hbar\omega_0/m^*}, \quad (12.27)$$

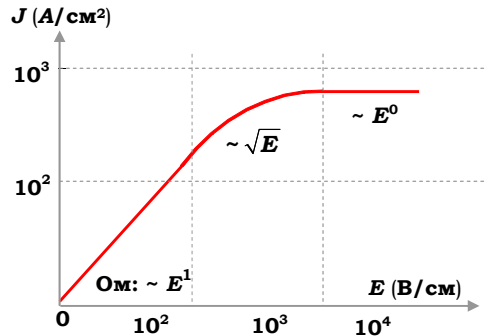
тобто густина струму наблизатиметься до насичення, значення якого дорівнює

$$J_c = env_c \quad (12.28)$$

і не залежить від напруженості поля. Інакше кажучи, при досягненні зовнішнім полем величини

$$E_c = \frac{v_0}{el} \left( m^* \hbar\omega_0 \operatorname{th} \frac{\hbar\omega_0}{2kT} \right)^{1/2} \quad (12.29)$$

настає насичення струму, отже і дрейфової швидкості. У ковалентних кристалах ці ефекти відбуваються в полях  $\sim 10^4$  В/см. При цьому швидкість дрейфу носіїв досягає значень  $\sim 10^7$  см/с. Зазначимо, що в іонних напівпровідниках за умов дії дуже сильних електричних полів ефекту насиченості струму не відбувається; навпаки, при зростанні поля зростає дрейфова швидкість носіїв і відбувається так званий ефект "втечі носіїв". Вольт-амперну характеристику в об'ємному зразку *n*-Ge, що вимірювалась за температури  $T = 77$  К у широкому інтервалі полів, подано на рис. 12.1.



**Рис. 12.1. Залежність електропровідності напівпровідника від величини електричного поля**

## ■ 12.2. Електронна температура

Розглянемо як приклад електронний напівпровідник. Нехай  $\tau_{ee}$  – середній час між зіткненнями електронів. Оскільки їх ефективні маси однакові, то при зіткненнях електрони обмінюються як енергією, так й імпульсом. Іншими словами,  $\tau_{ee}$  характеризує швидкість таких процесів обміну. Введемо характерні часи обміну енергією –  $\tau_E$  та обміну імпульсу –  $\tau_g$  при зіткненнях електронів із ґраткою. Тоді можна записати три граничних випадки співвідношень між цими характерними часами:

$$\tau_{ee} \ll \tau_g \ll \tau_E, \quad (12.30)$$

$$\tau_g \leq \tau_{ee} \ll \tau_E, \quad (12.31)$$

$$\tau_g \leq \tau_E \ll \tau_{ee}, \quad (12.32)$$

У першому випадку електрони встигають багаторазово обміняти енергією та імпульсом, доки у цих процесах більш-менш помітну роль почнуть відігравати ґратка та дефекти. У другому випадку електрони встигають багаторазово обміняти енергією, при цьому обмін імпульсом відбувається за активної участі ґратки та дефектів. У третьому випадку розсіювання носіїв один на одному є найповільнішим процесом і ним можна нехтувати – усі процеси релаксації відбуваються при розсіюванні носіїв на ґратці та дефектах.

У випадку (12.30) електронний газ можна розглядати як окрему термодинамічну підсистему, що слабко взаємодіє із ґраткою. За рахунок зіткнень між електронами в електронній підсистемі швидко встановлюється рівновага за енергіями та імпульсами – термодинамічна рівновага. Це означає, що електронну підсистему можна описати деякою температурою  $T_e$  (що називається електронною температурою), яка взагалі відрізняється від температури ґратки  $T$ . У цьому випадку вся електронна підсистема може характеризуватися швидкістю  $v_d$ , з якою електрони як ціле рухаються щодо ґратки. Це є повним аналогом руху рідини в ємності. Тому таке наближення часто називається гідродинамічним. Електронну температуру можна однозначно визначити із рівняння балансу енергії. Її можна ввести й у випадку (12.31). Це є так званим квазігідродинамічним наближенням.

Варто зазначити, що у власних або слабко легованих напівпровідниках, де концентрація вільних носіїв є невеликою, але спостерігається ефект розігріву електронного газу, в основному виконується умова (12.31). Стохастизація руху електронів відбувається шляхом його ізотропізації – відбувається релаксація імпульсу за умови збереження електронним газом високої середньої енергії.

Розглянемо схему визначення електронної температури у випадку виконання умов (12.30) і (12.31). Функція розподілу електронів має залежати від температури  $T_e$ , швидкості  $v_d$  та обертати в нуль інтеграл зіткнень. При  $v_d = 0$  це була б функція Фермі із заміною  $T$  на  $T_e$ . За переходу до рухомої системи координат, де електронний газ є нерухомим, знов маємо отримати функцію розподілу Фермі із заміною  $T$  на  $T_e$ . Тоді за наступного переходу до системи координат, пов'язаної із кристалом, необхідно зробити заміну  $\mathbf{v} \rightarrow \mathbf{v} - \mathbf{v}_d$  і записати функцію розподілу у вигляді

$$f(\mathbf{q}) = \left[ e^{\frac{E(\mathbf{v}-\mathbf{v}_d)-E_F}{kT_e}} + 1 \right]^{-1}. \quad (12.33)$$

Для невиродженого газу електронів

$$f(\mathbf{q}) = e^{-\frac{E(\mathbf{v}-\mathbf{v}_d)-E_F}{kT_e}}. \quad (12.34)$$

Якщо дрейфова швидкість руху електронів є малою порівняно із характерною тепловою швидкістю носіїв

$$v_{T_e} = \sqrt{\frac{kT_e}{m^*}}, \quad (12.35)$$

то функцію розподілу (12.34) можна розкласти у ряд та отримати

$$f(\mathbf{q}) = f_s + f_a = e^{-\frac{E(\mathbf{v}) - E_F}{kT}} + \frac{\mathbf{v}_d \cdot \nabla_{\mathbf{v}} E}{kT_e} e^{-\frac{E(\mathbf{v}) - E_F}{kT_e}}. \quad (12.36)$$

Перший доданок тут є симетричною функцією щодо зміни знаку імпульсу, тобто  $f_s(-\mathbf{q}) = f_s(\mathbf{q})$ , а другий – антисиметричною:  $f_a(-\mathbf{q}) = -f_a(\mathbf{q})$ . В умовах квазірівноваги  $f_s$  перетворюється на рівноважну функцію розподілу, а  $f_a$  – на малу поправку  $f_1$ . Якщо функцію розподілу відомо, то можна задати модель, в якій розраховуються часи релаксації  $\tau_E$  та  $\tau_q$ , а далі – використати умови балансу енергії та отримати рівняння щодо  $T_e$ . У випадку розігріву носіїв характерні часи розсіювання імпульсу  $\tau_q$  та енергії  $\tau_E$  будуть функціями електричного поля та електронної температури  $T_e$ . Тоді провідність напівпровідника також буде функцією електронної температури. У стаціонарному випадку питома потужність, яку електронний газ отримує від поля, з одного боку дорівнює джоулевій потужності  $P = \mathbf{j} \cdot \mathbf{E}$ , а з іншого – питомій потужності, яку електронний газ передає ґратці  $P(T_e)$

$$\sigma(T_e)E^2 = P(T_e). \quad (12.37)$$

Оскільки від електронної температури залежить час релаксації, то можна визначити залежність провідності від електронної температури як

$$\sigma(T_e) = \frac{ne^2}{m^*} \tau_q(T_e) = en\mu. \quad (12.38)$$

Енергія, яку в одиницю часу втрачає електронна підсистема (потужність, яку електрони передають ґратці),

$$P(T_e) = nc_e \frac{T_e - T}{\tau_E(T_e)}, \quad (12.39)$$

де  $c_e$  – теплоємність електронного газу. Тоді, використовуючи умови балансу енергії (12.37), отримаємо рівняння для визначення електронної температури

$$\frac{e^2}{m^*} \tau_{\mathbf{q}}(T_e) \cdot E^2 = c_e \frac{T_e - T}{\tau_E(T_e)}. \quad (12.40)$$

Із цього рівняння, застосовуючи ті чи інші моделі для визначення часів релаксації  $\tau_E(T_e)$  і  $\tau_{\mathbf{q}}(T_e)$ , можна знайти електронну температуру як функцію прикладеного зовнішнього електричного поля  $\mathbf{E}$ .

### ■ 12.3. Диференціальна провідність. N- та S-подібні вольт-амперні характеристики

Загалом існують три різних типи нелінійної поведінки  $J(E)$ . Нехай густина струму та електричне поле, що діє на напівпровідник, пов'язані матеріальним співвідношенням

$$J = \sigma_d E. \quad (12.41)$$

Тоді диференціальну провідність  $\sigma_d$  можна визначити як похідну

$$\sigma_d = \frac{dJ}{dE}. \quad (12.42)$$

Лінійна залежність (12.41) не означає, що провідність середовища не залежить від поля. Вона свідчить, що диференціальна провідність, яка визначає зв'язок між малими приростами повних значень струму  $dJ$  і поля  $dE$ , є константою, незалежною від амплітуди приросту. Узагалі у будь-якому кристалічному матеріалі диференціальна провідність є тензором і замість (12.41) маємо записати

$$\begin{pmatrix} dJ_x \\ dJ_y \\ dJ_z \end{pmatrix} = \begin{pmatrix} \sigma_{xx}^{(d)} & \sigma_{xy}^{(d)} & \sigma_{xz}^{(d)} \\ \sigma_{yx}^{(d)} & \sigma_{yy}^{(d)} & \sigma_{yz}^{(d)} \\ \sigma_{zx}^{(d)} & \sigma_{zy}^{(d)} & \sigma_{zz}^{(d)} \end{pmatrix} \cdot \begin{pmatrix} dE_x \\ dE_y \\ dE_z \end{pmatrix} \quad (12.43)$$

так, що

$$\sigma_{ij}^{(d)} = \frac{dJ_i}{dE_j}. \quad (12.44)$$

Але у більшості випадків доводиться мати справу з кристалами кубічної симетрії, де тензор (12.44) є діагональним. З іншого боку, часто апелюють до усереднених значень провідності, і розгляд скалярного зв'язку може бути достатнім.

Перший тип нелінійної поведінки характеризується позитивним значенням диференціальної провідності

$$\sigma_d = \frac{dJ}{dE} > 0. \quad (12.45)$$

При цьому, якщо зростання струму відбувається швидше за лінійне, то йдеться про *суперлінійну характеристику*. Якщо струм зростає повільніше за лінійний закон, то характеристика називається *сублінійною*. Іноді диференційну провідність записують у вигляді

$$\sigma_d = \sigma + E \frac{d\sigma}{dE}, \quad (12.46)$$

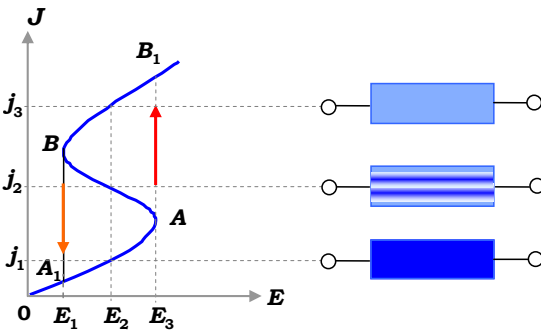
де  $\sigma$  – провідність,  $E$  – електричне поле. Тоді суперлінійній характеристиці відповідає позитивне значення похідної  $d\sigma/dE$ , а сублінійна характеристика визначається умовою  $d\sigma/dE < 0$ .

Другий тип нелінійної поведінки вольт-амперних характеристик визначається негативним значенням диференціальної провідності. Якщо в задачі задано струм та наявна ділянка характеристики, де

$$\sigma_d = \frac{dJ}{dE} < 0, \tag{12.47}$$

то йдеться про  $S$ - або  $N$ -подібну характеристику.

Почнемо із характеристики  $S$ -типу. Якщо ділянка з від'ємною диференціальною провідністю спостерігається на характеристиці, де задано напругу (або поле), то маємо  $S$ -подібну характеристику.



**Рис. 12.2. S-подібна вольт-амперна характеристика.**

У точці  $(E_2, j_2)$  диференціальна провідність від'ємна. Тут відбувається шнурування струму. Зразок розбивається на домени із великою  $j_3$  та малою  $j_1$  густиною струму

За заданого електричного поля  $E_1$  у системі, що описується такою характеристикою, могли б протікати струми  $j_1, j_2, j_3$  (рис. 12.2). Стани зі струмами  $j_1, j_3$  характеризуються позитивним значенням  $\sigma_d$  та є стійкими. Дійсно, якщо в системі виникає флуктуація заряду, то при  $\sigma_d > 0$  зі зростанням часу вона буде розсмоктуватись, згідно із законом  $\sim \exp(-t/\tau_M)$  із максвеллівським часом релаксації

$$\tau_M = \frac{\epsilon}{4\pi\sigma_d}, \tag{12.48}$$

де  $\epsilon$  – діелектрична стала кристала. Дійсно, нехай завдяки флуктуації у момент часу  $t = 0$  у напівпровіднику виникла область просторового заряду із густиною  $\rho(\mathbf{r}, 0)$ . Цей заряд породжує електричне поле, що пов'язано із густиною рівнянням Пуассона

$$\text{div}\mathbf{E} = 4\pi \frac{\rho}{\epsilon}. \tag{12.49}$$

Електричне поле генерує струм  $\mathbf{j} = \sigma_d \cdot \mathbf{E}$ , який задовольняє рівнянню неперервності

$$\frac{\partial \rho}{\partial t} = -\text{div}\mathbf{j} = -\text{div}(\sigma_d \mathbf{E}) = -\sigma_d \cdot \text{div}\mathbf{E} = -4\pi \frac{\sigma_d}{\epsilon} \rho. \tag{12.50}$$

Звідси знаходимо рівняння, що описує релаксацію флуктуації густини заряду до однорідної

$$\frac{\partial \rho}{\partial t} = -\frac{1}{\tau_M} \rho \tag{12.51}$$



і має розв'язок, який свідчить: флуктуація густини заряду релаксує до рівноважної за експоненціальним законом із часом релаксації  $\tau_M$

$$\rho(t) = \rho(0)e^{-t/\tau_M}. \quad (12.52)$$

Із (12.48) та (12.52) видно також, що при  $\sigma_d < 0$  флуктуація із часом не зменшується, а зростає за експоненціальним законом, тобто система, виведена зі стану рівноваги, не повернеться до цього стану – вона буде нестійкою. Із цієї причини стан, що характеризується струмом  $j_2$ , де диференціальна провідність від'ємна, буде нестійким. Оскільки полю  $E_1$  відповідають два стійких стани  $j_1$ ,  $j_3$  (рівноправні у всіх відношеннях), то система, намагаючись залишатись у стійкому стані, розіб'ється на області із великою (область великої концентрації носіїв) та з малою густиною струму (область низької концентрації носіїв). Зразок розшарується вздовж струму. Відбудеться так зване шнурування струму. Таким чином, за зростання поля від нуля до точки  $A_1$  на вольт-амперній характеристиці зразок буде однорідний, далі у зразку можуть утворитись шнури струму, і тільки за збільшення поля до величини  $B_1$  зразок знову стане однорідний. Взагалі кажучи, область однорідного струму може спостерігатись і за зростання поля до точки  $A$ , і за зменшення його до точки  $B$ .

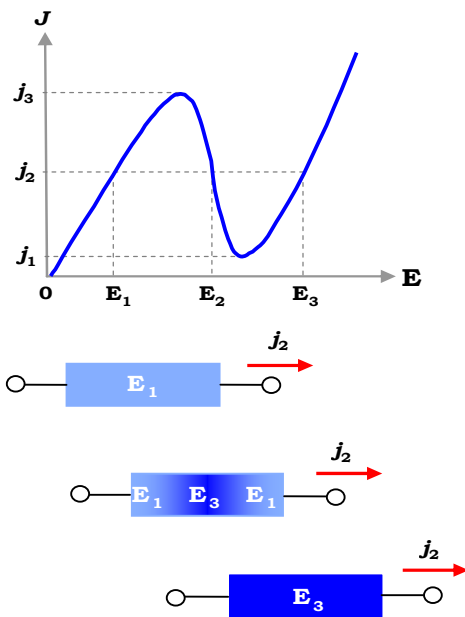


Рис. 12.3.  $N$ -подібна вольт-амперна характеристика. Утворення домену сильного поля

У напівпровіднику із  $N$ -подібною характеристикою заданому значенню струму  $j_2$  відповідають три значення поля –  $E_1$ ,  $E_2$ ,  $E_3$  (рис. 12.3). Стан, що характеризується полем  $E_2$ , має від'ємну диференціальну провідність, тому він є нестійким. У результаті об'єм кристала розіб'ється на домену. Зовні домену поле буде малим –  $E_1$ , а всередині домену – великим  $E_3$ . Якщо на зразок подіяти імпульсом електричної напруги, то домен сильного поля виникне біля катоду та рухатиметься у бік аноду. При утворенні домену струм в електричному ланцюгу (отже й у зразку) зменшиться й стане меншим за  $j_1$ . Таким чином у зразку може існувати лише один домен сильного поля. При зникненні домену струм знову зростає та утворюється домен сильного поля. У кристалі виникнуть коливання – струмова

нестійкість. Швидкість руху домену приблизно дорівнює дрейфовій швидкості електронів, тому частоту коливань можна оцінити як

$$v = \frac{v_{dr}}{l} = \frac{1}{\tau_{dr}}, \quad (12.53)$$

де  $l$  – довжина зразка,  $\tau_{dr}$  – час руху домену через зразок. Таким чином, для отримання коливання із великою частотою необхідно використовувати зразок із невеликою довжиною. Але існує деяка мінімальна довжина зразка, за якої домен до винесення його дрейфом за межі зразка ще не встигає сформуватися. Генерація коливань поля (або струму) може відбуватись і в кристалах із  $S$ -подібною характеристикою.

Розглянемо систему із позитивною диференціальною провідністю, в якій також можлива генерація коливань. Нехай провідність кристала

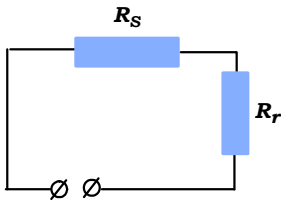
$$\sigma = e[\mu_n n(\mathbf{E}) + \mu_p p(\mathbf{E})]. \quad (12.54)$$

Оскільки в сильному полі рухливість і концентрація носіїв залежать від поля, виділимо два параметри, які протилежним чином залежать від поля – один зростає, а другий – зменшується за зростання поля. Нехай  $\sigma_n$  зменшується з полем, тобто

$$\frac{d\sigma_n}{dE} < 0, \quad (12.55)$$

а  $\sigma_p$  – зростає, тобто

$$\frac{d\sigma_p}{dE} > 0. \quad (12.56)$$



При цьому сумарна диференціальна провідність лишається позитивною. Включимо такий зразок до електричного ланцюга сталого струму (рис. 12.4). За виконання умов  $R_r \gg R_s$  поле всередині зразка визначається тільки його провідністю ( $j = \text{const}$ )

$$E = j/\sigma. \quad (12.57)$$

**Рис. 12.4.** Еквівалентна схема системи із позитивною диференціальною провідністю

Нехай у момент часу  $t_1 = 0$  виникла флуктуація в електронній підсистемі, яка викликала збільшення електронної компоненти провідності  $\sigma_n^{(1)} > \sigma_n$ . Це приведе до зменшення поля у зразку до величини

$E^{(1)} = j/(\sigma_p + \sigma_n^{(1)})$ , яке, у свою чергу, спричинить зменшення діркової компоненти провідності. Остання за час  $\tau_p$  зменшиться настільки, що в момент часу  $\tau_p$  поле у зразку матиме значення  $E^{(2)} = j/(\sigma_p^{(2)} + \sigma_n^{(1)}) > E > E^{(1)}$ .

Але збільшення поля зменшить електронну компоненту провідності та збільшить – діркову. Якщо сумарна провідність за деякий час знову стане більшою за  $R$ , то поле стане меншим за  $E$ . Таким чином, через

польовий зв'язок провідності у системі можливе виникнення коливань електричного поля на зразку, частота яких визначатиметься часом життя гарячих електронів  $\tau_n$  і дірок  $\tau_p$

$$\omega \approx \frac{1}{\sqrt{\tau_n \tau_p}}. \quad (12.58)$$

Далі визначимо формальні умови, за яких можливе виникнення нелінійних вольт-амперних характеристик  $N$ - та  $S$ -типів.

## ■ 12.4. Критерії виникнення нестійкості

Знайдемо деякі загальні співвідношення, що стануть у нагоді при визначенні можливості виникнення нестійкості у напівпровіднику або напівпровідниковій структурі. Знайдемо диференціальну провідність. Для цього використаємо (12.42) і запишемо його у формі (12.46)

$$\sigma_d = \frac{dJ}{dE} = \frac{d\sigma E}{dE} = \sigma + E \frac{d\sigma}{dE} \quad (12.59)$$

або

$$\sigma_d = \sigma + E \frac{d\sigma}{dT_e} \frac{dT_e}{dE}. \quad (12.60)$$

Для визначення  $dT_e/dE$  продиференціюємо (12.37) за  $E$

$$2\sigma E + E^2 \frac{d\sigma}{dT_e} \frac{dT_e}{dE} = \frac{dP}{dT_e} \frac{dT_e}{dE}. \quad (12.61)$$

Звідси отримуємо

$$\frac{dT_e}{dE} = \frac{2\sigma E}{\frac{dP}{dT_e} - E^2 \frac{d\sigma}{dT_e}}. \quad (12.63)$$

Підставимо (12.63) до (12.60). Маємо

$$\sigma_d = \sigma \left[ 1 + \frac{2E^2 \frac{d\sigma}{dT_e}}{\frac{dP}{dT_e} - E^2 \frac{d\sigma}{dT_e}} \right].$$

Після заміни  $E^2$  на  $P/\sigma$  отримуємо

$$\sigma_d = \sigma \left[ \frac{\frac{dP}{dT_e} + \frac{P}{\sigma} \frac{d\sigma}{dT_e}}{\frac{dP}{dT_e} - \frac{P}{\sigma} \frac{d\sigma}{dT_e}} \right] = \sigma \frac{\frac{dP}{dT_e} + P \frac{d \ln \sigma}{dT_e}}{\frac{dP}{dT_e} - P \frac{d \ln \sigma}{dT_e}}. \quad (12.63)$$

Ця формула, якщо її переписати у вигляді

$$\sigma_d = \sigma \frac{\frac{d \ln \sigma P}{d T_e}}{\frac{d \ln (P / \sigma)}{d T_e}}, \quad (12.64)$$

є зручною для отримання умов виникнення  $N$ - чи  $S$ -подібних вольт-амперних характеристик. Дійсно, у випадку  $N$ -подібної характеристики диференціальна провідність, проходячи через нуль, змінює знак, тобто для точки зміни знаку  $\sigma_d$  маємо

$$\frac{d}{d T_e} \ln \sigma P = 0. \quad (12.65)$$

Це рівняння можна переписати у зручнішому вигляді, якщо використати (12.38) та (12.39)

$$\sigma P = n^2 e c_e (T_e - T) \frac{\mu}{\tau_E}. \quad (12.66)$$

Згідно із (12.65) та використовуючи (12.66), запишемо

$$\frac{d \ln \sigma P}{d T_e} = \frac{1}{\sigma P} \frac{d \sigma P}{d T_e} = 0 \rightarrow \frac{d \sigma P}{d T_e} = 0 \rightarrow \frac{d}{d T_e} \left( \frac{n^2 \mu}{\tau_E} (T_e - T) \right) = 0$$

або

$$\frac{n^2 \mu}{\tau_E} + (T_e - T) \frac{d}{d T_e} \left( \frac{n^2 \mu}{\tau_E} \right) = 0. \quad (12.67)$$

Це й є умовою виникнення в системі нестійкості, що характеризується  $N$ -подібною вольт-амперною характеристикою.

Аналогічно знайдемо умови виникнення нестійкості  $S$ -типу. У випадку  $S$ -подібної характеристики диференціальна провідність змінює знак, проходячи через нескінченність. У цьому випадку умовою зміни знака  $\sigma_d$  буде

$$\frac{d}{d T_e} \ln (P / \sigma) = 0. \quad (12.68)$$

Зауважимо, що

$$\ln (P / \sigma) = \ln (c_e / e) + \ln (T_e - T) \frac{1}{\mu \tau_e},$$

а

$$\frac{d}{d T_e} [\ln (T_e - T) - \ln \mu \tau_e] = \frac{1}{T_e - T} - \frac{1}{\mu T_e} \frac{d}{d T_e} (\mu \tau_e),$$

тоді із (12.68) маємо

$$1 - \frac{T_e - T}{\mu \tau_e} \frac{d}{d T_e} (\mu \tau_e) = 0. \quad (12.69)$$

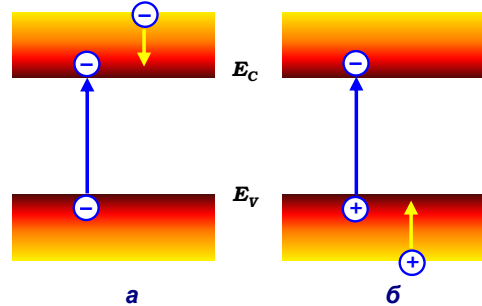
Таким чином, для виникнення  $N$ -подібної характеристики параметри напівпровідника мають задовольняти умові (12.67), а для виникнення  $S$ -подібної характеристики – умові (12.69). Для визначення можливості та типу нестійкості в конкретній системі необхідно в рамках тієї чи іншої моделі знайти рухливість, концентрацію носіїв заряду та час релаксації енергії як функції температури та використати (12.67) або (12.69).

Як приклад  $S$ -подібної нестійкості можна назвати низькотемпературний пробій, що виникає завдяки явищу ударної іонізації, а  $N$ -подібної характеристики – утворення домену сильного поля та виникнення ефекту Ганна. Ці ефекти детальніше розглянемо нижче.

## ■ 12.5. Ударна іонізація

Ефект ударної іонізації виникає в полях  $\sim 10^5$  В/см, коли носії набувають енергії, достатньої для іонізації атомів речовини, і може служити прикладом ефектів, пов'язаних зі зміною концентрації носіїв за дії на напівпровідник сильного електричного поля. У результаті виникають електронно-діркові пари, які у свою чергу прискорюються полем і генерують додаткові носії. Схематично генерацію носіїв у процесі ударної іонізації подано на рис. 12.5. Однак спостерігати такі

ефекти можна лише в  $p$ - $n$ -переходах (або на спеціально легованих гетерограницях), оскільки створення полів такої великої напруженості в однорідному напівпровіднику неможливе. У макроскопічно однорідному напівпровіднику цей ефект можна спостерігати за ударної іонізації домішкових центрів за низьких температур, коли домішки не іонізовані й потреби у сильних полях немає. Наприклад, ударну іонізацію домішок п'ятої групи в германії можна спостерігати при  $E \sim 5\text{--}10$  В/см, коли кожний одиничний акт іонізації є тричастинковим процесом. Ударна іонізація за низьких температур спостерігатиметься як різке збільшення густини струму під дією поля. На рис. 12.6 подано криву залежності густини струму у зразку електронного германію при  $T = 4,2$  К. При  $E \sim 5$  В/см відбувається різке збільшення густини струму, зумовлене іонізацією атомів домішки. З ефектом ударної іонізації пов'язаний



**Рис. 12.5.** Схема ефекту ударної іонізації.

Під дією сильного електричного поля електрон іонізує атом речовини із виникненням додаткової електрон-діркової пари (а). Ефект іонізації діркою (б)

ефект низькотемпературного пробую: при прикладенні до зразка електричного поля спостерігається немонотонне зростання струму. Електричне поле, набувши величини  $E_1$ , достатньої для ударної іонізації (рис. 12.7), стрибком зменшується до величини  $E_2$ , достатньої для підтримування пробую. Вертикальний підйом струму пов'язаний зі шнурованням.

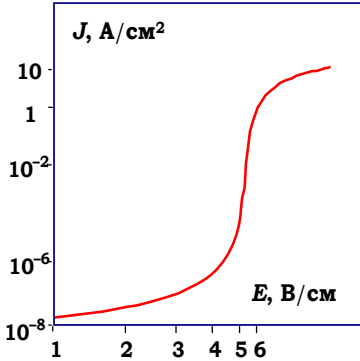


Рис. 12.6. Ударна іонізація в *n*-Ge при  $T = 4,2 \text{ K}$

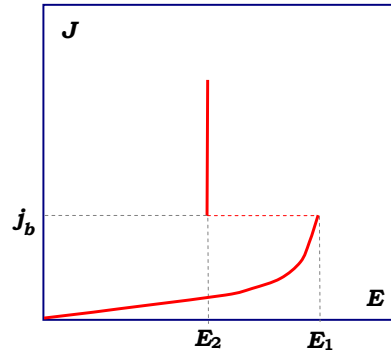


Рис. 12.7. Ефект низькотемпературного пробую

### ■ 12.6. Ефект Ганна

У напівпровідниках, де зона провідності має більше одного мінімуму, електрон із хвильовим вектором  $\mathbf{k}$  з однієї долини може за розсіювання потрапити до іншої та опинитися у стані  $\mathbf{k}'$ . Тобто може існувати так зване міждолинне розсіювання, що спричинить коливання електричного струму за дії на однорідний напівпровідник сильного сталого електричного поля. Таке явище називається ефектом Ганна. Воно вперше спостерігалось на зразках арсеніду галію. Для розуміння цього явища розглянемо зонну структуру GaAs у напрямку вісі  $[100]$  (рис. 12.8). Істотною рисою зонної структури тут є валентна зона із максимумом при  $k = 0$  і дві долини зони провідності: легких електронів (з ефективною масою  $m_1^* = 0,068m_0$ ) із мі-

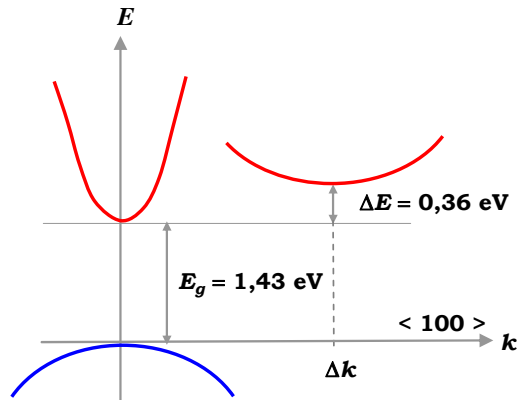


Рис. 12.8. Зонна структура GaAs у напрямку  $\langle 100 \rangle$

німумом при  $k = 0$ , що відділена від валентної зони енергетичною щілиною  $E_g = 1,43$  еВ і важких електронів (з ефективною масою  $m_2^* = 1,2m_0$ ) із мінімумом, що зсунутий на величину  $\Delta k$ . Мінімум лежить на величину  $\Delta E = 0,36$  еВ вище по енергії за мінімум долини легких електронів. Легкі електрони мають високі значення рухливості ( $\mu_1 = 4-8 \cdot 10^3$  см<sup>2</sup>/(В·с)), які майже на два порядки перевищують рухливість важких електронів ( $\mu_2 \approx 100$  см<sup>2</sup>/(В·с)). При цьому щільність станів верхньої долини приблизно в 700 разів вища за щільність станів нижньої долини. За низьких температур і невеликої напруженості електричного поля енергія електронів незначна. Її не вистачає для отримання імпульсу  $\sim \hbar \Delta k$  за розсіювання на фононах (для передачі частини енергії фонону) і збереження запасу енергії, якого б вистачило на перехід із валентної зони до верхньої долини зони провідності. Таким чином електрони накопичуватимуться на дні долини легких електронів (рис. 12.9 а). Густина електричного струму визначатиметься концентрацією електронів у нижній долині

$$n_1 = n_0(T), \quad n_2 = 0, \quad J = en_0 \mu_1 E. \quad (12.70)$$

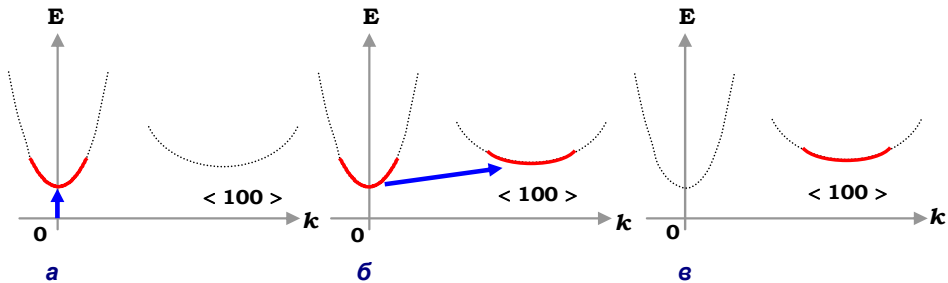


Рис. 12.9. Розподіл електронів у долинах GaAs за різних значень електричного поля:  
а –  $E < E_a$ ; б –  $E_a < E < E_b$ ; в –  $E > E_b$

Зі зростанням поля зростає й середня енергія електронів. Коли вона почне перевищувати величину енергетичного зазору між електронними долинами ( $E_a = \Delta E$ ), то частина електронів із нижньої долини почне переходити до верхньої долини, де електрони мають малу рухливість (рис. 12.9 б). Тоді в інтервалі полів  $[E_a, E_b]$  струм зменшуватиметься зі зростанням поля

$$n_0 = n_1 + n_2, \quad J = e \{ \mu_1 n_0 - (\mu_1 - \mu_2) n_2(E) \} E. \quad (12.71)$$

В області великих полів усі електрони перейдуть до верхньої долини зони провідності (рис. 12.9 в), і густина струму зі зростанням поля знов почне зростати

$$n_1 = 0, \quad n_2 = n_0, \quad J = en_0 \mu_2 E. \quad (12.72)$$

У результаті отримаємо  $N$ -подібну ВАХ напівпровідника (рис. 12.10), що свідчить про наявність нестійкості. Розглянемо фізичну причину такої нестійкості. В однорідному напівпровідниковому зразку довжиною  $L$  електричне поле приблизно однакове за всією довжиною зразка. Але в місці локальної неоднорідності, що характеризується підвищеним опором, напруженість електричного поля буде дещо вищою. Тобто критичне значення поля  $E_a$  виникне в області цієї неоднорідності. Тоді тут же з'являться важкі електрони. Рухливість носіїв зменшиться, а загальний опір підвищиться. Це приведе до збільшення напруженості поля, що спричинить інтенсивніші переходи електронів із долини легких до долини важких електронів. У результаті розподіл електричного поля стане сильно неоднорідний: виникне область сильного електричного поля – *електричний домен* (рис. 12.11), де рухатимуться важкі електрони із порівняно малою швидкістю. З лівого боку їх наздоганятимуть легкі електрони з області зовнішньої щодо домену, тобто утвориться область негативного об'ємного заряду. Праворуч від домену (що повільно рухається зліва направо) утвориться область, збіднена електронами (легкі електрони швидко рухаються, лишаючи позаду себе область домену).

Оскільки всередині домену напруженість поля швидко зростає, зростатиме й швидкість повільних електронів. Зростання спостерігатиметься до припинення зростання електричного поля. А зростання поля всередині домену припиниться, коли швидкість домену  $v_D$  зрівняється із дрейфовою швидкістю електронів зовні домену

$$v_D = v_{dr} \tag{12.73}$$

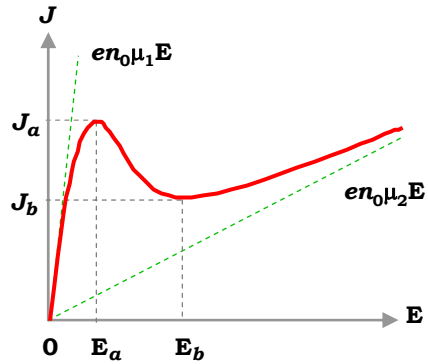


Рис. 12.10. Залежність густини струму від електричного поля у дводолінному напівпровіднику

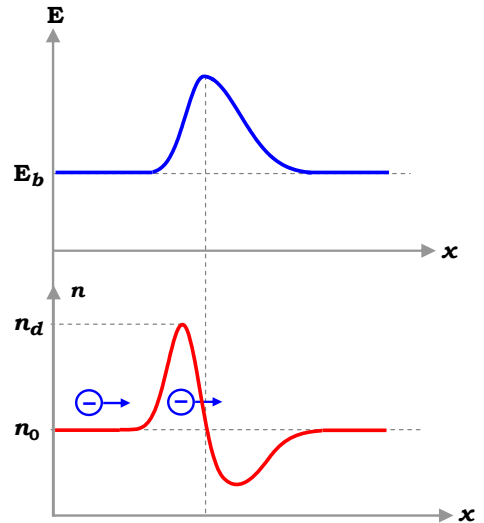


Рис. 12.11. Розподіл електричного поля (зверху) та електронної густини (знизу) вздовж зразка, в якому існує стабільний домен сильного поля, що рухається зі швидкістю  $v_D$



Стала швидкість руху домену буде меншою за швидкість електронів, яку вони мають у момент досягнення полем критичної позначки  $E_a$ , тому при підключенні до зразка напруги струм буде максимальний –  $I_{\max}$ . Почнеться утворення домену. При цьому процес визначається часом міждолинного переходу  $\tau_{w-w} \sim 10^{-12}$  с: струм швидко почне спадати до  $I_{\min}$ , що визначається швидкістю сталого руху домену

$$I_{\min} = en_0 v_D S, \quad (12.74)$$

де  $S$  – площа перерізу зразка. Мінімальне значення струму зберігатиметься, доки домен не пройде весь зразок і не зникне на правій його частині. У цей момент струм у зразку знов почне зростати, при досягненні  $I_{\max}$  у зразку утвориться новий електричний домен, і струм знизиться до  $I_{\min}$ . Таким чином отримуємо коливання струму, частота яких визначається довжиною зразка та швидкістю руху домену, що може досягати кількох гігагерц за довжини зразка порядку десятків мікрон. Треба зауважити, що у зразку утворюватиметься тільки один домен, незважаючи на можливість існування кількох областей неоднорідності поля.

Пояснення осциляцій Ганна переходом електронів між нееквівалентними долинами підтверджується експериментами, в яких різниця енергій між долинами зменшувалась під дією гідростатичного тиску. Оскільки добре відомо, що в разі прикладення гідростатичного тиску до зразку арсеніду галію основна долина (при  $\mathbf{k} = 0$ ) піднімається, а додаткова долина важких електронів опускається, і за тиску  $6 \cdot 10^9$  Па ці долини виявляються на одному рівні, що лежить на 0,49 еВ вище від положення основної долини за нульового тиску. Експериментально виявлено, що порогове поле для осциляцій Ганна зменшується зі збільшенням тиску, доки не досягає значення 1,4 кВ/см за тиску  $2,6 \cdot 10^9$  Па. За збільшення тиску осциляції не виникають. Оскільки виникнення домену сильного поля відбувається на одній із неоднорідностей зразка, то різне розташування неоднорідностей може спричиняти різні пролітні часи доменів, що означає виникнення коливань із різними періодами. Зрозуміло, що цей факт викличе "замивання" ефекту. Тому для експериментального спостереження ефекту Ганна варто використовувати чисті та однорідні зразки. На рис. 12.12. подано характерний вигляд осциляцій струму за

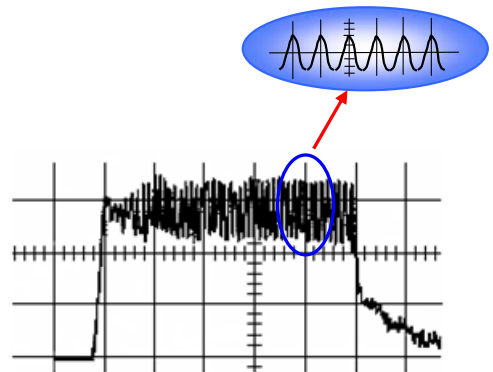
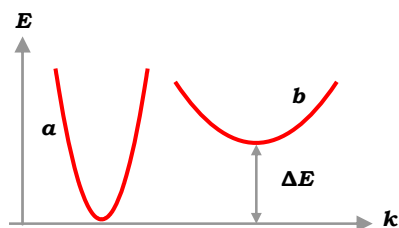


Рис. 12.12. Осцилограма осциляцій Ганна в GaAs

експериментального спостереження ефекту Ганна. Такі осциляції виникали при прикладенні до зразку арсеніду галію  $n$ -типу довжиною  $2,5 \cdot 10^{-3}$  см імпульсу напруги з амплітудою 16 В і тривалістю 10 нс. Частота осцилюючої компоненти дорівнювала 4,5 ГГц. Масштаб рисунку – 2 нс/см за горизонтальною й 0,23 А/см – за вертикальною осями. На врізанні показано розтягнуту частину осцилограми (виділено овалом). Як видно, із великою точністю осциляції можна вважати гармонічними із фіксованою частотою.

## ■ 12.7. Умови виникнення від'ємної диференціальної провідності

За уважного розгляду всіх обставин виникнення ефекту Ганна виникає запитання, чому він не виявляється у всіх напівпровідниках із зонами легких і важких носіїв, які розділені деяким енергетичним зазором. Які фактори електронних властивостей напівпровідника впливають на можливість виникнення від'ємного опору? Розглянемо напівпровідник, зона провідності якого складається із двох долин – легких  $a$ - та важких  $b$ -електронів (рис. 12.13) з ефективною щільністю станів  $N_a$  та  $N_b$ , відповідно, що розділені енергетичним проміжком  $\Delta E$ . Нехай ці долини характеризуються ефективними масами електронів  $m_a$  та  $m_b$ , рухливостями  $\mu_a$  та  $\mu_b$ . Якщо концентрація електронів у нижній долині дорівнює  $n_a$ , а у верхній –  $n_b$ , то



**Рис. 12.13. Особливості зонної структури, необхідні для виникнення від'ємної диференціальної провідності**

провідність такого напівпровідника визначатиметься формулою

$$\sigma = en_a \mu_a + en_b \mu_b. \quad (12.75)$$

Припустимо, що на зразок діє сильне електричне поле  $E$  так, що електрони із нижньої зони розігріваються та спроможні подолати бар'єр  $\Delta E$ . У результаті розігріву носіїв змінюватимуться їх концентрації та рухливості, що приведе до зміни провідності. Таким чином зміна провідності зі зміною електричного поля характеризуватиметься похідною

$$\frac{d\sigma}{dE} = e \left( \mu_a \frac{dn_a}{dE} + \mu_b \frac{dn_b}{dE} \right) + e \left( n_a \frac{d\mu_a}{dE} + n_b \frac{d\mu_b}{dE} \right). \quad (12.76)$$

Нехай

$$n(E) = n_a + n_b. \quad (12.77)$$

Припустимо, крім того, що рухливості носіїв в обох долинах мають ступеневу залежність від електричного поля

$$\mu_{a,b} = CE^p, \quad (12.78)$$

де  $C$  і  $p$  – константи. Тоді (12.76) можна представити у вигляді

$$\frac{d\sigma}{dE} = e(\mu_a - \mu_b) \frac{dn_a}{dE} + e(n_a \mu_a + n_b \mu_b) \frac{p}{E} = e(\mu_a - \mu_b) \frac{dn_a}{dE} + \sigma \frac{p}{E}. \quad (12.79)$$

Зауважимо, що оскільки густина струму  $j = \sigma \cdot E$ , то диференціальна провідність

$$\sigma_d = \frac{dj}{dE} = \sigma + E \frac{d\sigma}{dE}. \quad (12.80)$$

Умовою виникнення  $N$ -подібної ВАХ є від'ємність диференціальної провідності або необхідність виконання нерівності

$$\sigma + E \frac{d\sigma}{dE} < 0. \quad (12.81)$$

Підставляючи до цієї нерівності співвідношення (12.79), отримуємо

$$\frac{\mu_a - \mu_b}{\mu_a + f\mu_b} \cdot \left( -\frac{E}{n_a} \cdot \frac{dn_a}{dE} \right) - p > 1, \quad (12.82)$$

де  $f = n_b/n_a$  означає відносну частку електронів у верхній долині. З'ясуємо, за яких умов може виконуватись ця нерівність. Оскільки зі зростанням електричного поля все більше електронів переходять із долини  $a$  до долини  $b$ , але поле ще недостатньо велике для того, щоб викликати масовий перехід електронів із валентної зони до нижньої ( $a$ ) долини зони провідності, то похідна  $dn_a/dE < 0$ . Іншими словами, перший доданок у лівій частині нерівності (12.82) є позитивним. Величина  $p$  залежить від механізму розсіювання носіїв у напівпровіднику. Видно, що розсіювання на іонізованих домішках ( $p > 0$ ) не сприяє виникненню умови  $\sigma_d < 0$ . Коли рухливість електронів визначають процеси розсіювання на коливаннях ґратки ( $p < 0$ ), то вже за не дуже великих значень електричного поля рухливість носіїв формується за рахунок розсіювання на акустичних коливаннях ґратки, де  $p = -1/2$ , і в цьому випадку нерівність (12.82) вже може бути задоволена. Якщо ж рухливість електронів визначатиметься процесами розсіювання на оптичних коливаннях ґратки, коли  $p = -1$ , то нерівність (12.82) задовольняється автоматично (якщо, звісно, похідна  $dn_a/dE$  залишається від'ємною). Такі умови виконуються для напівпровідників типу  $A_{III}B_{V}$ , таких як GaAs, GaP та InP. Дійсно, оскільки в GaAs  $\Delta E = 0,36$  еВ при  $E_g = 1,45$  еВ й у полях  $\sim 7 \cdot 10^3$  В·см<sup>-1</sup> починається розігрів носіїв із переважним розсіюванням на оптичних фонах, то виникає від'єм-

на диференціальна провідність, і в напівпровіднику спостерігається ефект Ганна. У InSb  $\Delta E = 0,53$  еВ, ширина забороненої зони – 0,18 еВ, що унеможливає виконання умови (12.82). Дійсно, тут розігрів носіїв приводить не тільки (і не стільки) до переходу їх до верхньої долини, а, головним чином, – до генерації носіїв у нижній зоні через прямі переходи **зона–зона**, що призводить до того, що  $dn_a/dE > 0$ . Саме тому в антимоніді індію не спостерігається ефекту виникнення від'ємної диференціальної провідності.

## ■ 12.8. Задачі

1. Оцініть середній час зіткнень між носіями  $\tau_{ee}$ , вважаючи закон дисперсії носіїв квадратичним.

**Розв'язок.** Розглянемо рух двох заряджених частинок масою  $m_1^*$  та  $m_2^*$  у середовищі, що характеризується діелектричною функцією  $\epsilon$ . Зважаючи на наявність екранування, потенціальною енергією їхньої взаємодії є

$$U(r) = \frac{e^2}{\epsilon r} e^{-r/r_0}, \quad (12.83)$$

де  $r_0$  – радіус екранування. Задача про рух двох частинок у полі потенціалу  $U(r)$  розпадається на дві окремі: про рух центру мас системи та рух частинки зі зведеною  $m_r^* = m_1^* m_2^* / (m_1^* + m_2^*)$  масою в полі потенціалу  $U(r)$ . Оскільки потенціальна енергія (12.83) має такий самий вигляд, що і за розсіювання носіїв на іонізованих домішках, можна розглянути задачу про рух частинки із масою  $m_r^*$  у полі потенціалу (12.83), тобто скористатися результатом обчислення часу релаксації за розсіювання носіїв на іонізованих домішках, що наведено у додатку D. Згідно із (D23), час релаксації за розсіювання частинки масою  $m_r^*$  у полі потенціалу  $U(r)$  визначається формулою

$$\tau_{ee} = \frac{\epsilon^2 E^{3/2} (2m_r^*)^{1/2}}{\pi \cdot e^4 N_I \ln \left( \frac{8m_r^* E r_0^2}{\hbar^2} \right)}. \quad (12.84)$$

Замінімо енергію на її середнє значення  $(3/2kT)$ , а концентрацію розсіювачів  $N_I$  – на концентрацію носіїв заряду  $n$ , отримаємо

$$\tau_{ee} = \frac{\epsilon^2 (kT)^{3/2} (2m_r^*)^{1/2}}{\pi \cdot e^4 n \ln \left( \frac{8m_r^* kT \cdot r_0^2}{\hbar^2} \right)}. \quad (12.85)$$

2. Оцініть концентрації носіїв, за яких можливе застосування гідродинамічного та квазігідродинамічного наближень.

**Розв'язок.** Як і очікувалось, у невідродженій електронній системі час релаксації обернено-пропорційний до концентрації носіїв (див. (12.85)). Для визначення критичних концентрації оцінимо часи релаксації  $\tau_E$  та  $\tau_q$ . Нехай  $v_{T_e}$  – середня швидкість носіїв заряду з електронною температурою  $T_e$ . Тоді довжина вільного пробігу між зіткненнями, коли імпульс майже не змінюється

$$l_q = v_{T_e} \tau_q. \quad (12.86)$$

З іншого боку, згідно із (12.40), використаємо (12.86) та отримуємо

$$\frac{e^2}{m^*} \cdot \frac{l_q}{v_{T_e}} E^2 = c_e \frac{T_e - T}{\tau_E(T_e)}. \quad (12.87)$$

Беручи до уваги, що  $m^* v_{T_e} \approx 2 \cdot (3kT_e/2)^{1/2} (m^*)^{1/2}$  та  $T_e - T \approx T_e$ , отримуємо

$$\tau_E \approx \sqrt{6} \frac{c_e (kT_e)^{3/2} (m^*)^{1/2}}{k e^2 E^2 l_q}. \quad (12.88)$$

Визначимо критичну концентрацію  $n_E$ , за якої  $\tau_{ee} \approx \tau_E$ ,

$$n_E \approx \frac{\varepsilon^2 E^2 l_q}{e^2 \ln \left( \frac{8m_r^* T_e r_0^2}{\hbar^2} \right)}. \quad (12.89)$$

Критична концентрацію  $n_q$ , за якої  $\tau_{ee} \approx \tau_q$ ,

$$n_q \approx \frac{\varepsilon^2 (kT_e)^2}{e^4 l_q \ln \left( \frac{8m_r^* T_e r_0^2}{\hbar^2} \right)}. \quad (12.90)$$

Звідси маємо

$$\frac{n_E}{n_q} \approx \left( \frac{eEl_q}{kT_e} \right)^2. \quad (12.91)$$

З іншого боку, як впливає із (12.86) та (12.88), права частина співвідношення (12.91) пропорційна  $\tau_q/\tau_E$ . Таким чином, виконання умови квазігідростатичного наближення  $\tau_q \ll \tau_E$  потребує виконання умови  $n_E \ll n_q$ . Тобто критична концентрація, вище якої розсіювання енергії за рахунок міжелектронних взаємодій є домінуючим, має бути набагато меншою від критичної концентрації розсіювання імпульсу.

**3. Визначте умови утворення домену сильного поля в GaAs.**

**Указівка.** Врахуйте, що для утворення домену сильного поля необхідне виконання нерівності  $t_{dr} \gg \tau_M$ , де  $t_{dr}$  – час прольоту домену через зразок, а максвеллівський час релаксації визначається через диференціальну провідність

$$\tau_M = \frac{\varepsilon}{4\pi\sigma_d}.$$

### **Список літератури**

1. *Смит Р.* Полупроводники. – М.: Мир, 1982.
2. *Бонч-Бруевич В.Л., Калашников С.Г.* Физика полупроводников. – М.: Наука, 1990.
3. *Ю Питер, Кардона М.* Основы физики полупроводников. – М.: Физматлит, 2002.
4. *Ридли Б.* Квантовые процессы в полупроводниках. – М.: Мир, 1986.
5. *Kasap S.O.* Principles of Electronic Materials and Devices. – N.Y.: McGraw-Hill, 2005.

# ДОДАТКИ

## ■ Додаток А

### Теорема Блоха для вироджених систем

У розділі 2 ми сформулювали загальне положення: у трансляційно-інваріантній системі власні функції рівняння Шредингера можна записати у вигляді

$$|\mathbf{1}\rangle = e^{i\mathbf{k}\mathbf{l}} |0\rangle.$$

При виведенні цієї формули ми вважали, що стан системи  $|0\rangle$  є невиродженим. Доведемо це твердження у загальнішому випадку вироджених станів. Нехай стан  $|0\rangle$  – двократно вироджений, тобто існують дві функції  $|0\rangle_1$  і  $|0\rangle_2$ , що відповідають одній і тій самій енергії. Тоді трансляція на вектор ґратки  $\mathbf{a}_1$  може привести найбільше до представлення хвильової функції нового стану як лінійної комбінації двох функцій

$$\begin{aligned} |\mathbf{a}_1\rangle_1 &= T_1^{11} |0\rangle_1 + T_1^{12} |0\rangle_2 \\ |\mathbf{a}_1\rangle_2 &= T_1^{21} |0\rangle_1 + T_1^{22} |0\rangle_2 \end{aligned} \quad (\text{A.1})$$

Матриця коефіцієнтів  $\hat{\mathbf{T}}$  має бути унітарною, що вимагається умовами нормування на одиницю. Ці рівняння можна записати у матричному вигляді

$$(|\mathbf{a}_1\rangle) = \hat{\mathbf{T}}_1 (|0\rangle). \quad (\text{A.2})$$

Функції  $|0\rangle_1$  і  $|0\rangle_2$  не визначено однозначно, оскільки замість них можна використати будь-яку пару функцій, що є лінійною комбінацією функцій  $|0\rangle_1$  і  $|0\rangle_2$ . Наприклад, можна записати

$$\begin{aligned} |0\rangle_1 &= S^{11} |0\rangle_1 + S^{12} |0\rangle_2, \\ |0\rangle_2 &= S^{21} |0\rangle_1 + S^{22} |0\rangle_2. \end{aligned} \quad (\text{A.3})$$

Виберемо тепер унітарну матрицю  $\hat{\mathbf{S}}$  так, щоб матриця  $\hat{\mathbf{S}}\hat{\mathbf{T}}_1\hat{\mathbf{S}}^{-1}$  була діагональною. Нехай, наприклад

$$\hat{\mathbf{S}}\hat{\mathbf{T}}_1\hat{\mathbf{S}}^{-1} = \begin{pmatrix} e^{ik_1} & 0 \\ 0 & e^{ik'_1} \end{pmatrix}. \quad (\text{A.4})$$

Аналогічно (A.2) запишемо (A.3) у матричному вигляді

$$(|0\rangle) = \hat{\mathbf{S}} (|0\rangle). \quad (\text{A.5})$$

Зазначимо, що  $(|0\rangle)$  означає вектор-стовпчик

$$(|0\rangle) = \begin{pmatrix} |0\rangle_1 \\ |0\rangle_2 \end{pmatrix}. \quad (\text{A.6})$$

$$\text{Із (A.5) знайдемо} \quad (|0\rangle) = \hat{\mathbf{S}}^{-1}(|0\rangle). \quad (\text{A.7})$$

Підставляючи цей вектор-стовпчик до (A.2), маємо

$$(|\mathbf{a}_1\rangle) = \hat{\mathbf{T}}\hat{\mathbf{S}}^{-1}(|0\rangle). \quad (\text{A.8})$$

Подіємо тепер на цей вектор матрицею  $\hat{\mathbf{S}}$

$$\begin{pmatrix} |\mathbf{a}_1\rangle_1 \\ |\mathbf{a}_1\rangle_2 \end{pmatrix} = \hat{\mathbf{S}}\hat{\mathbf{T}}\hat{\mathbf{S}}^{-1}(|0\rangle) = \begin{pmatrix} e^{ik_1} |0\rangle_1 \\ e^{ik'_1} |0\rangle_2 \end{pmatrix}. \quad (\text{A.9})$$

Таким чином, стани  $|0\rangle_{1,2}$  перетворюються за трансляції так само, як і невироджені. Причому стану  $|0\rangle_1$  відповідає хвильовий вектор із компонентою  $k_1$  у напрямку  $\mathbf{a}_1$ , а стану  $|0\rangle_2$  – хвильовий вектор із компонентою  $k'_1$  у цьому самому напрямку.

Трансляція в іншому напрямку, наприклад  $\mathbf{a}_2$ , визначатиметься матрицею  $\hat{\mathbf{T}}_2$

$$(|\mathbf{a}_2\rangle) = \hat{\mathbf{T}}_2(|0\rangle). \quad (\text{A.10})$$

Розглянемо тепер перетворення за послідовних трансляцій на вектори  $\mathbf{a}_1$  і  $\mathbf{a}_2$ , і навпаки. Можна записати

$$(|\mathbf{a}_1 + \mathbf{a}_2\rangle) = \hat{\mathbf{T}}_1(|\mathbf{a}_2\rangle) = \hat{\mathbf{T}}_1\hat{\mathbf{T}}_2(|0\rangle). \quad (\text{A.11})$$

$$\text{З іншого боку} \quad (|\mathbf{a}_2 + \mathbf{a}_1\rangle) = \hat{\mathbf{T}}_2(|\mathbf{a}_1\rangle) = \hat{\mathbf{T}}_2\hat{\mathbf{T}}_1(|0\rangle). \quad (\text{A.12})$$

$$\text{А оскільки} \quad (|\mathbf{a}_1 + \mathbf{a}_2\rangle) = (|\mathbf{a}_2 + \mathbf{a}_1\rangle), \quad (\text{A.13})$$

то матриці  $\hat{\mathbf{T}}_1$  і  $\hat{\mathbf{T}}_2$  комутовують. Отже, існує унітарна матриця  $\hat{\mathbf{S}}$  така, що за допомогою перетворення типу (A.4) обидві матриці  $\hat{\mathbf{T}}$  можуть бути одночасно приведені до діагонального вигляду. Подібне твердження можна сформулювати і на випадок трансляцій уздовж усіх трьох векторів  $\mathbf{a}_i$ . Таким чином, аналогічно до (A.9) запишемо

$$\begin{aligned} & \begin{pmatrix} |\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3\rangle_1 \\ |\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3\rangle_2 \end{pmatrix} = \\ & = \hat{\mathbf{S}}\hat{\mathbf{T}}_1\hat{\mathbf{T}}_2\hat{\mathbf{T}}_3\hat{\mathbf{S}}^{-1} \begin{pmatrix} |0\rangle_1 \\ |0\rangle_2 \end{pmatrix} = \mathbf{S}\mathbf{T}_1\mathbf{S}^{-1}\mathbf{S}\mathbf{T}_2\mathbf{S}^{-1}\mathbf{S}\mathbf{T}_3\mathbf{S}^{-1} \begin{pmatrix} |0\rangle_1 \\ |0\rangle_2 \end{pmatrix} = \begin{pmatrix} e^{ik_1} e^{ik_2} e^{ik_3} |0\rangle_1 \\ e^{ik'_1} e^{ik'_2} e^{ik'_3} |0\rangle_2 \end{pmatrix}. \end{aligned} \quad (\text{A.14})$$

Звідси видно, що для кожної функції  $|0\rangle_i$  існує хвильовий вектор  $\mathbf{k}$  такий, що за трансляції на довільний вектор ґратки  $\mathbf{l}$  хвильова функція виродженого стану запишеться як

$$|\mathbf{l}\rangle = e^{i\mathbf{k}\mathbf{l}} |0\rangle. \quad (\text{A.15})$$



Ці рівняння складають суть загальної теореми:

**Будь-який розв'язок, що відповідає виродженому значенню енергії, можна представити у вигляді суперпозиції розв'язків, які відповідають тій самій енергії, причому коефіцієнтами в такій суперпозиції виступають комплексні експоненти  $\exp\{ik \cdot l\}$ .**

## ■ Додаток В

### Оператор координати в $\mathbf{k}$ -представленні

Нехай  $\psi_S(T)$  – власна функція оператора  $\hat{\mathbf{S}}$  у  $T$ -представленні, тобто

$$\hat{\mathbf{S}}(T)\psi_S(T) = S\psi_S(T), \quad (\text{B.1})$$

а  $\psi_T(S)$  – власна функція оператора  $\hat{\mathbf{T}}$  в  $S$ -представленні, що задовольняє умові

$$\hat{\mathbf{T}}(S)\psi_T(S) = T\psi_T(S). \quad (\text{B.2})$$

Тоді, як відомо з квантової механіки, виконується співвідношення

$$\psi_S(T) = \psi_T^*(S), \quad (\text{B.3})$$

тобто якщо  $\psi_{\mathbf{r}}(\mathbf{k})$  – власна функція оператора  $\hat{\mathbf{r}}(\mathbf{k})$  у  $\mathbf{k}$ -представленні, то

$$\psi_{\mathbf{r}}(\mathbf{k}) = \psi_{\mathbf{k}}^*(\mathbf{r}). \quad (\text{B.4})$$

Оскільки  $\psi_{\mathbf{k}}^*(\mathbf{r})$  є блохівською функцією

$$\psi_{\mathbf{k}}^*(\mathbf{r}) = e^{-i\mathbf{k}\cdot\mathbf{r}}\chi_{\mathbf{k}}^*(\mathbf{r}), \quad (\text{B.5})$$

маємо із (B.4)

$$\psi_{\mathbf{r}}(\mathbf{k}) = e^{-i\mathbf{k}\cdot\mathbf{r}}\chi_{\mathbf{k}}^*(\mathbf{r}). \quad (\text{B.6})$$

Тепер скористаємось тим фактом, що  $\psi_{\mathbf{r}}(\mathbf{k})$  є власною функцією оператора  $\hat{\mathbf{r}}(\mathbf{k})$

$$\hat{\mathbf{r}}(\mathbf{k})\psi_{\mathbf{r}}(\mathbf{k}) = \mathbf{r}\psi_{\mathbf{r}}(\mathbf{k}). \quad (\text{B.7})$$

Звідси, з огляду на співвідношення (B.4), отримуємо

$$\hat{\mathbf{r}}(\mathbf{k})\psi_{\mathbf{k}}^*(\mathbf{r}) = \mathbf{r}\cdot\psi_{\mathbf{k}}^*(\mathbf{r}), \quad (\text{B.8})$$

тобто оператор  $\hat{\mathbf{r}}(\mathbf{k})$  повинен мати таку структуру, за якої при дії на функцію  $\psi_{\mathbf{k}}^*(\mathbf{r})$  тільки за змінною  $\mathbf{k}$ , отримує ту саму функцію, помножену на вектор  $\mathbf{r}$ . Для визначення цієї структури згадаємо, що в квантовій механіці оператор координати в імпульсному представленні

має вигляд градієнта  $\hat{\mathbf{r}} = i\hbar\nabla_{\mathbf{p}}$ . Беручи до уваги (B.5), подіємо оператором  $\nabla_{\mathbf{k}} = d/d\mathbf{k}$  на функцію  $\psi_{\mathbf{k}}^*(\mathbf{r})$

$$\nabla_{\mathbf{k}}\psi_{\mathbf{k}}^*(\mathbf{r}) = -i\mathbf{r}\psi_{\mathbf{k}}^*(\mathbf{r}) + e^{-i\mathbf{k}\cdot\mathbf{r}}\nabla_{\mathbf{k}}\chi_{\mathbf{k}}^*(\mathbf{r}), \quad (\text{B.9})$$

помножуючи другий доданок у правій частині рівняння на  $\chi_{\mathbf{k}}^*(\mathbf{r})/\chi_{\mathbf{k}}^*(\mathbf{r})$ , отримуємо з (B.9)

$$\nabla_{\mathbf{k}}\psi_{\mathbf{k}}^*(\mathbf{r}) = -i\mathbf{r}\psi_{\mathbf{k}}^*(\mathbf{r}) + \psi_{\mathbf{k}}^*(\mathbf{r}) \cdot \nabla_{\mathbf{k}} \ln\chi_{\mathbf{k}}^*(\mathbf{r}). \quad (\text{B.10})$$

Звідси знайдемо 
$$\mathbf{r}\psi_{\mathbf{k}}^*(\mathbf{r}) = \{i\nabla_{\mathbf{k}} - i\nabla_{\mathbf{k}} \ln\chi_{\mathbf{k}}^*(\mathbf{r})\} \psi_{\mathbf{k}}^*(\mathbf{r}). \quad (\text{B.11})$$

Порівнюючи (B.8) та (B.11), отримаємо вираз для оператора координати в представленні хвильового вектора

$$\mathbf{r}(\mathbf{k}) = i\nabla_{\mathbf{k}} - i(\nabla_{\mathbf{k}} \ln\chi_{\mathbf{k}}^*). \quad (\text{B.12})$$

## ■ Додаток С

### Функція розподілу за багатозарядними домішковими центрами

Раніше (див. 8.5.2) було розглянуто простий випадок, коли домішковий центр нейтральний або одноразово іонізований, тобто може знаходитись тільки в двох зарядових станах. Розглянемо тепер складніший випадок багатозарядних домішкових центрів. Для обчислення ймовірності заповнення електронами багатозарядних центрів необхідно виходити з канонічного рівняння Гіббса.

Нехай ізольована система (що складається із двох підсистем) знаходиться в стані термодинамічної рівноваги. При цьому підсистеми можуть обмінюватись частинками та енергією. Числом частинок, що перейшли із підсистеми II до підсистеми I, є  $n$ . Тоді енергія підсистеми I збільшиться на величину  $E_{(n)}$ , а енергія підсистеми II зменшиться на величину  $E_T - E_{(n)}$  ( $E_T$  – повна енергія системи). Число частинок у підсистемі II тепер складає  $N_0 - n$ . Ентропія підсистеми II зміниться на величину  $\Delta S_{(n)}$ . Тоді ймовірність того, що система I складатиметься з  $n$  частинок  $f_{(n)}$  випливає із формули Больцмана  $S = -k \ln f_{(n)}$

$$f_{(n)} = C \cdot e^{\frac{\Delta S_{(n)}}{k}}. \quad (\text{C1})$$

Згадаємо, що вільна енергія системи записується як  $F = E - TS$ . Використовуючи цю формулу, визначимо, що зміна вільної енергії підсистеми II

$$\Delta F_{(n)} = -E_{(n)} - T\Delta S_{(n)}. \quad (C2)$$

Вільна енергія, що приходить на одну частинку (тобто хімпотенціал), запишеться як

$$\left(\frac{\partial F}{\partial n}\right)_{T,V} = E_F,$$

а вільна енергія підсистеми II – як

$$\Delta F_{(n)} = -nE_F.$$

Звідси знайдемо зміну ентропії

$$\Delta S_{(n)} = \frac{nE_F - E_{(n)}}{T}.$$

Підставляючи цей вираз до (C1), знайдемо функцію розподілу Гіббса для системи, що характеризується змінним числом частинок,

$$f_{(n)} = C \cdot e^{\frac{nE_F - E_{(n)}}{kT}}. \quad (C3)$$

Стаду  $C$  визначимо з умови нормування функції розподілу на одиницю

$$\sum_{n=0}^N f_{(n)} = 1 \rightarrow C^{-1} = \sum_{n=0}^N e^{\frac{nE_F - E_{(n)}}{kT}}.$$

Нехай  $g_n$  – кратність виродження  $n$ -разово зарядженого домішкового центра. Тоді із (C3) отримаємо

$$f_{(n)} = \frac{g_n e^{\frac{nE_F - E_{(n)}}{kT}}}{\sum_{n'=0}^N g_{n'} e^{\frac{n'E_F - E_{(n')}}{kT}}}. \quad (C4)$$

Оскільки через  $E_{(n)}$  тут позначено енергію системи, що складається з  $n$  частинок, то  $E_{(n)} = E_1 + E_2 + E_3 + \dots + E_n$ , де  $E_i$  – енергія  $i$ -го рівня в системі.

Наприклад, якщо центр може бути нейтральний, одно- або дворазово іонізований, то ймовірність того, що він:

♦ **нейтральний**

$$f_{(0)} = \frac{g_0 e^{\frac{0 \cdot E_F - 0}{kT}}}{g_0 + g_1 e^{\frac{E_F - E_1}{kT}} + g_2 e^{\frac{2E_F - E_2 - E_1}{kT}}}; \quad (C5)$$

$$\blacklozenge \text{ одноразово іонізований} \quad f_{(1)} = \frac{g_1 e^{\frac{E_F - E_1}{kT}}}{g_0 + g_1 e^{\frac{E_F - E_1}{kT}} + g_2 e^{\frac{2E_F - E_2 - E_1}{kT}}}; \quad (C6)$$

$$\blacklozenge \text{ дворазово іонізований} \quad f_{(2)} = \frac{g_2 e^{\frac{2E_F - E_1 - E_2}{kT}}}{g_0 + g_1 e^{\frac{E_F - E_1}{kT}} + g_2 e^{\frac{2E_F - E_2 - E_1}{kT}}}. \quad (C7)$$

Тут ми скористались очевидним фактом, що енергія "нульового" рівня на центрі (носій відсутній на центрі) дорівнює нулю.

## ■ Додаток D

### Час релаксації носіїв заряду за розсіювання на іонах домішок

Покажемо як обчислити час релаксації носіїв за розсіювання на іонах домішки. Задача полягає у знаходженні траєкторії руху зарядженої частинки в центральному полі іона домішки. Потенціальною енергією взаємодії іона з носієм заряду  $e$

$$U(r) = \pm \frac{Ze^2}{\epsilon_r r}, \quad (D1)$$

звідси швидкість руху частинки

$$\frac{dr}{dt} = \sqrt{\frac{2}{m} [E - U(r)] - \frac{M^2}{m^2 r^2}}. \quad (D3)$$

Тут  $M$  – момент імпульсу. Оскільки  $M = mr^2 \dot{\varphi}$ , то

$$d\varphi = \frac{M}{mr^2} dt. \quad (D4)$$

Із (D.3) отримуємо

$$dt = \frac{dr}{\sqrt{\frac{2}{m} [E - U(r)] - \frac{M^2}{m^2 r^2}}}.$$

Підставляючи цей вираз до (D.4) та інтегруючи за  $r$ , отримуємо

$$\varphi = \int \frac{\frac{M}{mr^2} dr}{\sqrt{\frac{2}{m}[E - U(r)] - \frac{M^2}{m^2 r^2}}} + \text{const} \cdot \quad (D5)$$

Запишемо, як змінюється кут  $\varphi$  за зміни відстані від частинки, що рухається за траєкторією розсіювання між двома симетричними точками, які характеризуються деякими значеннями радіус-вектора  $r_{\max}$  (рис. D1). Зважаючи на те, що траєкторія руху частинки є кривою, симетричною щодо лінії  $OA$ , напрямком якої задається найменшим значенням відстані між частинкою та центром розсіювання  $r_{\min}$ , запишемо

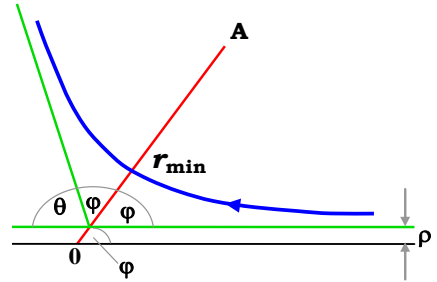


Рис. D1

$$\Delta\varphi = 2 \int_{r_{\min}}^{r_{\max}} \frac{\frac{M}{mr^2} dr}{\sqrt{\frac{2}{m}[E - U(r)] - \frac{M^2}{m^2 r^2}}} \cdot \quad (D6)$$

Позначимо енергію та момент імпульсу через прицільний параметр і швидкість на нескінченності

$$E = \frac{mv_{\infty}^2}{2}, \quad M = mv_{\infty}\rho,$$

тоді, маючи на увазі, що повна зміна кута  $\varphi$  за руху частинки від нескінченності до нескінченності дорівнюватиме  $2\varphi$ , (D.6) переписеться у вигляді

$$\varphi = \int_{r_{\min}}^{\infty} \frac{\frac{\rho}{r^2} dr}{\sqrt{1 - \frac{\rho}{r^2} - \frac{2U}{mv_{\infty}^2}}} \cdot \quad (D7)$$

Як зазначалось,  $r_{\min}$  – мінімальна відстань від центру розсіювання до траєкторії руху частинки. Якщо провести пряму, що з'єднає центр розсіювання з цією точкою, то кут між асимптотикою траєкторії на нескінченності й цією прямою дорівнюватиме  $\varphi$ , а точка повороту (відстань від неї до центру розсіювання дорівнює  $r_{\min}$ ) визначатиметься зміною знаку підкореневого виразу, тобто  $r_{\min}$  є коренем виразу, що стоїть під знаком радикала. Підставимо сюди потенціал як функцію відстані від центру розсіювання до частинки  $U = a/r$ . Введемо позначення  $\beta = a/(mv_{\infty}^2\rho)$  і запишемо інтеграл (D.7) у вигляді

$$\varphi = - \int_{r_{\min}}^{\infty} \frac{d \frac{\rho}{r}}{\sqrt{1 - \frac{\rho^2}{r^2} - 2\beta \frac{\rho}{r}}} = \int_0^{x_m} \frac{dx}{\sqrt{1 - (x + \beta)^2 + \beta^2}}. \quad (D.8)$$

Причому,  $x_m$ :  $x_m + \beta = \sqrt{1 + \beta^2}$ . Оскільки  $\int \frac{dx}{\sqrt{A^2 - x^2}} = \arcsin \frac{x}{A}$ , то інтегрування в (D.8) дає

$$\varphi = \arcsin \frac{x + \beta}{\sqrt{1 + \beta^2}} \Big|_0^{x_m} = \frac{\pi}{2} - \arcsin \frac{\beta}{\sqrt{1 + \beta^2}} = \arccos \frac{\beta}{\sqrt{1 + \beta^2}}. \quad (D.9)$$

Із цього рівняння маємо

$$\cos \varphi = \frac{B/\rho}{\sqrt{1 + B^2/\rho^2}}, \quad \cos^2 \varphi = \frac{B^2}{\rho^2} [1 - \cos^2 \varphi], \quad \rho^2 = B^2 \operatorname{tg}^2 \varphi. \quad (D.10)$$

Кут розсіювання  $\theta$  пов'язаний із кутом  $\varphi$  співвідношенням  $\theta = |\pi - 2\varphi|$ . Тоді із (D.10) отримуємо

$$\rho^2 = \frac{a^2}{m^2 v^4} \operatorname{ctg}^2 \frac{\theta}{2}. \quad (D.11)$$

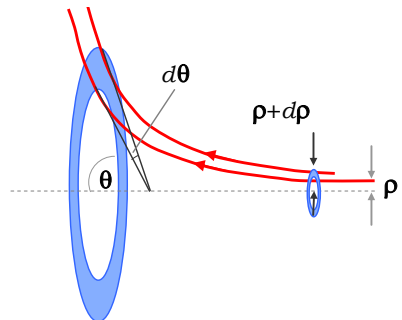
Оскільки для нашої задачі  $m = m^*$ ,  $a = Ze^2/\epsilon_r$ , то замість (D.11) можна записати

$$\rho^2 = \frac{Z^2 e^4}{\epsilon_r^2 m^2 v^4} \operatorname{ctg}^4 \frac{\theta}{2}. \quad (D.12)$$

Тобто при розсіюванні на центральносиметричному потенціалі  $U(r)$  на кут  $\theta$  розсіюються ті частинки, що налітають на центр розсіювання із прицільним параметром  $\rho$ . Візьмемо до уваги, що центр розсіювання є вісесиметричним. Тоді частинки, що розсіюються між кутами  $\theta$  та  $\theta = d\theta$ , попадають у тілесний кут

$$d\Omega = 2\pi \sin \theta d\theta. \quad (D.13)$$

Нехай  $dN$  – кількість частинок, що розсіюються за одиницю часу в конусний сегмент між кутами  $\theta$  та  $\theta + d\theta$ . Тоді, якщо  $dn$  – кількість частинок, які проходять за одиницю часу через одиничну площину перерізу падаючого пучка, що попадає в кільце між  $\rho$  та  $\rho + d\rho$  (рис. D.2), то відносною кількістю частинок, що розсіюються у конусний сегмент між



кутами  $\theta$  та  $\theta + d\theta$ , є  $\sigma = dN/dn$ .

Рис. D2

Нехай частинки рухаються хаотично із середньою швидкістю  $v$ . Тоді через одиничну площадку за одиницю часу пройде  $nv$  частинок. Повне число частинок, що відхилилися за одиницю часу на кут  $\theta$ , дорівнюватиме потоку частинок, що падають на кільце, яке визначається інтервалом прицільних параметрів  $(\rho, \rho + d\rho)$ . Площа кільцевої ділянки  $dS = 2\pi \cdot \rho |d\rho|$ . Оскільки повне число частинок зберігається, то числом частинок, що набігають на центр розсіювання за одиницю часу, є  $nv d\Omega$ . Тоді переріз розсіювання, що визначається як відношення числа відхилених частинок на число частинок, які набігають на розсіювач,

$$\sigma(\theta) = \frac{nv2\pi\rho |d\rho|}{nv\sin\theta d\theta}. \quad (\text{D.14})$$

Використовуючи (D.12), маємо

$$2\rho |d\rho| = \frac{Z^2 e^4}{\varepsilon_r^2 m^2 v^4} \cdot \text{ctg} \frac{\theta}{2} \frac{1}{\sin^2 \frac{\theta}{2}} d\theta. \quad (\text{D.15})$$

Підставимо цей вираз до (D.14) і, зважаючи на те, що  $\text{ctg}(\theta/2)/\sin^2(\theta/2) = \sin\theta/(2\sin^4(\theta/2))$ , остаточно отримуємо

$$\sigma(\theta) = \frac{Z^2 e^4}{\varepsilon_r^2 m^2 v^4} \cdot \frac{1}{\sin^4 \frac{\theta}{2}}. \quad (\text{D.16})$$

Використаємо формулу, що пов'язує час релаксації з ефективним перерізом провідності (див. (9.35) та (9.37)). Але, якщо безпосередньо підставити до інтегралу

$$\sigma_C = 2\pi \cdot \int_0^\pi \sigma(\vartheta) [1 - \cos\vartheta] \sin\vartheta \cdot d\vartheta \quad (\text{D.17})$$

функцію  $\sigma(\theta)$ , то побачимо, що на нижній границі він розходиться. Для уникнення такої ситуації зауважимо: насправді переріз розсіювання обмежений, що пов'язано з існуванням деякого максимального значення прицільного параметра. Останній відповідає мінімальному значенню кута, на який у кристалі відхиляються носії. Якщо іони домішки розподілені рівномірно за об'ємом кристала зі щільністю  $N_I$ , то середньою відстанню між ними буде  $(N_I)^{-1/3}$ . Зрозуміло, що максимальний прицільний параметр тепер дорівнює  $(1/2)(N_I)^{-1/3}$ , тобто використовуючи (D.11), маємо

$$\text{ctg} \frac{\theta_{\min}}{2} = \frac{\varepsilon_r m^* v^2}{Ze^2} \rho_{\max} = \frac{\varepsilon_r m^* v^2}{2Ze^2 N_I^{1/3}}. \quad (\text{D.18})$$

При обчисленні  $\sigma_C$  маємо інтеграл

$$\int_{\theta_{\min}}^{\pi} d\theta \frac{(1 - \cos\theta)\sin\theta}{\sin^4 \frac{\theta}{2}} = 8 \int_{\theta_{\min}}^{\pi} d\theta \frac{\cos \frac{\theta}{2}}{\sin \frac{\theta}{2}} = -4 \ln \left( \sin^2 \frac{\theta_{\min}}{2} \right). \quad (\text{D.19})$$

$$\text{Але } -4 \ln \left( \sin^2 \frac{\theta_{\min}}{2} \right) = 4 \ln \frac{1}{\sin^2 \frac{\theta_{\min}}{2}} = 4 \ln(1 + \text{ctg}^2 \frac{\theta_{\min}}{2}) = 4 \ln \left( 1 + \frac{\varepsilon_r^2 (m^*)^2 v^4}{4Z^2 e^4 N_I^{2/3}} \right). \quad (\text{D.20})$$

Таким чином, остаточно знаходимо переріз розсіювання та час релаксації носіїв заряду на іонізованій домішці

$$\sigma_c = 2\pi \left( \frac{Ze^2}{\varepsilon_r m^* v^2} \right)^2 \ln \left( 1 + \frac{\varepsilon_r^2 (m^*)^2 v^4}{4Z^2 e^4 N_I^{2/3}} \right) \quad (\text{D.21})$$

і час релаксації електрона при розсіюванні на іонізованій домішці

$$\tau_I = \frac{\varepsilon_r^2 (m^*)^2 v^3}{2\pi \cdot Z^2 e^4 N_I \ln \left( 1 + \frac{\varepsilon_r^2 (m^*)^2 v^4}{4Z^2 e^4 N_I^{2/3}} \right)}. \quad (\text{D.22})$$

Оскільки енергія частинки у невиродженому електронному газі складає  $\sim kT$ , то другий доданок в аргументі логарифму є великим, й одиницею можна нехтувати. Крім того, при врахуванні ефекту екранування, тобто при використанні потенціалу типу

$$U(r) = \frac{\alpha}{r} e^{-r/r_0}$$

(D22) набуде вигляду

$$\tau_I = \frac{\varepsilon_r^2 E^{3/2} (2m^*)^{1/2}}{\pi \cdot Z^2 e^4 N_I \ln \left( \frac{8m^* E r_0^2}{\hbar^2} \right)}, \quad (\text{D23})$$

де  $E$  – енергія електрона.

## ■ Додаток Е

### Енергія, яку втрачає електрон за розсіювання на оптичному фоні

Для з'ясування походження формули для енергії, яку втрачає носій за одиничний акт зіткнення  $\Delta W = \hbar\omega_0 \text{th}(\hbar\omega_0/2kT)$  (див. (12.24)), припустимо, що ймовірністю переходу електрона зі стану  $\mathbf{k}$  до стану  $\mathbf{k}'$  з



емітуванням фонона, що характеризується імпульсом  $\mathbf{q}$ , є  $W^-(\mathbf{k}, \mathbf{k}')$  ( $\mathbf{k} = \mathbf{k} - \mathbf{q}$ ), а ймовірністю переходу електрона із поглинанням фонона є  $W^+(\mathbf{k}, \mathbf{k}')$ . Тоді енергія, що передається ґратці за один акт зіткнення електрона із ґраткою,

$$\Delta E = \hbar\omega_0 \frac{W^-(\mathbf{k}, \mathbf{k} - \mathbf{q}) - W^+(\mathbf{k}, \mathbf{k} + \mathbf{q})}{W^-(\mathbf{k}, \mathbf{k} - \mathbf{q}) + W^+(\mathbf{k}, \mathbf{k} + \mathbf{q})}. \quad (\text{E1})$$

Ймовірність переходу електрона зі стану  $\mathbf{k}$  до стану  $\mathbf{k}'$  з емітуванням фонона, що характеризується імпульсом  $\mathbf{q}$ , можна обчислити як

$$W^-(\mathbf{k}, \mathbf{k} - \mathbf{q}) \cong \omega(\mathbf{q})(N_q + 1). \quad (\text{E2})$$

Аналогічно для переходу електрона із поглинанням фонона

$$W^+(\mathbf{k}, \mathbf{k} + \mathbf{q}) \cong \omega(\mathbf{q})N_q. \quad (\text{E3})$$

Тоді

$$\frac{W^- - W^+}{W^- + W^+} = \frac{N_q + 1 - N_q}{N_q + 1 + N_q} = \frac{1}{2N_q + 1} = \frac{1}{2 \frac{1}{e^{\hbar\omega_0/kT} - 1} + 1}.$$

Спростимо отриманий вираз

$$\frac{1}{2 \frac{1}{e^{\hbar\omega_0/kT} - 1} + 1} = \frac{e^{\hbar\omega_0/kT} - 1}{e^{\hbar\omega_0/kT} + 1} = th\left(\frac{\hbar\omega_0}{2kT}\right).$$

Підставляючи цей вираз до (E1), знайдемо частину енергії, яку передає електрон ґратці за розсіювання на оптичному фоні

$$\Delta W = \hbar\omega_0 th(\hbar\omega_0/2kT). \quad (\text{E4})$$

## ■ Додаток F

### Визначення відношення коефіцієнта дифузії до рухливості

Еволюція профілю просторового розподілу неосновних носіїв заряду в області, що дрейфує під дією електричного поля, визначається рівнянням

$$\delta p = \frac{\delta p_0}{2\sqrt{\pi D_p t}} \exp(-t/\tau_p) \exp(-x^2/4D_p t).$$

Знайдемо півширини профілів розподілу дірок у два моменти часу  $t_1$  та  $t_2$ :

$$\Delta x(t_1) = 4\sqrt{D_p t_1 \ln 2}, \quad (\text{F.1})$$

$$\Delta x(t_2) = 4\sqrt{D_p t_2 \ln 2}. \quad (\text{F.2})$$

Використовуючи ці формули, маємо

$$\Delta x^2(t_1) - \Delta x^2(t_2) = 16D_p \ln 2 (t_1 - t_2). \quad (\text{F.3})$$

Перепишемо (11.4) у вигляді  $\Delta x(t_1) = v_{dr} \Delta t_1$ . Тоді, беручи до уваги, що вимірювання проводяться за незмінної величини струму, тобто сталої дрейфової швидкості  $v_{dr} = \mu_p E$ , отримуємо із (F.3)

$$\mu_p^2 E^2 (\Delta t_1^2 - \Delta t_2^2) = 16D_p \ln 2 (t_1 - t_2). \quad (\text{F.4})$$

Виразимо електричне поле через напругу та відстань між емітером і колектором, скажімо у першому вимірюванні. Отримаємо

$$\mu_p v_{dr} \frac{U_1}{l_1} = \frac{16D_p \ln 2 (t_1 - t_2)}{\Delta t_1^2 - \Delta t_2^2},$$

звідки знаходимо остаточно

$$\frac{D_p}{\mu_p} = \frac{U_1}{16 \ln 2} \frac{\Delta t_1^2 - \Delta t_2^2}{t_1 (t_1 - t_2)}. \quad (\text{F.5})$$