

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

В. О. Москалюк, В.І. Тимофєєв, Т.А. Саурова

ФІЗИКА ЕЛЕКТРОННИХ ПРОЦЕСІВ

*Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як підручник для студентів, які навчаються
за спеціальністю 153 «Мікро- та наносистемна техніка»,
спеціалізацією «Електронні біомедичні системи і технології»,
«Інформаційні технології проектування в електроніці та
наносистемах»*

Київ
КПІ ім. Ігоря Сікорського
2020

Рецензенти: Осинський В. І., д. т. н., професор, директор Центру оптоелектронних технологій НДІ Мікроприладів» НТК «ІМК» НАН України
Гончаров О. А., д.ф.-м.н., професор, головн. наук. співробітник інституту фізики НАНУ

Відповідальний редактор В.І. Тимофєєв, д-р.т.н., професор

Гриф надано Методичною радою КПІ ім. Ігоря Сікорського (протокол № від XX.XX.2020 р.) за поданням Вченої ради факультету електроніки (протокол №XX/2020 від XX.XX.2020 р.)

Електронне мережне навчальне видання

Москалюк Володимир Олександрович, канд. техн. наук, професор
Тимофєєв Володимир Іванович, доктор техню наук, професор
Саурова Тетяна Асадівна, канд. техн. наук, доцент

ФІЗИКА ЕЛЕКТРОННИХ ПРОЦЕСІВ

Фізика електронних процесів,: [Електронний ресурс]: підручник для студ. спеціальності 153 «Мікро- та наносистемна техніка», спеціалізації «Електронні біомедичні системи і технології», «Інформаційні технології проектування в електроніці та наносистемах»/ В. О. Москалюк, В.І.Тимофєєв, Т.А.Саурова; КПІ ім. Ігоря Сікорського. – Електронні текстові данні (1 файл: 7,21 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2020. –324 с.

Наведено текстовий та ілюстративний матеріал до лекцій з курсу «Фізика електронних процесів». За змістом розглянуто властивості усупільнених і локалізованих станів електронів у твердих тілах, в тому числі у кристалах низької розмірності. Опис ведеться з використанням дисперсійних кривих, наведено особливості законів дисперсії для основних напівпровідникових матеріалів. Викладено найпростіші моделі дрейфового і дифузійного руху носіїв; на основі опису рівноважного стану та механізмів розсіювання носіїв розглянуто питання провідності твердих тіл. Розглянуто дві групи ефектів у сильному електричному полі: особливості процесів транспорту гарячих електронів та польової генерації носіїв. Окремий розділ присвячено фотоелектричним ефектам, їх порівняльним властивостям та характерним фотоприймачам. В останньому розділі розглянуто фізичні процеси у найпростіших структурах твердотільних компонентів, а також в електронно-дірковому переході і сучасних транзисторних структурах.

© В. О. Москалюк, В.І.Тимофєєв, Т.А.Саурова; 2020

© КПІ ім. Ігоря Сікорського, 2019

ЗМІСТ

ПЕРЕДМОВА	6
1. ЕЛЕКТРОННІ СТАНИ У КРИСТАЛАХ	8
1.1. Усуспільнені стани	8
1.1.1. Утворення енергетичних зон	8
1.1.2. Загальні властивості хвильової функції	12
1.1.3. Закон дисперсії електрона у кристалі	18
1.1.4. Ефективна маса електрона	22
1.1.5. Розподіл енергетичних станів	26
1.1.6. Зонна структура напівпровідників	31
1.2. Локалізовані стани електронів	41
1.2.1. Донорні домішки	41
1.2.2. Акцепторні домішки	44
1.2.3. Багатозарядні домішки, глибокі рівні	46
1.2.4. Вплив концентрації на зонну структуру	48
1.3. Електронні стани на поверхні	51
1.3.1. Макроскопічна енергетична модель поверхні	51
1.3.2. Таммівські поверхневі стани	54
1.3.3. Адсорбовані поверхневі стани	55
1.4. Електронні стани та механізми провідності	58
1.4.1. Зв'язок зонної структури та провідності	58
1.4.2. Електрони у повністю заповненій зоні	61
1.4.3. Види провідності і носії заряду	63
<i>Контрольні питання та вправи</i>	67
2. ЕЛЕКТРОННИЙ ТРАНСПОРТ	69
2.1. Параметри плазми	73
2.2. Дрейф носіїв. Одночасткова модель	74
2.3. Дрейф носіїв. Колективна модель	79
2.3.1. Кінетичне рівняння	79
2.3.2. Релаксаційні рівняння, часи релаксації	88
2.4. Розсіювання носіїв, рухливість	93
2.4.1. Фононні спектри	94
2.4.2. Основні закономірності	98
2.4.3. Види розсіювання	100
2.4.4. Рухливість електронів	110
2.5. Дифузія носіїв заряду	118
2.6. Рівняння неперервності	123
<i>Контрольні питання та вправи</i>	124
3. РІВНОВАЖНИЙ СТАН НОСІЇВ	126
3.1. Основні положення статистики електронів	126
3.2. Розрахунок концентрації рівноважних носіїв	130
3.2.1. Електрони у зоні провідності	130
3.2.2. Дірки у валентній зоні	133
3.2.3. Електрони у кристалах низької розмірності	137
3.2.4. Заповнення локалізованих станів	139
3.3. Провідність металів	141
3.4. Власна провідність напівпровідників	145
3.5. Домішкова провідність	148

3.5.1. Донорні напівпровідники	153
3.5.2. Акцепторні напівпровідники	158
3.6. Рівноважні носії в електричному полі	160
3.6.1. «Викривлення» енергетичних зон	160
3.6.2. Зміна концентрацій рухливих носіїв	162
3.7. Термоелектронна емісія	164
3.7.1. Закон термоелектронної емісії	165
3.7.2. Вплив електричного поля на роботу виходу	170
3.7.3. Термоелектронні емітери	175
<i>Контрольні питання та вправи</i>	<i>179</i>
4. ГАРЯЧІ НОСІЇ У СИЛЬНОМУ ЕЛЕКТРИЧНОМУ ПОЛІ.....	180
4.1. Дрейф носіїв.....	180
4.1.1. Розігрівання електронного газу	181
4.1.2. Насичення дрейфової швидкості	188
4.1.3. Динамічна провідність	194
4.1.4. «Сплеск» дрейфової швидкості	198
4.1.5. Ефект Ганна.....	201
4.1.6. Дифузія гарячих електронів	204
4.2. Польова генерація.....	206
4.2.1. Ударна іонізація.....	206
4.2.2. Електростатична іонізація	212
4.2.3. Іонізація у домішкових напівпровідниках	214
4.2.4. Автоелектронна емісія	215
<i>Контрольні запитання та вправи</i>	<i>218</i>
5. РЕКОМБІНАЦІЯ НОСІЇВ ЗАРЯДУ.....	220
5.1. Загальні закономірності рекомбінації	220
5.2. Електронно-діркова рекомбінація у напівпровідниках	223
5.3. Рекомбінація електронів та іонів у газовій плазмі	229
<i>Контрольні запитання та вправи</i>	<i>230</i>
6. ОПТИЧНА ГЕНЕРАЦІЯ.....	231
6.1. Особливості поглинання світла	231
6.2. Властивості фотоефектів	237
6.2.1. Поведінка збуджених електронів	237
6.2.2. Інерційність фотоефектів	239
6.2.3. Лінійність світлової характеристики	243
6.2.4. Чутливість фотоефектів	245
6.2.5. Гранична частота фотоефектів	246
6.3. Фотоеlementи.....	249
6.3.1. Фотоемітери.....	249
6.3.2. Фоторезистивні матеріали	253
<i>Контрольні запитання та вправи</i>	<i>255</i>
7. ТВЕРДОТІЛЬНІ СТРУКТУРИ І ПРОЦЕСИ	256
7.1. Ефект поля	256
7.2. Контакт двох металів	266
7.3. Інжекція нерівноважних носіїв	269
7.4. Електронно-дірковий перехід	271
7.4.1. Рівноважний стан переходу	272
7.4.2. Товщина електронно-діркового переходу	277
7.4.3. Вольт-амперна характеристика переходу	280

7.4.4. Ємності електронно-діркового переходу.....	287
7.5. Контакти метал- напівпровідник	289
7.6. n^+n – контакти	293
7.7. Гетеропереходи	295
7.8. Транзисторні структури	300
7.8.1. Гетеробіполярні структури	300
7.8.2. Субмікронні польові транзистори з бар'єром Шоттки	302
7.8.3. Гетероструктурні польові транзистори.....	307
<i>Контрольні запитання та вправи</i>	<i>311</i>
Д1. Модель Кроніга-Пенні	313
Д2. Інтеграл Фермі-Дірака.....	317
Д3. Надбар'єрне проходження електронів	319
РЕКОМЕНДОВАНА ЛІТЕРАТУРА.....	323

ПЕРЕДМОВА

Сучасні тенденції розвитку електроніки викликані нестримним бажанням людства підвищувати швидкодію комп'ютерів і освоювати нові інформаційні технології, покращувати якість зв'язку і передачі інформації, підвищувати надійність локаційних засобів і вирішувати нові технічні проблеми за допомогою електронної техніки. Такі завдання вимагають відповідних швидкостей роботи активних електронних компонентів, що у свою чергу змушує використовувати для їх створення як нові матеріали так і усе більш тонкі фізичні ефекти. Крім того, цілком природним і необхідним виглядає також перехід від мікрометрових характерних розмірів приладів до нанометрових, які освоюються сучасною технологією. За цих розмірів все помітніше проявляються квантові ефекти, які можуть бути використані для створення принципово нових приладів.

Навчальна література, не лише призначена для спецкурсів, але й та, яку прийнято відносити до фундаментальної, не може залишатися у стороні від рівня і тенденцій розвитку електроніки. Пропонований підручник є спробою відстежити ці тенденції і дати майбутнім фахівцям основи знань, необхідних для розуміння фізичних процесів, особливо динамічних, що відбуваються у сучасних приладах.

У першій частині розглянуто усупільнені і локалізовані електронні стани у кристалах, які зазвичай вивчають у курсах з фізики твердого тіла. Але на відміну від традиційного викладання, опис ведеться не на основі зонної теорії за допомогою «одновимірних» енергетичних діаграм, а з використанням закону дисперсії, який дозволяє розглядати не лише зміни енергії, але і імпульсу. Без цього неможливе розуміння особливостей процесів взаємодії електронів як з кристалічною ґраткою так і з різними частинками і квазічастинками. Крім того, знання деталей законів дисперсії дозволяє зрозуміти специфічні властивості різних напівпровідникових матеріалів. Надано увагу також електронним

станам в одно- і двовимірних кристалах, які перспективні для створення твердотільних нанометрових структур.

Детальніше, ніж за традиційного викладання, розглянуті напівкласичні рівняння руху носіїв і механізми провідності, описано дві моделі дрейфу носіїв в статичних полях, одночастинкова та колективна, і особливості дифузійного руху заряджених часток, взаємодія дифузійних та дрейфових процесів. Завершується перша частина розглядом особливостей рівноважного стану електронно-діркових колективів у твердих тілах і їх провідності.

Важливим змістовним розділом, якому надається недостатньо уваги у навчальній літературі, є розділ про процеси у сильних електричних полях, які мають місце практично в кожному приладі з мікрометровими розмірами. Розглянуто процеси розігрівання електронного газу, насичення дрейфової швидкості електронів, умови та особливості її перевищення над стаціонарними значеннями, балістичний проліт електронів, питання динамічної провідності, у тому числі від'ємної, особливості дифузії гарячих електронів.

У розділі про фотоелектричні процеси наголос робиться на відмінностях характерних властивостей та параметрів внутрішнього та зовнішнього фото ефектів, наводяться типові структури фотоприймачів.

Останній розділ присвячено фізичним процесам у найпростіших структурах. Це, по-перше, ефект поля, як основа функціонування численної групи електронних компонентів – польових приладів. По-друге, електронно-дірковий перехід, як вихідна фізична модель другої важливої групи – біполярних приладів. Розглянуто властивості деяких видів контактів, серед яких контакт Шоттки, гетероконтакти тощо. Також розглянуто три типи транзисторних структур: гетеробіполярний транзистор, польовий транзистор з затвором Шоттки та гетероструктурний польовий транзистор, яким у навчальній літературі надається недостатньо уваги, непропорційно їх ролі в електронній техніці.

1. ЕЛЕКТРОННІ СТАНИ У КРИСТАЛАХ

1.1. УСУСПІЛЬНЕНІ СТАНИ

1.1.1. Утворення енергетичних зон

Зонна теорія описує характер і особливості електронних станів у кристалічних твердих тілах. Основні властивості поведінки електронів у кристалах можна зрозуміти з досить простих загальних уявлень. Відомо, що енергетичний спектр електронів в окремому атомі є дискретним. У твердому тілі валентні електрони, які знаходились у своїх атомах на однакових рівнях, у разі об'єднання атомів у кристалічні ґратки, усупільнюються в колективи і мають бути рівноправними, залишаючись в той же час квантовими частками, енергія яких квантується. Але з іншого боку електрони одного колективу не можуть знаходитися в однакових квантових станах відповідно до принципу Паулі. Узгоджує ці дві, на перший погляд протирічні, вимоги модель, відповідно до якої енергетичні стани електронів у кристалі порівняно з атомарними станами дещо змінюються внаслідок взаємодії атомів. За таких умов кожен електронний енергетичний рівень ізольованого атома з об'єднанням атомів у кристал розщеплюється на підрівні, які і утворюють енергетичну полосу або зону.

Міра розщеплювання атомарного рівня на підрівні, тобто ширина енергетичної зони, залежить від взаємодії між атомами. Чим ближче розташовані атоми, тим більша енергія взаємодії між ними і тим істотніше розщеплювання рівня. Це показано на рис. 1.1, звідки видно, що енергетичні рівні електронів в ізольованих атомах зі зменшенням відстані між ними перетворюються на енергетичні зони. Енергетичні зони, які утворилися з різнойменних рівнів, можуть залежно від певних умов перекриватися або не перекриватися. В останньому випадку, зони дозволених значень енергії розділені так званими *забороненими зонами*, які є інтервалами значень енергії, яких електрони мати не можуть. Треба відзначити, що наведені міркування стосуються дозволених рівнів (як заповнених, так і вакантних), а

характер розподілу електронів у зонах визначається конкретним типом кристала.

Кількісний аналіз електронних станів у твердому тілі можна провести на основі рівняння Шредінгера. Але його точний розв'язок для кристала неможливий, а спрощений аналіз проводиться за деяких припущень і наближень. Найчастіше використовується адіабатичне і одноелектронне наближення.

Адіабатичне наближення полягає в тому, що не враховується тепловий рух кристалічної ґратки. Це виправдано тим, що швидкість руху електронів значно вища за швидкість ядер, і за час, потрібний для зміни стану електронів, положення ядер практично не змінюється.

Одноелектронне наближення дозволяє аналізувати поведінку електронного колективу в твердому тілі на основі розв'язку рівняння Шредінгера для одного електрона

$$\nabla^2 \psi + \frac{2m}{\hbar^2} [E - U(\vec{r})] \psi = 0, \quad (1.1)$$

де вигляд функції потенціальної енергії $U(\vec{r})$ враховує силовий вплив на даний електрон деякого усередненого електричного поля, яке створюється іншими електронами і атомними залишками кристалічної ґратки.

Але навіть ці досить сильні припущення ще не дають можливості розв'язати рівняння (1.1) через невизначеність і складність знаходження потенціальної функції $U(\vec{r})$ для реальних кристалів. Існують різноманітні методи, що відрізняються у першу чергу видом апроксимації функції $U(\vec{r})$, яка дозволяє розв'язати рівняння Шредінгера. Найбільш наочними з методичної точки зору є два методи, що призводять до уявлення про зонний характер енергетичного спектру електронів у кристалі.

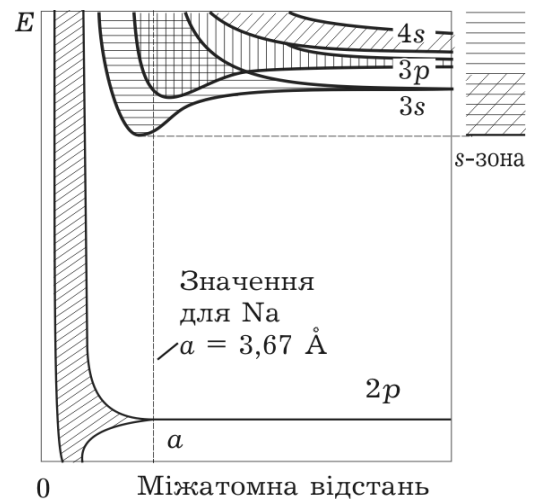


Рис. 1.1. Розщеплення енергетичних станів атомів натрію зі зменшенням міжатомних відстаней. 3s-зона заповнюється лише наполовину

У методі сильного зв'язку приймається, що електрон сильно зв'язаний з атомом і рухається у кулонівському полі ядра та електронних оболонок (як в ізольованому атомі), а дія інших часток ґратки розглядається як мале збурення. За таких умов $U(\vec{r})$ представляється у вигляді суми потенціальної функції електрона в ізольованому атомі $U_a(\vec{r})$ і деякої функції $\Delta U(\vec{r})$, що враховує вплив інших атомів. Природно очікувати, що у цьому випадку енергетичний спектр електронів залишається в основному подібним до спектра в ізольованому атомі, а збурення, що вноситься присутністю сусідніх атомів, буде тим більшим, чим далі електрон знаходиться від ядра. Якісна картина результатів такого аналізу наведена на рис. 1.2, де для порівняння зображені енергетичні діаграми електрона в ізольованому атомі (а) і в ланцюжку близько розташованих атомів (б).

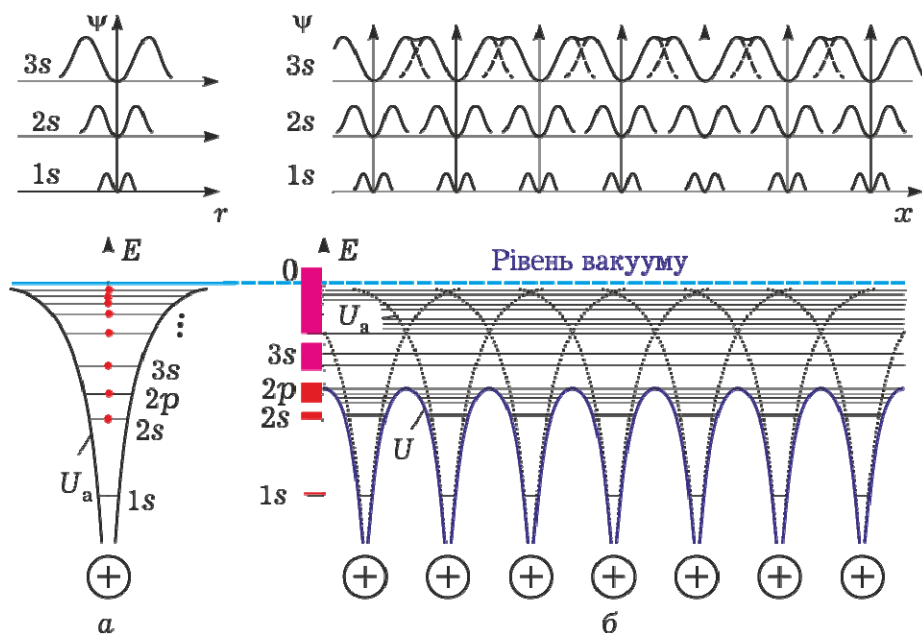


Рис. 1.2. Потенціальна енергія електрона у полі окремого атома (а) і періодичний потенціальний рельєф кристала (б). Утворені з атомних рівнів зони тим ширші, чим сильніше перекриваються відповідні хвильові функції, умовно зображені у верхній частині рисунка

Збурення, викликане присутністю сусідніх атомів, призводить до утворення між атомами потенціальних бар'єрів скінченної висоти, які є «прозорими» для електронів завдяки тунельному ефекту. Це означає, що електрони, які знаходилися в ізольованих атомах в однакових квантових станах, утворюють зв'язану систему, для якої енергетичний спектр є сукупністю близько розташованих підрівнів (енергетичною зоною). Природно, що це збурення є тим більшим, чим

сильніше перекриваються хвильові функції електронів у відповідних станах. Тому більш широкими є зони, утворені з вище розташованих рівнів. Цілком закономірно, що дане наближення виконується тим точніше, чим сильніше електрон зв'язаний з атомом, тому цей метод дає задовільні результати для внутрішніх електронних оболонок.

Метод слабого зв'язку використовується для аналізу енергетичного спектру валентних електронів, які значною мірою колективізовані і слабо зв'язані зі своїми атомами. Потенціальна функція $U(\vec{r})$ у цьому випадку представляється у вигляді суми постійного середнього значення U_{cp} і деякої функції $U_-(\vec{r})$, що характеризує періодичність поля кристалічної ґратки (рис. 1.3), а рівняння Шредінгера набуває вигляду диференціального рівняння з періодичними коефіцієнтами. У теорії таких рівнянь доводиться, що рішенням є функція, яку у фізиці твердого тіла називають *функцією*, або *хвилею Блоха*

$$\psi(\vec{r}) = \varphi(\vec{r}) \exp(i\vec{k}\vec{r}) \quad (1.2)$$

і яка є плоскою хвилею з амплітудою $\varphi(\vec{r})$ у вигляді періодичної функції з тим же періодом, що і у кристалічної ґратки. Але рішення такого виду можливе не за всіх значень параметра рівняння E , а тільки у деяких його інтервалах, які відповідають дозволеним зонам. У проміжку між цими інтервалами розв'язок рівняння Шредінгера також можливий, але він відповідає уявним значенням хвильового вектора k і представляється через функції $\exp(\pm i\vec{k}\vec{r})$. Такі функції не задовольняють умові нормування хвильових функцій через необмежене зростання і не можуть бути придатними для опису *стаціонарного* стану електрона у нескінченно довгому кристалі. Тому енергетичні інтервали, що відповідають цим типам розв'язків, називають *забороненими*.

Типовим прикладом такого рівняння, яке часто розглядається в математиці, є так зване рівняння Мат'є, до якого зводиться рівняння (1.1) для потенціальної функції $U_-(\vec{r})$ у вигляді звичайної

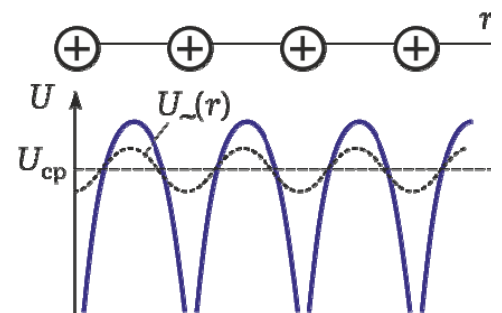


Рис. 1.3. Періодичність поля кристалічної ґратки у перетинах, що проходять через площину розташування іонів (суцільна лінія) і між ними (штрихова лінія)

синусоїдальної. Іншим прикладом, поширеним у навчальній літературі, є модель Кроніга-Пенні (див. додаток до гл. 1), в якій функція $U(\vec{r})$ приймається у вигляді періодичної послідовності прямокутних потенціальних бар'єрів.

Таким чином, використовуючи наближення як сильного, так і слабого зв'язку, ми приходимо до уявлення про зонний характер енергетичного спектру електронів у кристалі: дозвалені значення енергії чергуються зі значеннями, яких електрон мати не може.

1.1.2. Загальні властивості хвильової функції електронів у кристалі

Головний висновок зонної теорії про зонний характер дозволених значень енергії є досить плідним для пояснення багатьох фізичних процесів. Але такий «одновимірний» висновок недостатній для опису деяких більш тонких явищ, коли треба застосовувати знання форми закону дисперсії, тобто залежності повної енергії електрона від його хвильового вектора. Цей «двовимірний» результат у принципі впливає і з розглянутих методів сильного і слабого зв'язку, але через прийняті наближення він має обмежене значення.

Досить потужні висновки, що мають важливе значення, слідують з найзагальніших властивостей кристала. Якщо розглядати ідеальний нескінченний кристал, то однією з найбільш важливих його властивостей є *періодичність* кристалічної ґратки, або так звана *трансляційна симетрія*, яка означає, що зміщення (трансляція) ґратки на ціле число періодів у той чи інший бік нічого не змінює. Ця властивість потенціальної енергії електрона означає, що

$$U(\vec{r}) = U(\vec{r} + \vec{a}), \quad (1.3)$$

де $\vec{a} = n_x \vec{a}_x + n_y \vec{a}_y + n_z \vec{a}_z$ називається вектором трансляції, що складається з будь-якого цілого числа (n_x, n_y, n_z) базисних векторів $\vec{a}_x, \vec{a}_y, \vec{a}_z$ цієї кристалічної ґратки (рис. 1.4). Це означає, що будь-яка точка з координатою \vec{r} і трансляційно симетрична з нею точка $\vec{r} + \vec{a}$ є фізично еквівалентними і нерозрізненими для електрона. У свою чергу, це означає, що ймовірність знаходження електрона в околі цих точок однакова, тобто $|\psi(\vec{r})|^2 = |\psi(\vec{r} + \vec{a})|^2$. Рівність модулів не означає рівності хвильових функцій, вони можуть відрізнитися

деяким фазовим множником з фазою, яку природньо вважати пропорційною вектору трансляції

$$\psi(\vec{r} + \vec{a}) = \psi(\vec{r}) \exp(i\vec{k}\vec{a}) \quad (1.4)$$

Хвильову функцію електрона можна у загальному вигляді представити як добуток деякої амплітуди $\varphi(\vec{r})$, що просторово змінюється, на хвильовий множник

$$\psi(\vec{r}) = \varphi(\vec{r}) \exp(i\vec{k}\vec{r}) \quad (1.5)$$

Для з'ясування властивостей $\varphi(\vec{r})$ у кристалі змістимося на деякий вектор \vec{a} . Тоді відповідно до (1.4) вираз (1.5) набере вигляду $\psi(\vec{r} + \vec{a}) = \varphi(\vec{r}) e^{i\vec{k}\vec{r}} e^{i\vec{k}\vec{a}}$. Але з іншого боку (1.5) залишається справедливим і для точки $\vec{r} + \vec{a}$, тому $\psi(\vec{r} + \vec{a}) = \varphi(\vec{r} + \vec{a}) e^{i\vec{k}(\vec{r} + \vec{a})}$. Порівнюючи останні вирази, бачимо, що

$$\varphi(\vec{r}) = \varphi(\vec{r} + \vec{a}), \quad (1.6)$$

тобто амплітуда хвильової функції є *періодичною*. Цей же результат можна отримати, використовуючи метод слабкого зв'язку, але тут він представляється універсальним, впливаючим з найзагальніших властивостей кристалічної ґратки. Хвильову функцію електрона у такому вигляді називають *хвилею Блоха*, її умовне зображення наведено на рис. 1.5.

Важливість отриманого результату полягає в тому, що просторова густина ймовірності $|\psi(\vec{r})|^2 = |\varphi(\vec{r})|^2$ є періодичною функцією, а усереднена по періоду ґратки - *постійною уздовж усього кристала*. Це означає, що в ідеальній кристалічній ґратці електронна хвиля поширюється без розсіювання, тому ідеальний кристал повинен мати нульовий опір. Причиною можливого розсіювання електронних хвиль є будь-які *дефекти* кристалічної ґратки, які викликають порушення трансляційної симетрії.

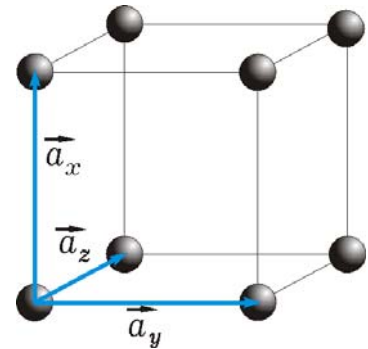


Рис. 1.4. Базисні вектори кристалічної ґратки мають довжину, рівну відповідному періоду ґратки, і напрямки, що збігаються з ребрами елементарної комірки

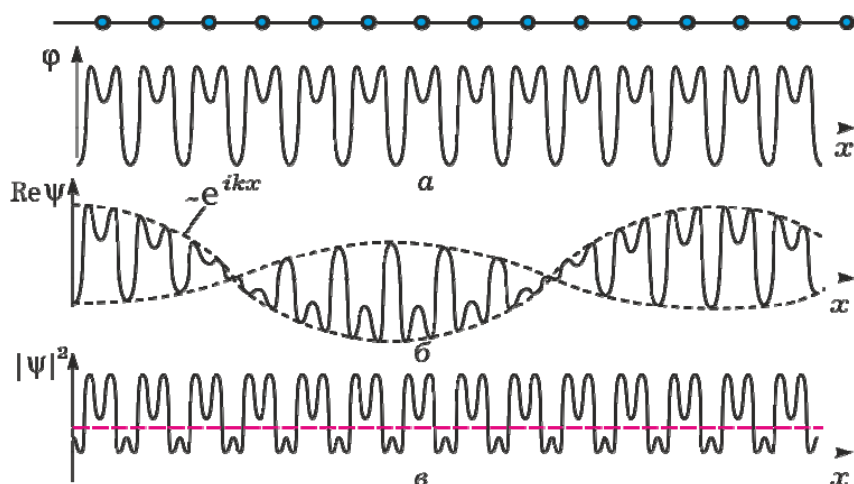


Рис. 1.5. Періодичність амплітуди хвилі Блоха відповідає періодичності кристалічної ґратки (а). Умовний вигляд хвилі Блоха (б) і квадрат її амплітуди (в) показують, що середня густина ймовірності знаходження електрона уздовж ідеального кристала постійна

Особливості поведінки електронів у кристалі не можна зрозуміти без з'ясування деяких характерних властивостей хвильового вектору, які витікають також з властивостей симетрії кристала. Для спрощення розглядатимемо одновимірний кристал, тобто лінійний нескінченно довгий ланцюжок атомів, розташованих уздовж осі x з періодом a_x , і отримані результати будемо узагальнювати для інших напрямів.

Перша властивість хвильового вектору полягає в тому, що неможливо вказати його значення однозначно, а тільки з точністю до деякого доданку. Щоб це зрозуміти, поставимо питання, чи може електрон у станах з різними значеннями k' і k'' описуватися однаковою хвильовою функцією. Тобто припустимо, що $\psi_{k'}(x) = \psi_{k''}(x)$. Розглянемо цю рівність і в точці $x + a$: $\psi_{k'}(x + a) = \psi_{k''}(x + a)$, де a у даному випадку - одновимірний вектор трансляції. Використовуючи властивість (1.4), останній вираз подамо у вигляді: $\psi_{k'}(x)e^{ik'a} = \psi_{k''}(x)e^{ik''a}$, де показники експонент мають бути рівними. Звідси отримуємо, що наше припущення можливе не лише для тотожного випадку $k' = k''$, але і для більш загального, коли $k'a = k''a + 2\pi n$, де n - будь-яке ціле число. Отже, якщо

$$k' = k'' + n(2\pi/a), \quad (1.7)$$

то хвильові функції $\psi_{k'}$ і $\psi_{k''}$ електрона однакові, а його стани з такими значеннями k нерозрізненні.

Мінімальна відстань уздовж осі k_x між фізично еквівалентними станами (рис. 1.6) дорівнює $2\pi/a_x$. Область в k -просторі, яка включає нееквівалентні значення хвильового вектора, називають зоною Бріллюена. Прийнято зони Бріллюена розташовувати симетрично відносно початку координат. На рис. 1.7 зображено перші три зони для лінійного ланцюжка атомів (а) і двовимірної прямокутної ґратки (б), а також перші зони для простої кубічної (в) і гранецентрованої кубічної ґратки (г).

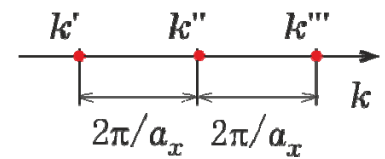


Рис. 1.6. Стани електронів з хвильовими векторами k' , k'' і k''' , які відрізняються на $2\pi/a$, описуються однаковими хвильовими функціями і є нерозрізненими

Важливість уявлення про зону Бріллюена полягає у тому, що розгляд усіх можливих переміщень електрона по станах достатньо проводити тільки в одній зоні Бріллюена (як правило в першій), оскільки інші зони містять стани, що повторюються. З відміченою особливістю хвильового вектору пов'язано також те, що величина $\vec{p} = \hbar\vec{k}$ не має звичайних властивостей імпульсу і називається квазіімпульсом, а \vec{k} - квазіхвильовим вектором.

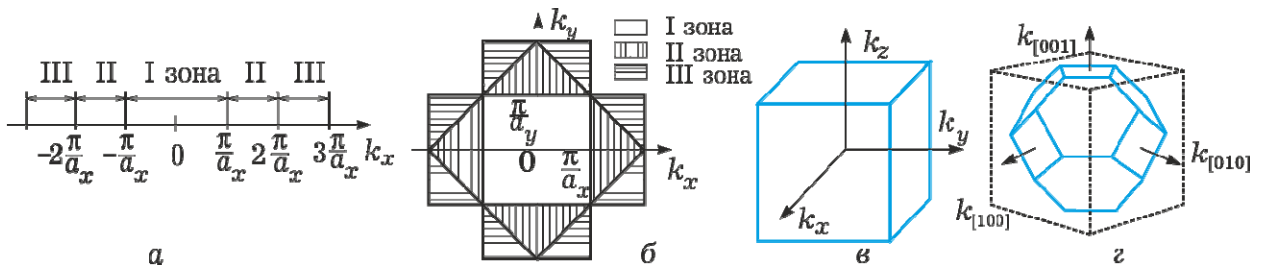


Рис. 1.7. Зображення перших трьох зон Бріллюена для лінійного ланцюжка атомів (а), двовимірної квадратної ґратки (б) і перших зон для простої кубічної (в) і гранецентрованої кубічної ґратки типу алмаза (г)

Друга властивість хвильового вектору, дискретність його можливих значень, проявляється для кристалів обмежених розмірів. Для підрахунку числа цих значень використаємо наступний прийом. Розглянемо кристал (одновимірний) завдовжки L_x , уздовж якого укладається $N_x = L_x/a_x$ атомів. Послідовними переміщеннями кристала покриємо кристалічною ґраткою увесь простір. Це дає можливість використовувати для опису електронних станів блоківські

функції (1.5). Але не усі ці функції, які придатні для нескінченного кристала, можуть нас задовольнити. З них треба відібрати тільки ті, для яких виконується умова $\psi(x) = \psi(x + L_x)$, оскільки точки x і $x + L_x$ є не просто фізично еквівалентними з однаковою густиною ймовірності, але і *тотожними*, з однаковими значеннями хвильових функцій. Використовуючи форму (1.5) представлень блохівських функцій, перепишемо цю умову у вигляді $\varphi(x)e^{ik_x x} = \varphi(x + L_x)e^{ik_x(x+L_x)}$. Враховуючи властивість (1.6) періодичності амплітуди $\varphi(x)$, отримуємо, що хвильові функції для кристала завдовжки L_x мають бути такими, що $e^{ik_x L_x} = 1$, звідки

$$k_x = n_x(2\pi/L_x) = n_x(2\pi/N_x a_x), \quad n_x = 0, \pm 1, \pm 2, \dots \quad (1.8, a)$$

Цей результат показує, що для кристала скінченних розмірів можливі тільки такі електронні стани, які відповідають дискретному набору значень k_x (рис. 1.8), можуть бути пронумеровані і кількість яких, взагалі кажучи, нескінченна. Але якщо обмежитися лише значеннями, що не повторюються, тобто знаходяться тільки в першій зоні Бріллюена ($-\pi/a_x \leq k_x \leq \pi/a_x$), то нумерацію їх потрібно проводити в межах $-N_x/2 < n_x < N_x/2$, а число цих станів дорівнює числу атомів N_x у даному кристалі.

Як *третю властивість* слід зазначити, що ці дозволені стани розподілені рівномірно уздовж осі k_x з густиною $\Delta n_x / \Delta k_x = N_x a_x / 2\pi$. У реальних ситуаціях кількість атомів у кристалі дуже велика, тому стани вважають розташованими *квазінеперервно*.

У тривимірному кристалі отримані результати природно узагальнити і для інших напрямів k -простору:

$$k_y = n_y(2\pi/L_y) = n_y(2\pi/N_y a_y), \quad n_y = 0, \pm 1, \pm 2, \dots, \pm N_y/2, \quad (1.8, б)$$

$$k_z = n_z(2\pi/L_z) = n_z(2\pi/N_z a_z), \quad n_z = 0, \pm 1, \pm 2, \dots, \pm N_z/2. \quad (1.8, в)$$

Тут, як і раніше, число станів у першій зоні Бріллюена уздовж кожної з осей k -простору дорівнює числу атомів уздовж відповідних напрямів у кристалі. Оскільки ці стани нумеруються незалежно в усіх напрямках, то повне число станів дорівнює добутку $N_x N_y N_z$, тобто повному числу атомів N у кристалі. Усі стани розміщені рівномірно у k -просторі (рис. 1.9) з об'ємною густиною:

$$g(k) = \frac{\Delta n}{\Delta V_k} = \frac{\Delta n_x \Delta n_y \Delta n_z}{\Delta k_x \Delta k_y \Delta k_z} = \frac{L_x L_y L_z}{8\pi^3} = \frac{V}{8\pi^3}. \quad (1.9)$$

Тут $V = L_x L_y L_z$ - геометричний об'єм кристала. Якщо прийняти $V = 1$, то отримуємо нормовану величину густини станів

$$g(k) = 1/(2\pi)^3. \quad (1.10, a)$$

Сучасна технологія дозволяє виготовляти так звані структури низької розмірності: двовимірні (2D) і одновимірні (1D). Вони є тонкими (порядку одиниць або десятків міжатомних відстаней) шарами або нитками з напівпровідників, які обмежені потенціальними бар'єрами, утвореними більше широкозонними матеріалами, які їх оточують. Зазвичай ці бар'єри досить високі, щоб локалізувати електрони в поперечному перерізі цих структур.

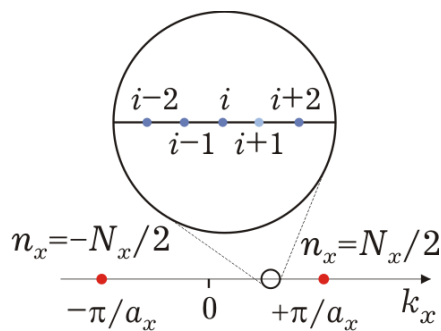


Рис. 1.8. У кристалі скінченних розмірів k -стани розташовані рівномірно, а їх кількість у першій зоні Брілюєна дорівнює кількості атомів у кристалі

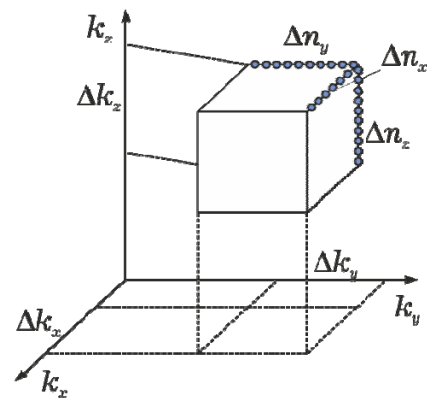


Рис. 1.9. Елементарний об'єм у k -просторі для підрахунку густини станів

Для плоскої кристалічної ґратки, періодичної у напрямках x і y та обмеженої у напрямі z , розв'язок рівняння Шредінґера дає спектр k_z станів з дискретним розподілом k_i , у деякій мірі подібним до розподілу в окремому атомі. У напрямках x і y відбувається усупільнення станів, спектр яких має зонний характер з блохівськими хвильовими функціями. Тому кожному k_{iz} -стану властива двовимірна густина квазінеперервних станів, яка відповідно до (1.8, а і б) має значення $g_i(k_x, k_y) = L_x L_y / (2\pi)^2$, або для кристала одиничної площі

$$g_i(k_x, k_y) = 1/(2\pi)^2. \quad (1.10, б)$$

Для одновимірного кристала, обмеженого в напрямках y і z , зберігається дискретність станів k_{ij} упродовж k_y і k_z , але в поздовжньому напрямі x спостерігається розщеплювання станів у квазінеперервний спектр. Для кожного k_{ij} -стану придатне співвідношення (1.10, а), згідно з яким для кристала одиничної довжини

$$g_{ij}(k_x) = 1/(2\pi). \quad (1.10, в)$$

З останніх виразів видно, що у k -просторі електронні стани розподілені рівномірно, з однаковою густиною. Відзначимо, що ці значення густини треба подвоїти, враховуючи можливість заповнення станів двома електронами з протилежними значеннями спінових моментів.

1.1.3. Закон дисперсії електрона у кристалі

Явище дисперсії характерне практично для будь-яких хвильових процесів і проявляється у залежності фазової швидкості поширення хвилі від її частоти. Оскільки $v_\phi = \omega/k$, то закон дисперсії можна представити у вигляді залежності $k(\omega)$. У фізиці твердого тіла прийнято закон дисперсії виражати через зворотню функцію $\omega(k)$ або (з урахуванням співвідношення де Бройля $E = \hbar\omega$) як залежність $E(k)$. Важливість знання закону дисперсії полягає в тому, що він дозволяє описувати основні особливості електронних хвиль у кристалі: фазову швидкість $v_\phi = (E/k)/\hbar$; групову швидкість $v_{gr} = (dE/dk)/\hbar$; ефективну масу електрона $m^* = \hbar^2 / (d^2E/dk^2)$ (див. наступний розділ) і його енергетичний спектр, а також більш детальні особливості поведінки електрона за різних значень енергії.

Знаходження законів дисперсії або дисперсійних кривих для конкретних речовин є складною задачею і проводиться спільно обчислювальними і експериментальними методами. Ми з'ясуємо лише основні властивості дисперсійних кривих, що витікають із вже відомих властивостей хвильового вектору.

1. Дисперсійна крива *періодична*, тобто $E(k) = E(k + n2\pi/a)$. Ця властивість виходить з того, що стани k і $k + n2\pi/a$ у кристалі нерозрізненні і їм відповідають однакові хвильові функції, тобто і

однакові значення енергії. Період дисперсійної кривої дорівнює одній зоні Бріллюена.

2. Дисперсійна крива *парна функція*, тобто $E(k) = E(-k)$. Це легко зрозуміти, оскільки у будь-якому кристалі рух електронів «вперед» і «назад» абсолютно рівноправний і станам $+k$ і $-k$ відповідає однакова енергія.

3. Для кристала скінченних розмірів дисперсійна крива - *дискретна*, або точніше - визначена на точковій множині. Ця властивість безпосередньо виходить з дискретності хвильового вектору для обмеженого кристала.

Усі перераховані властивості умовно зображені на рис. 1.10. Дисперсійна крива дозволяє знайти розподіл енергетичних станів електронів у зоні, виходячи з густини k -станів (1.10). Це питання буде розглянуто далі, а поки обмежимося наближеними оцінками.

Оскільки число атомів дорівнює числу дозволених k -станів у кристалі, то і число енергетичних станів у дозволених зоні має такий же порядок. У кристалі об'ємом $\sim 1 \text{ мм}^3$ знаходиться приблизно 10^{20} атомів, тому в енергетичній зоні шириною близько 1 еВ енергетичні стани розташовані на відстані порядку 10^{-20} еВ. Такі малі енергетичні інтервали дозволяють вважати енергетичний спектр електрона у межах дозволеної зони *квазінеперервним*.

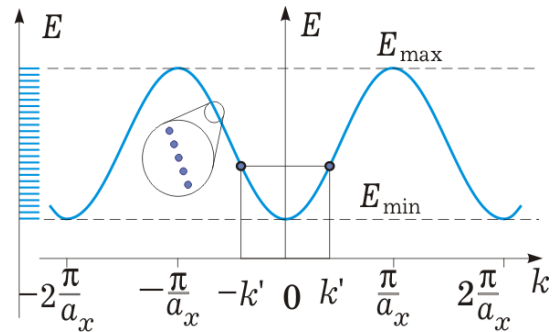


Рис. 1.10. Умовне зображення основних особливостей дисперсійної кривої: періодичності, парності та дискретності

Наочне зображення закону дисперсії для одновимірного випадку представляють графічно, як на рис. 1.10. Проте, якщо зображувати його для декількох дозволених енергетичних зон (рис. 1.11, а), то отримаємо так звану схему «зон, що повторюються», яка іноді незручна своєю багатозначністю. Уникнути цієї багатозначності можна використанням схеми «розширених зон» (рис. 1.11, б), в якій дисперсійна крива однозначно задана на всій осі k так, що кожній енергетичній зоні відповідає тільки одна зона Бріллюена. Але найчастіше дисперсійні криві зображують тільки в межах першої зони Бріллюена (рис. 1.11, в), тобто за допомогою схеми «зведених зон». На тому ж малюнку справа (г) наведена звичайна «одновимірна» енергетична діаграма, яка показує тільки

енергетичний спектр електрона у вигляді дозволених (заштриховані) і заборонених зон, що чергуються.

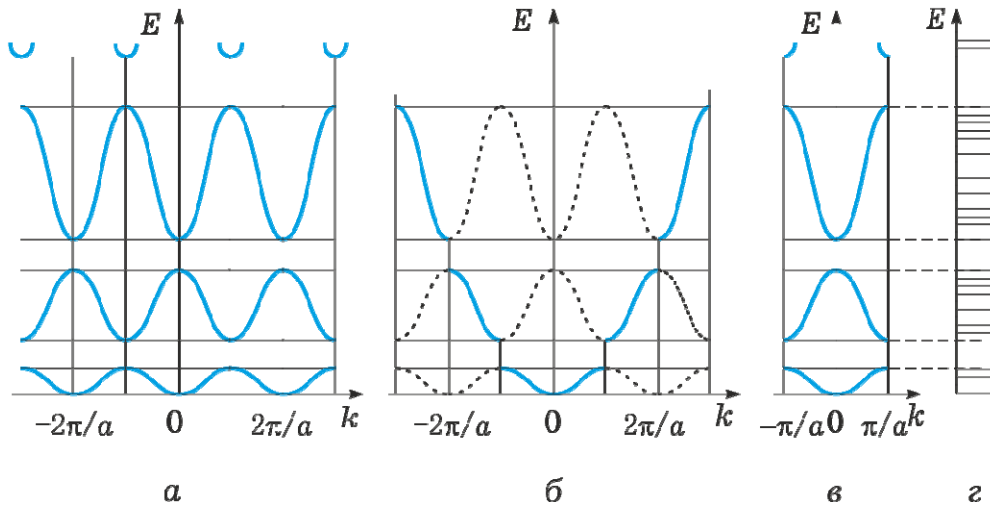


Рис. 1.11. Зображення закону дисперсії для одновимірного кристала у вигляді схем повторюваних (а), розширених (б) і зведених (в) зон

Для зображення дисперсійної кривої у вигляді функції складових хвильового вектору (наприклад, k_x і k_y) доводиться використовувати косокутні проєкції (рис. 1.12) або сукупність ліній однакових енергій (ізоенергетичних), що є слідами перетину поверхні $E(k_x, k_y)$ площинами постійної енергії.

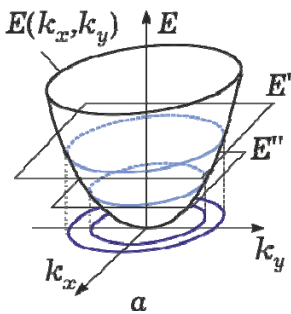


Рис. 1.12. Спосіб зображення закону дисперсії у двовимірному k -просторі за допомогою ліній рівної енергії

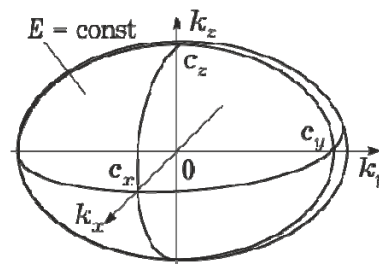


Рис. 1.13. Поверхні рівної енергії у тривимірному k -просторі поблизу екстремумів закону дисперсії мають вигляд еліпсоїдів

Функція $E(k_x, k_y, k_z)$ є чотиривимірною поверхнею і не може бути зображена на плоскому малюнку. Для цього в довідниках зазвичай використовують набір одновимірних графіків $E(k)$ для найбільш характерних напрямів зони Бріллюена. Ще одна можливість криється в зображенні $E(k_x, k_y, k_z)$ за допомогою

ізоенергетичних поверхонь, тобто поверхонь постійної енергії $E = \text{const}$ у просторі змінних k_x , k_y і k_z (рис. 1.13).

Питання про форму ізоенергетичних поверхонь у конкретних речовинах є дуже складним. Але оскільки основні електричні властивості кристала пов'язані з електронними станами *поблизу краю енергетичних зон*, деякі загальні висновки можна досить просто отримати, розкладаючи функцію $E(k_x, k_y, k_z)$ в степеневий ряд в околі екстремуму (верхнього або нижнього краю) енергетичної зони. Для спрощення вважатимемо, що цей екстремум лежить у центрі зони Бріллюена ($k = 0$), тобто $E_m = E(0)$. Оскільки закон дисперсії є парною функцією, то в розкладанні будуть відсутні непарні степені. Обмеживши ряд другими степенями аргументів, отримуємо:

$$E(k_x, k_y, k_z) = E_m + \frac{1}{2} \left(\left. \frac{\partial^2 E}{\partial k_x^2} \right|_{k_x=0} k_x^2 + \left. \frac{\partial^2 E}{\partial k_y^2} \right|_{k_y=0} k_y^2 + \left. \frac{\partial^2 E}{\partial k_z^2} \right|_{k_z=0} k_z^2 \right). \quad (1.11)$$

У цей вираз не включені змішані похідні типу $\partial^2 E / \partial k_i \partial k_j$, оскільки їх завжди можна зробити рівними нулю відповідним вибором системи координат. Доцільно (як далі буде видно) також позначити:

$$\frac{\partial^2 E}{\partial k_i^2} = \frac{\hbar^2}{m_i^*}, \quad i = x, y, z \quad (1.12)$$

після чого закон дисперсії набуває вигляду:

$$E(\vec{k}) = E_m + \frac{\hbar^2}{2m_x^*} k_x^2 + \frac{\hbar^2}{2m_y^*} k_y^2 + \frac{\hbar^2}{2m_z^*} k_z^2. \quad (1.13)$$

Рівняння ізоенергетичної поверхні для фіксованого значення E' безпосередньо виходить з (1.13). Представивши його в канонічній формі:

$$\frac{k_x^2}{c_x^2} + \frac{k_y^2}{c_y^2} + \frac{k_z^2}{c_z^2} = 1, \quad (1.14)$$

де $c_i = \sqrt{2m_i^*(E' - E_m) / \hbar^2}$, ($i = x, y, z$), отримуємо, що у загальному випадку ізоенергетична поверхня є еліпсоїдом з півосями c_x , c_y і c_z (рис. 1.13). Це твердження, по-перше, є справедливим також для екстремумів, розташованих не лише в центрі зони Бріллюена, але обов'язково для абсолютних. По-друге, воно зберігає свою силу не

лише для нижнього краю енергетичної зони, де $E' > E_{\min}$ і $\partial^2 E / \partial k_i^2 > 0$. Біля верхнього краю $E' < E_{\max}$, але оскільки другі похідні в околі максимуму від'ємні, то величини c_i і в цьому випадку залишаються дійсними. Для локальних екстремумів дисперсійних кривих, які можуть мати місце усередині енергетичних зон, в деяких напрямках значення c_i^2 можуть бути від'ємними, а ізоенергетичні поверхні стають іншими поверхнями другого порядку.

На закінчення відмітимо, що типовими формами ізоенергетичних поверхонь для багатьох напівпровідників є еліпсоїд обертання (з двома однаковими півосями) і сфера. Форму сфери мають ізоенергетичні поверхні і для вільного електрона. Оскільки вільній частці властива тільки кінетична енергія, то

$$E = \frac{mv^2}{2} = \frac{p^2}{2m} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2). \quad (1.15)$$

Для фіксованого значення $E = E'$ цей вираз є в k -просторі рівнянням сфери з радіусом $\sqrt{2mE' / \hbar^2}$. З порівняння (1.13) і (1.15) видно, що введені нами величини m_i^* займають таке ж місце в законі дисперсії для електронів у кристалі, що і звичайна маса для вільного електрона.

Наведена аналогія дозволяє за допомогою так званого *методу ефективної маси* описати поведінку електронів поблизу краю енергетичних зон, застосовуючи модель вільних електронів. Природно, що цей метод зручний у межах квадратичності закону дисперсії, хоча іноді його використовують, коли в розкладанні (1.11) в ряд враховують четверті степені.

1.1.4. Ефективна маса електрона

На електрон у кристалі діє сила періодичного поля ґратки і (за її наявності) зовнішня сила, яка може бути викликана, наприклад, електричним полем, створеним деяким стороннім джерелом. Аналіз поведінки електрона у разі спільної дії таких сил є дуже складною задачею для квантової механіки. Але є дуже простий метод, який дозволяє істотно її спростити. Для його розуміння спочатку встановимо деякі формальні співвідношення.

Зонна теорія визначає для періодичного кристалу *стаціонарні* електронні \vec{k}_i -стани і відповідно до закону дисперсії – енергетичні стани $E(\vec{k}_i)$. Електрони в них залишаються як завгодно довго, якщо немає ніякої зовнішньої дії. Зміна цих станів можлива тільки під дією деякої зовнішньої сили \vec{F} , наприклад, електричного поля. Розглядаючи для спрощення одновимірний випадок, маємо

$$F = \frac{dp}{dt} = \hbar \frac{dk}{dt}. \quad (1.16)$$

Це співвідношення нагадує звичайний закон Ньютона, але сюди входить тільки зовнішня сила (періодична сила ґратки враховується зонною теорією у визначенні k -станів) і *квазіімпульс*, а не імпульс.

З іншого боку швидкість електрона відповідає груповій швидкості електронної хвилі де Бройля і є рівною

$$v = d\omega / dk = (dE / dk) / \hbar, \quad (1.17)$$

а прискорення електрона, враховуючи неявну залежність енергії від часу через хвильовий вектор $k(t)$, можна представити у такому вигляді:

$$a = \frac{dv}{dt} = \frac{1}{\hbar} \cdot \frac{d}{dt} \left(\frac{dE}{dk} \right) = \frac{1}{\hbar} \cdot \frac{dk}{dt} \cdot \frac{d}{dk} \left(\frac{dE}{dk} \right) = \frac{1}{\hbar} \cdot \frac{d^2E}{dk^2} \cdot \frac{dk}{dt}. \quad (1.18)$$

Порівняння виразів (1.16) і (1.18) виявляє пропорційність між зовнішньою силою, що діє на електрон, і його прискоренням. Коефіцієнт пропорційності має розмірність маси і дорівнює раніше введений (1.12) нами величині m^* :

$$m^* = \frac{\hbar^2}{d^2E / dk^2}. \quad (1.19)$$

Ця величина і називається *ефективною масою електрона*. Таким чином, суть прийому, званого *методом ефективної маси*, полягає в тому, що, приписуючи електрону деяку фіктивну величину m^* , ми дістаємо можливість абстрагуватися від дії періодичного поля кристалічної ґратки і розглядати його рух як рух вільної частки з масою m^* під дією *тільки сторонніх* сил. Закономірно, що у цьому випадку задача спрощується, оскільки

ефективну масу вважаємо відомою. Але це можливо лише за відомого закону дисперсії для цього кристала.

У більш загальному тривимірному випадку справедливе векторне співвідношення:

$$\vec{F} = \mathbf{m}^* \cdot \vec{a} = \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{pmatrix} m_x^* & 0 & 0 \\ 0 & m_y^* & 0 \\ 0 & 0 & m_z^* \end{pmatrix} \cdot \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}, \quad (1.20)$$

де m^* - тензорна величина, що включає тільки діагональні компоненти

$$m_i^* = \frac{\hbar^2}{(\partial^2 E / \partial k_i^2)}, \quad (1.21)$$

Позадіагональні компоненти дорівнюватимуть нулю, якщо осі k_x , k_y і k_z співпадатимуть з головними осями еліпсоїда, як показано на рис. 1.13.

У тривимірному випадку групову швидкість електрона знаходять не з (1.17), а через градієнт функції $E(\vec{k})$:

$$\vec{v} = \frac{1}{\hbar} \text{grad}_{\vec{k}} E(\vec{k}) = \frac{1}{\hbar} \left(\vec{e}_x \frac{\partial E}{\partial k_x} + \vec{e}_y \frac{\partial E}{\partial k_y} + \vec{e}_z \frac{\partial E}{\partial k_z} \right). \quad (1.22)$$

Для електронних станів поблизу краю зони, де має місце квадратичний закон дисперсії (1.13):

$$\vec{v} = \hbar \left(\vec{e}_x \frac{k_x}{m_x^*} + \vec{e}_y \frac{k_y}{m_y^*} + \vec{e}_z \frac{k_z}{m_z^*} \right). \quad (1.23)$$

Останні співвідношення показують, що в кристалі вектор швидкості спрямований перпендикулярно до ізоенергетичної поверхні $E(\vec{k}) = \text{const}$ у бік зростання енергії. Якщо ізоенергетичні поверхні є сферами (як і для вільних електронів), що можливо за рівних ефективних мас, то $v(\vec{k})$ і \vec{k} спрямовані в один бік.

Через тензорний характер ефективної маси вектори прискорення і зовнішньої сили можуть не співпадати за напрямом. Більше того, вони можуть бути протилежні один одному, коли

ефективна маса від'ємна. Це має місце, наприклад, біля верхнього краю дозволеної енергетичної зони (рис. 1.14, а), де через опуклість дисперсійної кривої $d^2E/dk^2 < 0$. У цьому немає нічого дивного, якщо пам'ятати, що електрон знаходиться в полі зовнішніх сил і в полі кристалічної ґратки, але результат їх дії, тобто результуюче прискорення електрона, ми приписуємо тільки зовнішній силі. Ця ситуація ($m^* < 0$) реалізується, коли уся робота зовнішніх сил і крім того частина внутрішньої кінетичної енергії електрона витрачається на збільшення потенціальної енергії електрона. Він у цьому випадку сповільнюється і зовнішньою силою сприймається як частка з від'ємною масою.

Якщо частина роботи, що здійснюється зовнішньою силою, збільшує кінетичну енергію електрона, то напрям сили співпадає з напрямом прискорення і $m^* > 0$. Таке положення характерне для електронних станів біля нижнього краю енергетичної зони, де $d^2E/dk^2 > 0$ (рис. 1.14, б). У цьому випадку частина роботи зовнішніх сил збільшує потенціальну енергію електрона, і його прискорення буде меншим, ніж у вільного електрона, тому $m^* > m$.

Якщо ж кінетична енергія електрона збільшується не лише за рахунок роботи зовнішніх сил, але і за рахунок потенціальної енергії його взаємодії з ґраткою, то швидкість зростає так, ніби він легший за звичайний, тобто $m^* < m$.

Треба пам'ятати, що застосування поняття ефективної маси зовсім не означає зміни гравітаційної маси електрона, а лише відображає факт зміни його динамічних властивостей, тобто чутливості до зовнішніх сил за рахунок впливу періодичного поля кристалічної ґратки.

Важливість наведеного розгляду полягає в тому, що ми отримуємо так звані *напівкласичні рівняння руху* (1.20) і (1.22). Вони за формою нагадують звичайні рівняння класичної механіки, але враховують квантові особливості поведінки електронів у періодичному полі ґрат, і на сьогодні є основним засобом для

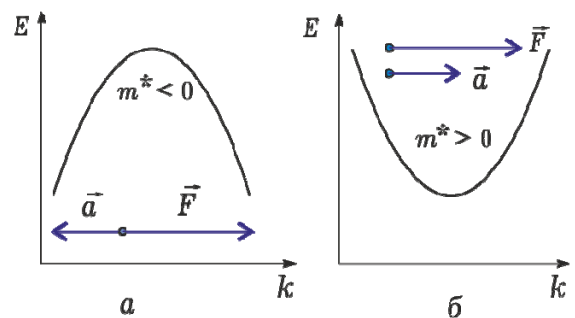


Рис. 1.14. Ефективна маса електрона біля стелі зони від'ємна і його прискорення направлено протилежно силі (а), біля дна зони додатна (б), а сила і прискорення збігаються

проведення аналізу поведінки електронів у кристалі під дією зовнішніх сил.

1.1.5. Розподіл енергетичних станів

Розглянемо питання про розподіл електронних станів, який дуже важливий для кількісного обчислення основних електричних параметрів речовини. Вище ми вже з'ясували, що у k -просторі електронні стани розподілені рівномірно з густиною $g(k) = 1/(2\pi)$ для кожного напрямку (формули 1.10). Оскільки ми не звертали увагу на спіні, то цей результат треба подвоїти, оскільки кожен стан може бути заповнений двома електронами з протилежними напрямками спінових моментів

$$g(k) = 2/(2\pi)^n, \quad (1.24)$$

де n відповідає «розмірності» кристалу: $n = 1$ для одновимірних, $n = 2$ для двовимірних і $n = 3$ для звичайних об'ємних кристалів. Реальними об'єктами U наноелектроніці є так звані квантові нитки (1D) або квантові шари (2D).

У багатьох випадках інтерес викликає розподіл не k -станів, а розподіл енергетичних станів, що характеризується густиною

$$g(E) = dn_c/dE, \quad (1.25)$$

яка має сенс кількості електронних станів в одиничному енергетичному інтервалі. Перехід до функції $g(E)$ від $g(\vec{k})$ можна зробити, використовуючи відповідний закон дисперсії $E(\vec{k})$.

Проста модель *одновимірних* (1D) кристалів є лінійним ланцюжком атомів (рис. 1.15, а), що має властивість трансляційної симетрії тільки в одному x -напрямі.

Локалізація електронів в y - і z -напрямах обумовлює квантування поперечних станів k_{iy} і k_{jz} . Їм відповідає дискретний енергетичний спектр $E_{ij}(k_{iy}, k_{jz})$, у деякій мірі подібний до спектру електронів в окремому атомі. Але взаємодія валентних електронів в поздовжньому напрямі викликає їх усупільнення і зонний спектр енергій їх руху уздовж кристала. Відповідно хвильова функція має вигляд хвилі Блоха, що біжить вздовж осі x : $\psi(\vec{r}) \sim \exp(ik_x x)$. Квадратичний закон дисперсії у таких «одновимірних» зонах (іноді

їх називають підзонами) у наближенні ефективної маси можна представити у вигляді квадратичних функцій зі своїм початком відліку E_{ij} (рис. 1.15, б)

$$E(k_x) = E_{ij}(k_{iy}, k_{jz}) + \hbar^2 k_x^2 / 2m_x^* . \quad (1.26)$$

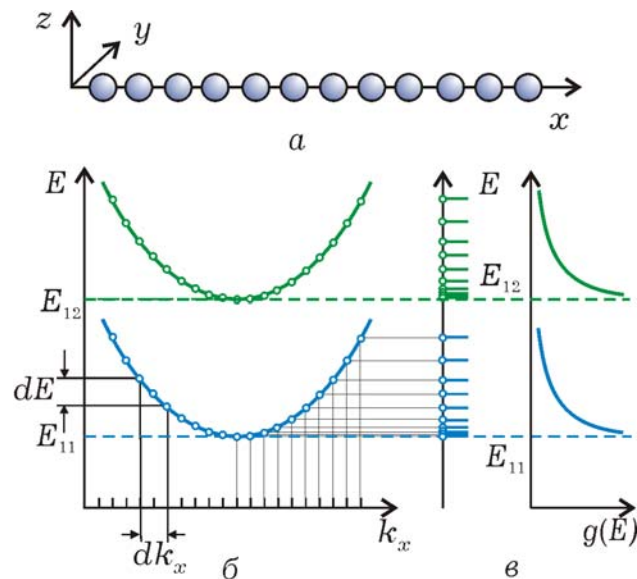


Рис. 1.15. В одновимірному ланцюжку атомів (а) закони дисперсії для поздовжнього руху у кожній підзоні є квадратичними (б) зі спадаючою густиною енергетичних станів (в)

Густину енергетичних станів можна визначити із співвідношення $g_{ij}(E)dE = g_{ij}(k_x)dk_x$, де густина k_x -станів задається значенням (1.24, в). Після обчислення похідної dk_x/dE і нескладних перетворень отримаємо

$$g_{ij}(E) = \sqrt{2m_x^*/\hbar^2} (E - E_{ij})^{-1/2}. \quad (1.27)$$

Як видно, густина енергетичних станів спадає з віддаленням від краю зони (рис. 1.15, в) і вони «локалізуються» поблизу рівня E_{ij} . Треба також пам'ятати, що густина одновимірних станів нормована на одиницю довжини і має розмірність $1/\text{м}\cdot\text{Дж}$.

Модель двовимірного (2D) кристала є атомарною структурою, періодичною у двох напрямках (x і y), і одноатомним шаром у z -вимірі (рис.1.16, б). Локалізація електрона у поперечному напрямі кристала викликає квантування k_{iz} -станів, яким відповідає дискретний енергетичний спектр E_i . Але уздовж кристала електронні стани усупільнюються, тому кожному $E_i(k_{iz})$ -стану відповідає своя підзона неперервних (для нескінченного в x - і y -

напрямах кристала) станів, дозволених для руху електронів у площині кристала. Для них хвильова функція має вигляд хвиль Блоха $\psi(\vec{r}) \sim \exp[i(k_x x + k_y y)]$, які поширюються у площині кристала. У наближенні ефективної маси біля краю кожної i -ої дозволеної підзони закон дисперсії можна вважати, подібно до (1.13), квадратичним, але двовимірним (рис.1.16, б):

$$E(k_x, k_y) = E_i(k_{iz}) + \hbar^2 k_x^2 / 2m_x^* + \hbar^2 k_y^2 / 2m_y^* \quad (1.28)$$

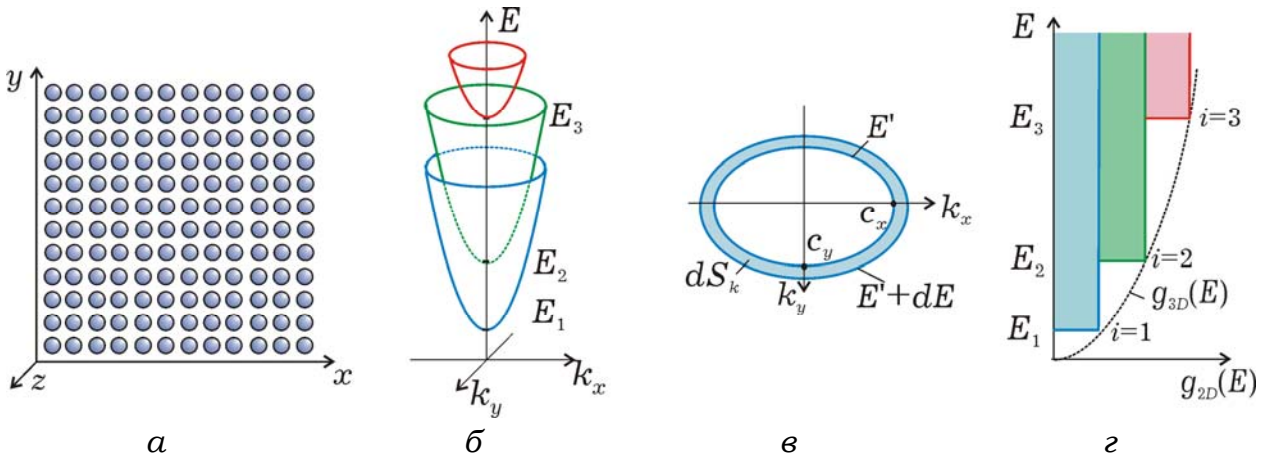


Рис 1.16. В одноатомному шарі (а) закони дисперсії у кожній підзоні є параболічними у напрямках k_x та k_y (б) з еліптичними ізоенергетичними лініями (в); у підзонах розподіл енергетичних станів є рівномірним

Відповідно до цього закону ізоенергетична лінія $E = \text{const}$ (рис. 1.16, в) має форму еліпса $(k_x/c_x)^2 + (k_y/c_y)^2 = 1$, де півосі c_x і c_y мають той же сенс, що і в (1.14), а його площа

$$S_k = \pi c_x c_y = 2\pi \sqrt{m_x^* m_y^*} (E - E_i) / \hbar^2.$$

Надаючи енергії приріст dE (рис. 1.16, в), отримуємо приріст площі еліпса $dS_k = 2\pi(\sqrt{m_x^* m_y^*} / \hbar^2) dE$. У цьому шарі знаходиться деяка кількість станів $dn_c = g_i(E) dE = g_i(k_x, k_y) dS_k$, звідки можна отримати двовимірну густину енергетичних станів у кожній i -зоні: $g_i(E) = g_i(k_x, k_y) dS_k / dE$. Обираючи значення $n = 2$ в (1.24) для густини k -станів, отримаємо

$$g_i(E) = 4\pi m_{di}^* / \hbar^2, \quad (1.29)$$

де $m_{di}^* = \sqrt{m_x^* m_y^*}$ – деяка середня величина, яку називають ефективною масою двовимірної густини станів. Звідси видно, що для двовимірних кристалів характерним є *рівномірний розподіл* енергетичних станів у зонах (рис. 1.16, з). Відзначимо також, що їхня густина є нормованою на одиницю площі і має розмірність $1/\text{м}^2 \cdot \text{Дж}$.

Для *тривимірного* кристалу функцію розподілу густини енергетичних станів $g(E)$ можна порівняно легко знайти для станів біля краю енергетичних зон, де закон дисперсії квадратичний (1.13), використовуючи *рівномірний розподіл* k -станів. Для цього розглянемо деякий об'єм k -простору dV_k між двома ізоенергетичними поверхнями, які відповідають значенням енергії E і $E + dE$ (на рис. 1.17 зображена тільки восьма частина цього об'єму). Число станів dn_c у цьому шарі, з одного боку, дорівнює $g(k)dV_k$, а з іншого – $g(E)dE$. Оскільки це ті ж самі стани, то

$$g(E) = g(k) dV_k/dE. \quad (1.30)$$

Залишається з'ясувати, як змінюється об'єм V_k усередині ізоенергетичної поверхні зі зміною значення енергії. Об'єм еліпсоїда визначається його півосями (1.14), тому

$$V_k = \frac{4}{3} \pi c_x c_y c_z = \frac{4}{3} \pi \frac{(m_x^* m_y^* m_z^*)^{1/2}}{\hbar^3} [2(E - E_m)]^{3/2}$$

Диференціюючи цей вираз по E і підставляючи похідну в (1.30), після деяких перетворень отримаємо шукану функцію розподілу густини енергетичних станів:

$$g(E) = 4\pi (2m_d^* / \hbar^2)^{3/2} (E - E_m)^{1/2}. \quad (1.31)$$

У цей вираз для компактності запису введена величина

$$m_d^* = \sqrt[3]{m_x^* m_y^* m_z^*}, \quad (1.32)$$

що є середньою з компонентів тензора ефективної маси, звана *ефективною масою густини станів*. Відмітимо, що густина

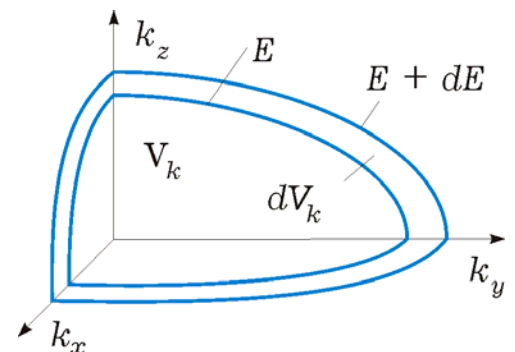


Рис. 1.17. Елементарний об'єм k -простору, що відповідає збільшенню енергії, для підрахунку густини станів є еліпсоїдальним шаром

енергетичних станів нормована на об'єм і має розмірність $1/\text{м}^3 \cdot \text{Дж}$, або $1/\text{м}^3 \cdot \text{еВ}$.

Отримана функція розподілу в принципі придатна як для нижнього, так і для верхнього краю енергетичної зони. Але для станів біля верхнього краю, де $m^* < 0$ і $E < E_{\text{max}}$ ця формула незручна і доцільно представити її, помінявши знаки

$$g(E) = 4\pi \left(2m_{pd}^* / h^2 \right)^{3/2} (E_{\text{max}} - E)^{1/2}. \quad (1.33)$$

Тут $m_{pd}^* = -m_d^*$ і $(E_{\text{max}} - E)$ додатні величини.

Таким чином, з виразів (1.31) і (1.33) виходить, що густина станів у дозволених зонах параболічно зростає з віддаленням від краю всередину зони (рис. 1.18, а). У забороненій зоні стани, природно, відсутні. Вдалині від краю зони питання про знаходження функції $g(E)$ аналітично дуже складне і зазвичай вирішується експериментально.

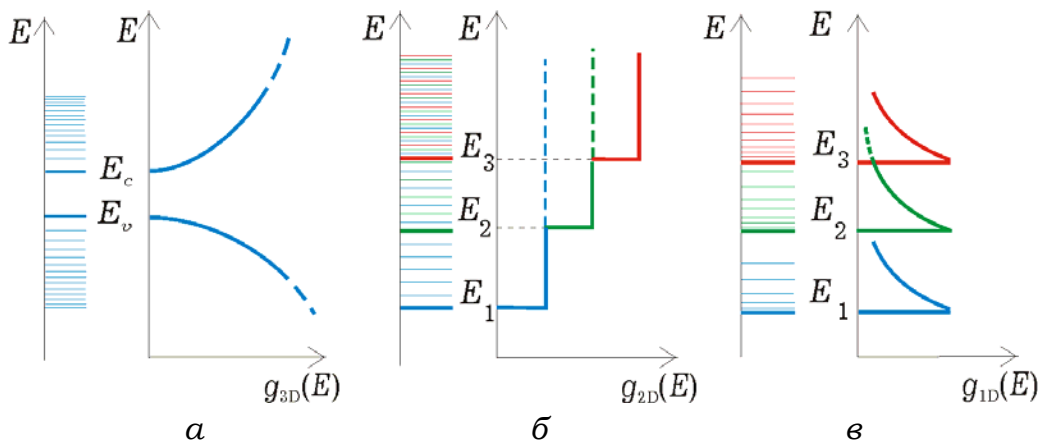


Рис. 1.18. Густина енергетичних станів біля країв зон з віддаленням від краю параболічно зростає для тривимірних (3D) кристалів (а), постійна для двовимірної (2D) ґратки (б) і спадає для одновимірного (1D) ланцюжка атомів (в)

На рис. 1.18 для порівняння наведено також розподіл енергетичних станів електронів у дво- та одновимірних кристалах. Для дірок у валентній зоні енергетичний розподіл станів є симетричним електронному розподілу з відповідною заміною початку і напрямку відліку та ефективної маси.

1.1.6. Зонна структура напівпровідників

Напівпровідникові властивості мають прості речовини: бор В, вуглець (графіт) С, кремній Si, германій Ge, олово (сіре) Sn, фосфор Р, миш'як As, сурма Sb, сірка S, селен (червоний) Se, теллур Те, йод J. До напівпровідників відносяться також бінарні з'єднання типу A_xB_{8-x} , де А - елемент групи x, а В - елемент групи (8-x). Найбільш поширені і застосовні з них - речовини типу $A^{III}B^V$, до яких відносяться з'єднання азоту (нітриди), фосфору (фосфіди), миш'яку (арсеніди), сурми (антимоніди) з елементами третьої групи (бору, алюмінію, галію, індію). Перспективними є з'єднання типу $A^{II}B^{VI}$, а також $A^{IV}B^{IV}$. Напівпровідникові властивості проявляють потрібні і більш складні з'єднання, а також деякі органічні речовини.

Широке застосування в електроніці мають атомарні напівпровідникові речовини (германій і кремній). Структура верхніх оболонок цих атомів, а також інших атомів цієї підгрупи IV групи має вигляд ns^2np^2 . Через близьке енергетичне розташування s- і p-рівнів ці атоми є чотиривалентними. Об'єднуючись в кристал за допомогою ковалентних зв'язків, вони утворюють ґратку таку ж, як і у алмазу. Вона містить дві підґратки, складені з двох гранецентрованих кубічних осередків, зміщених один відносно одного на чверть діагоналі куба (рис. 1.19). У такому кристалі кожен атом пов'язаний з чотирма найближчими, які просторово розташовані так, як це наведено на рис. 1.20, а. У кожному зв'язку беруть участь два електрони, по одному від кожного атома, що є характерним для ковалентного зв'язку. Цю ситуацію часто зображують на площині умовно так, як це наведено на рис. 1.20, б.

У разі утворення кристалічної ґратки алмазу з N однакових атомів s-рівні атомів енергетично розщеплюються на дві групи (або зони) підрівнів (рис. 1.21). Для розуміння цього слід нагадати характер взаємодії атомів водню у разі їх об'єднання в молекулу ковалентними зв'язками. За однакових напрямів спінових моментів їх валентних електронів спостерігатиметься збільшення потенціальної енергії, тобто відштовхування атомів. І лише співпадання напрямів приводить до зменшення енергії і, відповідно, до утворення стійкого стану молекули. З точки зору можливих енергетичних станів цю ситуацію розглядають як розщеплення атомарного стану на два молекулярних, нижній з яких є стійким.

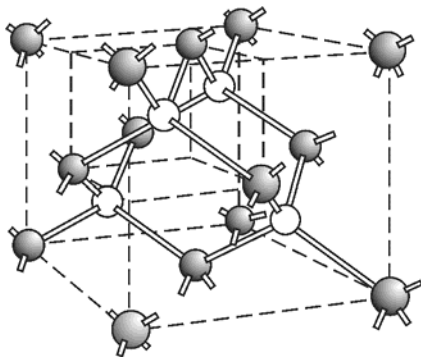


Рис. 1.19. Кристалічна структура алмазу. Затушовано атоми, що утворюють гранецентровану кубічну підгратку. Не затушовані атоми утворюють другу підгратку

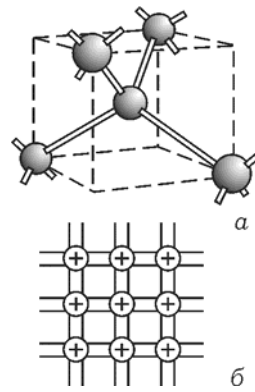


Рис. 1.20. Просторове розташування найближчих атомів у гратці алмазу (а) і її умовне зображення на площині (б)

Подібна ситуація спостерігається і для ковалентних зв'язків у кристалах, тільки замість двох дискретних станів утворюється дві групи станів. Верхній групі, яка складається з N підрівнів, відповідають антисиметричні хвильові функції, яким у разі їх перекривання властиві сили відштовхування. Нижня містить також N станів, електрони в яких мають симетричні хвильові функції, що обумовлює сили тяжіння у разі їх взаємодії.

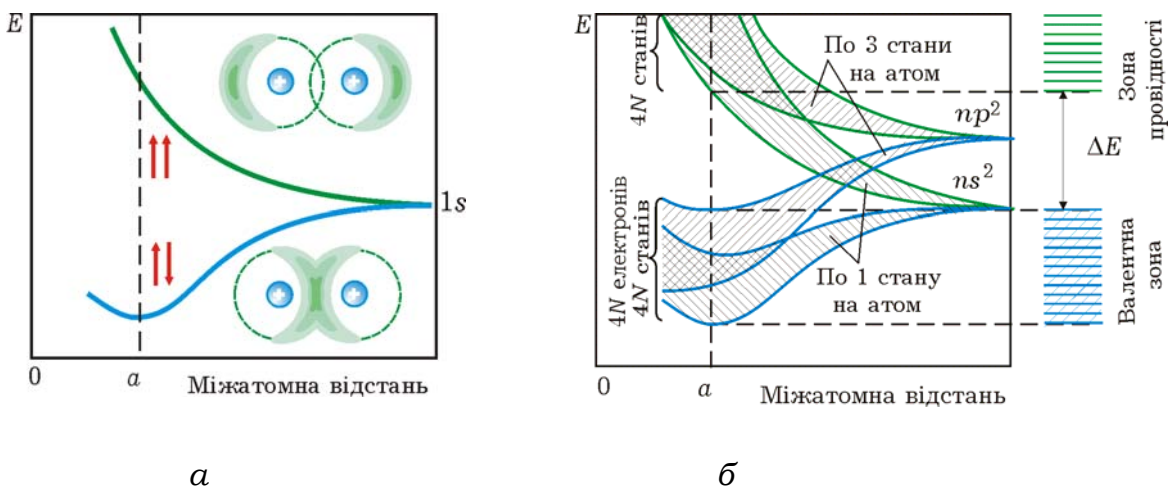


Рис. 1.21. Розщеплення s -станів для ковалентного зв'язку (на прикладі молекули водню) відбувається в залежності від взаємодії спінових моментів (а) Утворення енергетичних зон у разі зближення атомів у ковалентних кристалах елементів IV групи з гратці типу алмазу (б). Зона провідності та валентна зона включають гібридні sp^3 -стани, утворені з атомарних s - і p -станів

Цим двом типам хвильових функцій відповідає істотно різна енергія, тому між двома зонами підрівнів існує заборонена зона. Так само відбувається розщеплювання p -станів на дві зони, у кожній з яких є по 3 стани на атом. У результаті перекриття цих зон, як це показано на рис. 1.21, утворюються дві гібридні sp^3 -зони, які складаються з $4N$ станів кожна (з них N подібні до s -, а $3N$ до p -станів атомів). Природно, що всі $4N$ валентних електрона кристала заповнюють нижні стани, тому нижня зона, яку називають *валентною*, виявляється повністю заповненою, а верхня зона (зона *провідності*) вільною.

Симетричний характер хвильових функцій електронів у валентній зоні призводить до того, що густина електронного заряду буде максимальною уздовж прямих, які умовно зображують (рис. 1.19) зв'язки між атомами у кристалічній ґратці. На плоскій моделі кристалічної ґратки розподіл заряду електронів валентної зони наведений на рис. 1.22, *а*. Для електронів, які знаходяться у станах зони провідності, густина заряду (рис. 1.22, *б*) дорівнює нулю уздовж лінії зв'язку, що веде до переважання сил відштовхування між атомами. Тому перехід електронів з валентної зони у зону провідності призводить до послаблення зв'язків між атомами у кристалі.

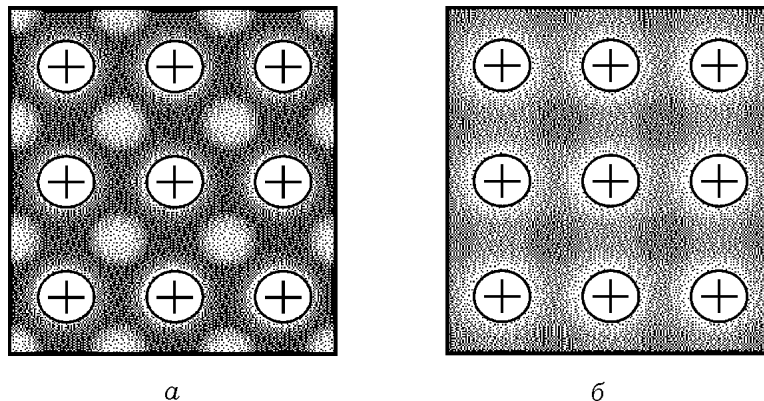


Рис. 1.22. Просторовий розподіл густини заряду електронів валентної зони (*а*) і зони провідності (*б*) на плоскій моделі кристала

Дуже корисно з точки зору описаної картини утворення енергетичних зон порівняти між собою елементи четвертої групи (деякі їх параметри наведені у таблиці 1.1).

Зі збільшенням порядкового номера елементів зростає номер і радіус орбіти, на якій знаходяться валентні електрони. Відповідно ростуть «розміри» атомів і зменшується енергія їх електронної спорідненості. Це супроводжується зменшенням ширини забороненої зони, а значить і енергії зв'язку між атомами в ґратах (кількісною

мірою енергії зв'язку може бути температура плавлення, яка зменшується зі зростанням номера періоду). Відмінність у ширині забороненої зони призводить до того, що алмаз є типовим діелектриком, германій і кремній - типовими напівпровідниками. Олово з ґраткою типу алмазу (α -модифікація або сіре олово) є низькоомним напівпровідником з дуже вузькою забороненою зоною, близькою до 0. Часто зустрічається така модифікація кристалічної ґратки олова (біле олово), за якої воно є типовим металом. Свинець, що замикає наведену таблицю, є також типовим металом.

Таблиця 1.1

№	Елемент	Валентна оболонка	Параметр ґратки, Å	Заборонена зона, еВ	Електричні властивості	$T_{\text{плав}}, \text{K}$
2	Вуглець С (алмаз) $2s^2 2p^2$	3,6	5,4	Діелектрик	4300
3	Кремній Si $3s^2 3p^2$	5,4	1,2	Високоомний напівпровідник	1683
4	Германій Ge $4s^2 4p^2$	5,6	0,78	Напівпровідник	1211
5	Олово Sn (сіре) $5s^2 5p^2$	6,5	~0,1	Низькоомний напівпровідник	905
6	Свинець Pb $6s^2 6p^2$	7,0	0	Метал	601

Описана модель енергетичних зон пояснює наявність від'ємного температурного коефіцієнта ширини забороненої зони. Відомо, що ширина забороненої зони напівпровідника змінюється з температурою приблизно лінійно:

$$\Delta E(T) = \Delta E(0) + \alpha T, \quad (1.34)$$

причому для кремнію і германію величина α від'ємна. Це пояснюється тим, що з підвищенням температури внаслідок теплового розширення збільшується відстань між атомами і відповідно до рис. 1.21 зменшується ширина забороненої зони.

Розглянемо особливості зонної структури найбільш поширених напівпровідникових матеріалів. Як відзначалося вище, періодичність закону дисперсії електронів у кристалі дає можливість його зображення у межах тільки першої зони Бріллюена. Для ґратки типу алмазу вона має вигляд октаедра зі зрізаними вершинами (рис. 1.23).

Тут же показано, як такими фігурами можна заповнити весь k -простір. Замість проєкцій хвильового вектора на основні напрями,

задані числовим набором кристалографічних індексів, часто використовують іншу систему позначень. У ній характерні точки і лінії симетрії у зоні Бріллюена позначаються буквами, а закон дисперсії задається у вигляді одновимірної функції уздовж тих чи інших напрямів між двома точками. Центр зони Бріллюена прийнято позначати буквою Γ , напрями еквівалентні $[100]$ - Δ , а шість точок на поверхні зони, що їм відповідають - X. Вісім еквівалентних напрямів $[111]$ позначають буквою - Λ , а точки їх перетину з поверхнею - L.

Кремній.

Вигляд дисперсійних кривих кремнію для двох напрямів хвильового вектору наведено на рис. 1.22. Наявність декількох кривих у валентній зоні і зоні провідності може бути пов'язана з описаною вище картиною утворення гібридних sp^3 -зон і відповідно з наявністю s- і p-станів, яким властиві різні закони дисперсії. Треба відзначити, що у різних джерелах вид дисперсійних кривих іноді помітно відрізняється, що може пояснюватися як відмінністю в методах обчислення цих кривих, так і різною мірою їх деталізації. Не торкаючись цих відмінностей, розглянемо тільки найбільш суттєві особливості законів дисперсії.

Хвильові функції електронів і відповідні їм дисперсійні криві для валентної зони можуть бути побудовані на підставі хвильових функцій електронів в атомних p-станах. Їм відповідають три криві на рис. 1.24. Дисперсійна крива для s-станів лежить біля нижнього краю валентної зони і на рисунку не показана. Максимум енергії валентної зони вироджений і знаходиться в центрі зони Бріллюена для двох гілок, для третьої за рахунок спин-орбітальної взаємодії він опускається на 0,035 eV. Різна кривизна цих гілок в максимумі дає підставу говорити про різну ефективну масу рухливих носіїв: для «легких» дірок $m_{p,l}^* \approx 0,16m$, для «важких» $m_{p,t}^* \approx 0,5m$. Для обчислення ефективної густини станів біля краю валентної зони зручно користуватися деякою усередненою масою:

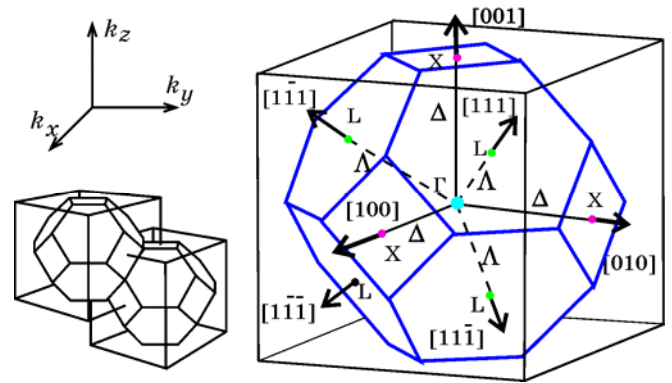


Рис. 1.23. Форма першої зони Бріллюена для кристалічної ґратки типу алмазу і позначення характерних точок і напрямків симетрії

$$m_{pd}^* = \left(m_{pl}^{*3/2} + m_{pt}^{*3/2} \right)^{2/3} \approx 0,56m.$$

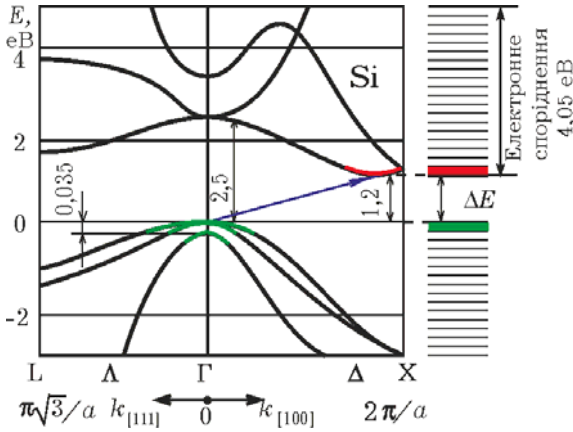


Рис. 1.24. Дисперсійні криві у зоні провідності і валентній зоні кремнію в Λ - і Δ -напрямах.

Відстань між абсолютними екстремумами кривих відповідає забороненій зоні

Структура зони провідності ще складніша, ніж валентної, через те, що тут спостерігається перекриття не лише однієї s - і трьох p -підзон, але і вищих. Абсолютний мінімум в зоні провідності має одна з p -гілок у Δ -напрямі на відстані 0,855 відрізка Γ - X . Беручи до уваги симетрію зони Бріллюена (рис. 1.23), можна вважати, що цей мінімум складається з шести еквівалентних мінімумів, або, як часто їх називають, *долин*. На рис. 1.25 ці долини зображені за допомогою ізоенергетичних поверхонь, які мають форму еліпсоїда обертання. Оскільки півосі еліпсоїда

характеризують компоненти тензора ефективної маси, зручно говорити про поздовжню і поперечну ефективні маси електронів, які для кремнію мають значення відповідно $m_l^* \approx 0,98m$ і $m_t^* \approx 0,19m$.

Іноді для визначення питомої електропровідності користуються деякою усередненою скалярною величиною ефективної маси, яку називають *омічною* (див. розділ 2.3) і зворотню величину якої знаходять як середнє арифметичне з обернених значень ефективних мас у трьох напрямках. Тому $m_c^* = 3m_t^*m_l^* / (2m_l^* + m_t^*) \approx 0,26m$.

Наявність шести еквівалентних долин біля дна зони провідності означає шестикратне збільшення кількості станів. Це враховують, зберігаючи незмінним вид формули (1.27) для функції розподілу густини станів і використовуючи ефективну масу густини, як середнє геометричне з компонентів тензора ефективної маси, помножене на коефіцієнт $M^{2/3}$. Для кремнію $M = 6$, тому $m_d^* = 6^{2/3}(m_t^{*2}m_l^*)^{1/3} \approx 1,08m$.

Ширина забороненої зони для кремнію різна у різних точках зони Бріллюена. У центрі зони для $k = 0$ вона дорівнює $\approx 2,5$ еВ. Проте

шириною забороненої зони, як правило, вважають відстань між абсолютними екстремумами, тобто для кремнію відстань від максимуму валентної зони в точці Γ до мінімуму зони провідності в Δ -напрямі. Ця відстань дорівнює $\approx 1,2$ еВ за абсолютного нуля і $\approx 1,1$ еВ за кімнатної температури, зменшуючись із зростанням температури за законом (1.34), де $\alpha = -4 \cdot 10^{-4}$ еВ/К.

З точки зору форми забороненої зони кремній відносять до так званих *непрямозонних* напівпровідників. Це означає, що отримавши енергію, рівну мінімальній відстані між зонами, валентний електрон може здійснити перехід у зону провідності тільки за умови зміни також і його квазіімпульса. Такий перехід, позначений стрілкою на рис. 1.24, називається *непрямим*. Прямий перехід може відбутися без зміни квазіімпульса, але за більшої енергії. Ці поняття істотні для розгляду різних механізмів генерації і рекомбінації рухливих носіїв заряду. Важливим параметром з точки зору емісійних властивостей є енергія *електронної спорідненості*, яка визначається як відстань від дна зони провідності до рівня вакууму (див. розділ 1.3.1) і для кремнію дорівнює 4,05 еВ.

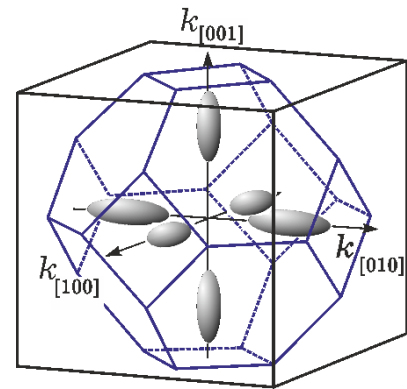


Рис. 1.25. Розташування ізоенергетичних поверхонь, що відповідають Δ -долинам біля дна зони провідності кремнію, у першій зоні Бріллюена

Германій. Закони дисперсії і структура енергетичних зон германію багато в чому подібні до кремнію. Валентна зона (рис. 1.26) має вироджений максимум у центрі зони Бріллюена і характеризується наявністю «легких» і «важких» дірок ($m_{pl}^* = 0,04m$, $m_{pt}^* = 0,34m$), що відповідає ефективній масі густини станів $m_{pd}^* = (m_{pl}^{*3/2} + m_{pt}^{*3/2})^{2/3} \approx 0,36m$. Третя гілка відстає від краю зони на величину енергії спин-орбітальної взаємодії, яка дорівнює 0,28 еВ.

Найбільш суттєва відмінність від кремнію полягає у тому, що як мінімум в центрі зони Бріллюена, так і абсолютний мінімум зони провідності утворені s-гілкою, причому останній лежить на межі першої зони Бріллюена у точці L. Ізоенергетичні поверхні, що відповідають цьому мінімуму, зображені на рис. 1.27 і мають вигляд

еліпсоїдів обертання. Анізотропія ефективної маси у германію виражена ще більше ($m_l^* = 1,58m$, $m_t^* = 0,08m$), тому скалярне значення омічної ефективної маси $m_c^* \approx 3m_t^*/2 \approx 0,12m$. Ефективна маса густини станів електронів у зоні провідності $m_d^* = 4^{2/3} (m_t^{*2} m_l^*)^{1/3} \approx 0,56m$ враховує наявність восьми еквівалентних долин, кожна з яких лежить лише наполовину в першій зоні Бріллюена (рис. 1.26, б), тому $M = 8/2 = 4$.

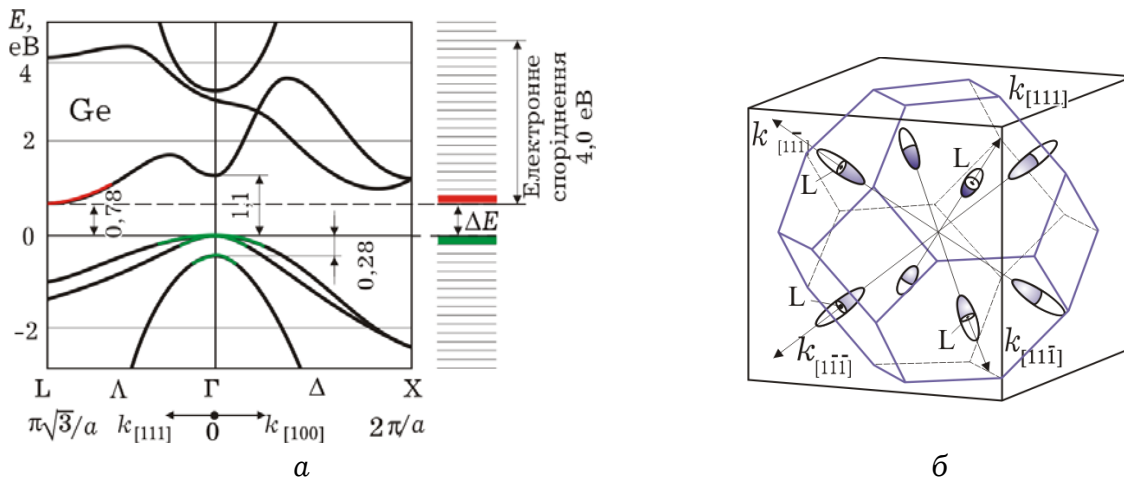


Рис. 1.26. Дисперсійні криві у зоні провідності і валентній зоні германію в Λ - і Δ - напрямках (а). Ширина забороненої зони германію - відстань між абсолютними екстремумами кривих. L-долинам біля дна зони провідності відповідають ізоенергетичні поверхні (б), які лежать лише наполовину в першій зоні Бріллюена

Найбільш суттєва відмінність від кремнію полягає у тому, що як мінімум у центрі зони Бріллюена, так і абсолютний мінімум зони провідності утворені s-гілкою, причому останній лежить на межі першої зони Бріллюена у точці L. Ізоенергетичні поверхні, що відповідають цьому мінімуму, зображені на рис. 1.26, б і мають вигляд еліпсоїдів обертання. Анізотропія ефективної маси у германії виражена ще більше ($m_l^* = 1,58m$, $m_t^* = 0,08m$), тому скалярне значення омічної ефективної маси $m_c^* \approx 3m_t^*/2 \approx 0,12m$. Ефективна маса густини станів електронів у зоні провідності $m_d^* = 4^{2/3} (m_t^{*2} m_l^*)^{1/3} \approx 0,56m$ враховує наявність восьми еквівалентних долин, кожна з яких лежить лише наполовину в першій зоні Бріллюена (рис. 1.26, б), тому $M = 8/2 = 4$.

Ширина забороненої зони германію в центрі зони Бріллюена 1,1 еВ. Відстань між абсолютними екстремумами енергетичних зон, яку зазвичай називають забороненою зоною, дорівнює 0,78 еВ за нульової температури і 0,66 еВ для кімнатної. Температурний коефіцієнт ширини забороненої зони практично такий саме, як і у кремнію, енергія електронної спорідненості у германію 4,0 еВ.

Арсенід галію. GaAs є найбільш поширеним напівпровідником з групи з'єднань $A^{III}B^V$, які іноді називаються інтерметалами. Ці з'єднання зазвичай утворюють кристалічну ґратку типу цинкової обманки, яка за формою подібна до ґратки алмазу і відрізняється від неї тільки тим, що атоми різного виду чергуються, утворюючи дві окремі кубічні гранецентровані підґратки Ga і As. За такої структури кожен атом одного виду пов'язаний з чотирма найближчими сусідами іншого виду по суті ковалентними зв'язками, хоча різна кількість валентних електронів, які усупільнюються від атомів Ga і As (три і п'ять відповідно), обумовлює помітний вклад іонного зв'язку.

У структурі енергетичних зон атомарних напівпровідників IV групи і з'єднань $A^{III}B^V$ багато спільного. Так, наприклад, характер дисперсійних кривих у валентній зоні арсеніду галію аналогічний германію і кремнію. У центрі зони (рис. 1.27) знаходяться максимуми двох гілок, що відповідають «легким» ($m_{рл}^* \approx 0,1m$) і «важким» ($m_{рв}^* \approx 0,5m$) діркам. Третя гілка за рахунок спінорбітальної взаємодії лежить нижче на 0,33 еВ.

На відміну від германію і кремнію більшість з'єднань $A^{III}B^V$ є *прямозонними* напівпровідниками, тобто абсолютний мінімум зони провідності лежить в центрі зони Бріллюена (точка Γ). Для GaAs відстань між екстремумами складає за нульової температури

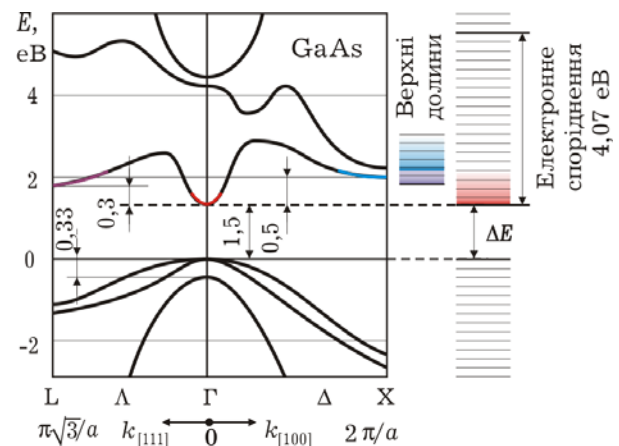


Рис. 1.27. Характер дисперсійних кривих для арсеніду галію. Абсолютні екстремуми (Γ -долини) знаходяться у центрі зони Бріллюена. Ізоенергетичні поверхні, що відповідають найближчим верхнім L-долинам зони провідності, розташовані так само, як і у германії

1,5 еВ, зменшуючись до 1,45 еВ для кімнатної. Ефективна маса електрона поблизу абсолютного екстремуму (у Γ -долині) є скалярною величиною ($m^* \approx 0,067m$), оскільки відповідні ізоенергетичні поверхні практично сферичні.

Однією з найбільш яскравих особливостей арсеніду галію є наявність верхньої долини на висоті $\approx 0,3$ еВ від краю зони провідності в точці L. Ізоенергетичними поверхнями біля дна цієї долини є еліпсоїди обертання, ще більш витягнуті, ніж у германію. Це говорить про велику відмінність поздовжньої і поперечної ефективних мас, які мають значення відповідно $m_l^* = 1,9m$ та $m_t^* = 0,075m$. Так само, як і в германії, розраховується омічна ефективна маса $m_c^* \approx 3m_l^*/2 \approx 0,11m$ і ефективна маса густини станів $m_d^* = 4^{2/3} (m_t^{*2} m_l^*)^{1/3} \approx 0,56m$.

З цією особливістю Γ - і L-долин пов'язаний ефект виникнення *від'ємної диференціальної рухливості*. У зоні провідності GaAs є і вищі долини, наприклад, на відстані $\approx 0,5$ еВ в X-напрямі, але помітний вплив на електричні властивості вони мають тільки за наявності дуже сильного електричного поля. Енергія електронної спорідненості складає 4,07 еВ.

Нітрид галію має дві модифікації кристалічної структури: кубічну, подібну до ґратки для арсеніду галію, або ґратки типу «цинкової обманки», і гексагональну типу «вюрциту». Для *кубічних* ґраток параметри зонної структури наведено на рис. 1.28, для якої в зоні провідності в точках симетрії зони Бріллюена є Γ -, X- і L-долини. Дну зони провідності відповідає центральна Γ -долина з ізотропною ефективною масою $m_\Gamma^* = 0,13m$. Наступні шість X-долин знаходяться на енергетичній відстані $\Delta_{\Gamma X} = 1,3$ еВ на краю зони Бріллюена. Їм відповідає помітна анізотропія ефективних мас $m_l^*/m_t^* = 0,5/0,3$, для якої омічна ефективна маса $m_c^* = 0,35m$, а ефективна маса густини станів з врахуванням кількості долин ($M_X = 6/2$) $m_d^* = M^{2/3} \sqrt[3]{m_t^{*2} m_l^*} = 0,76m$. Більш високі L-долини знаходяться на 2,3 еВ вище, ніж Γ -долина.

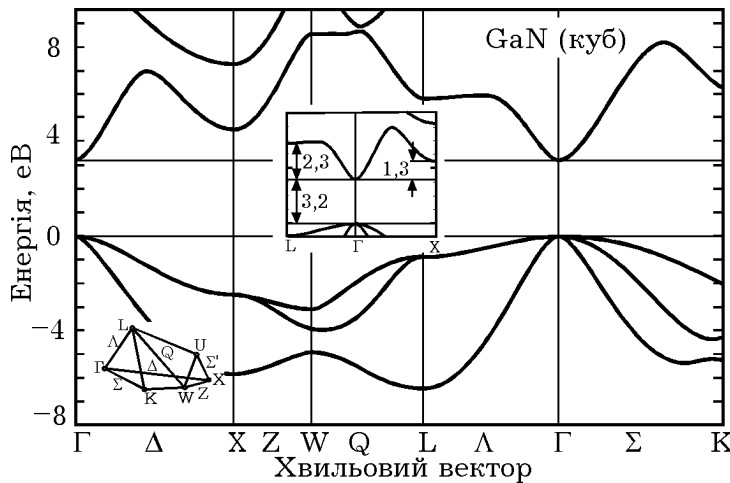


Рис. 1.28. Закон дисперсії для нітриду галію з кубічною структурою кристалічної ґратки. На вставках зображено фрагменти дисперсійних кривих біля країв забороненої зони а також мінімальний елемент симетрії зони Бріллюена з характерними напрямками та точками симетрії

Важливим параметром з точки зору емісійних і контактних властивостей є енергія електронної спорідненості, яка визначається як відстань від дна зони провідності до рівня вакууму і для нітриду галію дорівнює 4,1 еВ.

1.2. ЛОКАЛІЗОВАНІ СТАНИ ЕЛЕКТРОНІВ

Головні висновки зонної теорії отримані для ідеальної кристалічної ґратки. Проте практично усі кристали, які використовуються в електроніці, далекі від структурної досконалості. Це обумовлено, по-перше, технологічними труднощами отримання досконалих кристалів і, по-друге, спеціальним створенням дефектів для надання речовині потрібних властивостей. Різні дефекти (домішки, дислокації, поверхневі дефекти, межзеренні межі та ін.) впливають на зонну структуру, а значить, і на електричні властивості речовини; їх роль особливо зростає у разі використання плівкової технології. Найбільш вивченим є вплив домішкових атомів, які спеціально вводяться для цілеспрямованої зміни електричних властивостей речовини.

1.2.1. Донорні домішки

Класичним і простим прикладом донорного напівпровідника є германій або кремній з ґратками типу алмазу, в яких частина атомів напівпровідника заміщена п'ятивалентними домішковими атомами. Чотири валентні електрони кожного домішкового атома беруть участь в утворенні зв'язків з сусідніми атомами ґрат, п'ятий валентний електрон залишається пов'язаним зі своїм атомом. Проте енергія

цього зв'язку через «відволікаючий» вплив інших атомів у кристалі набагато менша, ніж для вільних атомів, і складає, як буде показано далі, величини порядку 10^{-2} еВ. У результаті деякого збудження домішковий атом може бути іонізований і п'ятий електрон стає вільним. На мові енергетичної моделі напівпровідника процес іонізації означає появу рухливого електрона у зоні провідності. Тому такий домішковий атом називають *донором*, а напівпровідник з п'ятивалентною домішкою - *донорним*.

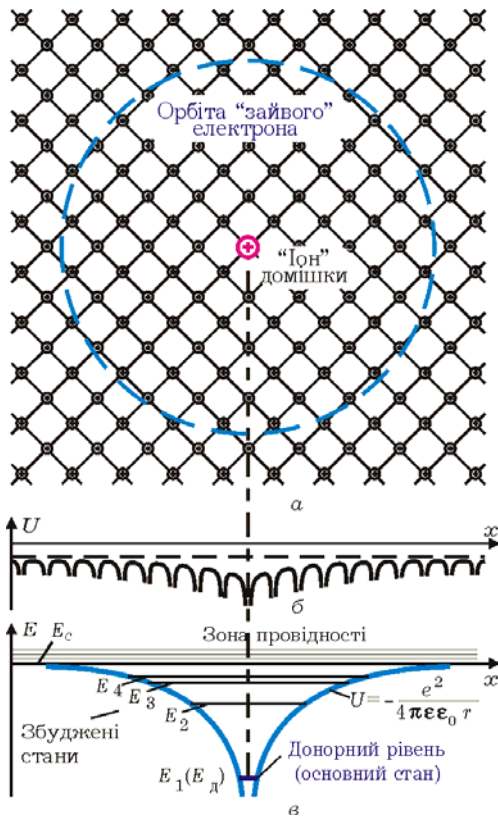


Рис. 1.29. П'ятивалентний атом домішки у чотиривалентному напівпровіднику. «Зайвий» електрон знаходиться на відстанях близько десятка періодів кристалічної ґратки (а). Спотворення періодичного поля ґратки полем іона домішки (б) і енергетичний спектр «зайвого» електрона (в). Масштаб (в) у порівнянні з (б) збільшено

Для визначення можливих енергетичних станів «зайвого» електрона необхідно розв'язати рівняння Шредінгера у полі кристалічної ґратки, періодичність якої спотворена присутністю домішкового атома. Характер збурення зображений на рис. 1.29, б і може розглядатися як накладення на періодичне поле ґратки кулонівського поля одноразово зарядженого по відношенню до електрона домішкового атомного залишку. Для спрощення рівняння Шредінгера у такому полі приймемо наступні допущення. По-перше, врахуємо періодичне поле ґратки, використовуючи метод ефективної маси, розглядаючи рух електрона як вільної частки з масою m^* в центральному полі атомного залишку. Друге припущення відноситься до вибору збурюючої потенціальної функції $U(\vec{r})$ у цьому рівнянні. Поле атома домішки

послаблюється екрануючим впливом сусідніх атомів кристала, що можна врахувати введенням діелектричної проникності ϵ матеріалу основної речовини.

Основою для використання такого макроскопічного параметра, як діелектрична проникність, є той факт, що п'ятий електрон домішки віддалений від свого атома і радіус його орбіти складає декілька періодів кристалічної ґратки (рис. 1.29, а). Враховуючи це послаблення, запишемо вираз для потенціальної енергії електрона

$$U(\vec{r}) = -e^2/4\pi\epsilon\epsilon_0 r. \quad (1.35)$$

Початок відліку цієї енергії відповідає мінімальній енергії вільного від атома електрона, тобто нижньому краю зони провідності E_c (рис. 1.29, в). Підставляючи (1.35) у рівняння Шредінгера, отримаємо

$$-\frac{\hbar^2}{2m^*} \nabla^2 \psi - \frac{e^2}{4\pi\epsilon\epsilon_0 r} \psi = E\psi, \quad (1.36)$$

що співпадає з рівнянням для атома водню з тією різницею, що замість маси електрона тут стоїть ефективна маса, а замість діелектричної проникності вакууму ϵ_0 - діелектрична проникність напівпровідника $\epsilon\epsilon_0$. Природньо, що і результат розв'язку (1.36) буде таким самим, як і для атома водню з урахуванням вказаних змін.

Головним підсумком є висновок про те, що енергетичний спектр електрона буде дискретним і визначається квантовим числом $n = 1, 2, 3, \dots$ (рис. 1.29, в):

$$E_n = -\frac{m^* e^4}{8\epsilon^2 \epsilon_0^2 \hbar^2 n^2} = -\frac{m e^4}{8\epsilon_0^2 \hbar^2} \frac{m^*}{m} \frac{1}{\epsilon^2} \frac{1}{n^2} = -13,6 \frac{m^*}{m} \frac{1}{\epsilon^2} \frac{1}{n^2} \text{ (eV)}. \quad (1.37)$$

В основному стані ($n = 1$) електрон знаходиться на рівні E_1 (рис. 1.29, в), який називають *донорним рівнем* E_d і який при $m^* \approx 0,1m$ і $\epsilon \approx 10$ лежить приблизно на 0,01 еВ нижче за дно зони провідності. Ця величина $E_c - E_d = \Delta E_d$ дорівнює енергії іонізації донорного атома. Значення енергії іонізації донорних атомів в германії і кремнії наведені у таблиці. 1.2.

Таблиця 1.2

Донори	Енергія іонізації, еВ		Акцептори	Енергія іонізації, еВ	
	Германій	Кремній		Германій	Кремній
P	0,0120	0,044	B	0,0104	0,045
As	0,0127	0,049	Al	0,0102	0,057
Sb	0,0096	0,039	Ga	0,0108	0,065
Bi	0,0127	0,039	In	0,0112	0,160
			Tl		0,0131

Вид хвильової функції електрона в основному стані E_1 , як і в атомі водню, характеризується множителем $\exp(-r/r_m)$, де r_m має той же сенс, що і радіус першої борівської орбіти $a_0 = h^2 \epsilon_0 / \pi m e^2 = 0,53 \text{ \AA}$, і відповідає відстані, на якій радіальна густина ймовірності $4\pi r^2 |\psi|^2$ знаходження електрона в точці r максимальна. З урахуванням прийнятих припущень відносно діелектричної проникності і ефективної маси

$$r_m = \frac{h^2 \epsilon \epsilon_0}{\pi m^* e^2} = a_0 \epsilon \frac{m}{m^*}. \quad (1.38)$$

Для $m^* \approx 0,1m$ і $\epsilon \approx 10$ радіус орбіти «зайвого» електрона $r_m \approx 53 \text{ \AA}$, що складає приблизно десять періодів кристалічної ґратки. З цих оцінок виходить, що сферична область локалізації електрона містить більше тисячі атомів напівпровідника, і тому використання у цій моделі такого макроскопічного поняття, як діелектрична проникність цілком виправдане.

Окрім основного стану E_d електрона у домішковому атомі існують ще і *збуджені стани*, які відповідають (1.37) для $n = 2, 3, 4, \dots$: $E_n = E_d / n^2$. Спектр цих станів зображений на рис. 1.29. Наявність рівнів збудження призводить до появи додаткових ліній в експериментальних спектрах поглинання світла домішковими напівпровідниками.

1.2.2. Акцепторні домішки

Схожу ситуацію можна спостерігати у напівпровідниках четвертої групи у разі заміщення деякої їх частини атомами третьої групи. Три валентних електрони домішкового атома беруть участь в утворенні зв'язків з чотирма сусідніми атомами, при цьому один валентний зв'язок залишається вільним. Ця вакансія може бути заповнена електроном з іншого валентного зв'язку під дією якогось збурення, внаслідок чого в системі зв'язків створюється вільний стан, який може рухатися під дією, наприклад, електричного поля. Цей рух зручно інтерпретувати як рух деякої квазічастинки, званої *діркою*. На мові енергетичної моделі такий процес відповідає переходу електрона з валентної зони у більш високий вакантний стан, пов'язаний з атомом домішки. У результаті домішковий атом стає від'ємно зарядженим іоном, а у валентній зоні утворюється

вільна дірка. Це пояснює, чому такий атом називають *акцептором*, а напівпровідник з домішкою атомів III групи *акцепторним*.

Для аналізу електронних станів в акцепторі треба мати на увазі, що у безпосередньому оточенні біля атома всі валентні зв'язки заповнені, а вакантний стан («дірка») знаходиться, як і п'ятий електрон у донорі, досить далеко від центру (рис. 1.30). Таке локальне зміщення електронної густини призводить до виникнення від'ємного електричного заряду в околі атома домішки і викликає порушення періодичного потенціалу ґратки, як це зображено на рис. 1.30, б. Це збурення можна описати кулонівським розподілом потенціалу, використовуючи поняття діелектричної проникності, але потенціальна енергія дірки у полі від'ємного заряду відрізняється від (1.35) тільки знаком і початком відліку. У результаті ми приходимо знову до моделі атома водню з тією різницею, що поле «ядра» спрямоване протилежно.

Хвильові функції і відповідний до них спектр енергії електрона у такій системі можуть бути знайдені, як і у попередньому випадку, за допомогою рівняння Шредінгера для моделі воднеподібного атома у наближенні ефективної маси. Враховуючи, що остання біля верхнього краю валентної зони є від'ємною, і змінюючи знаки, легко отримати рівняння Шредінгера для дірки, яке не відрізнятиметься від електронного в задачі з донорним атомом.

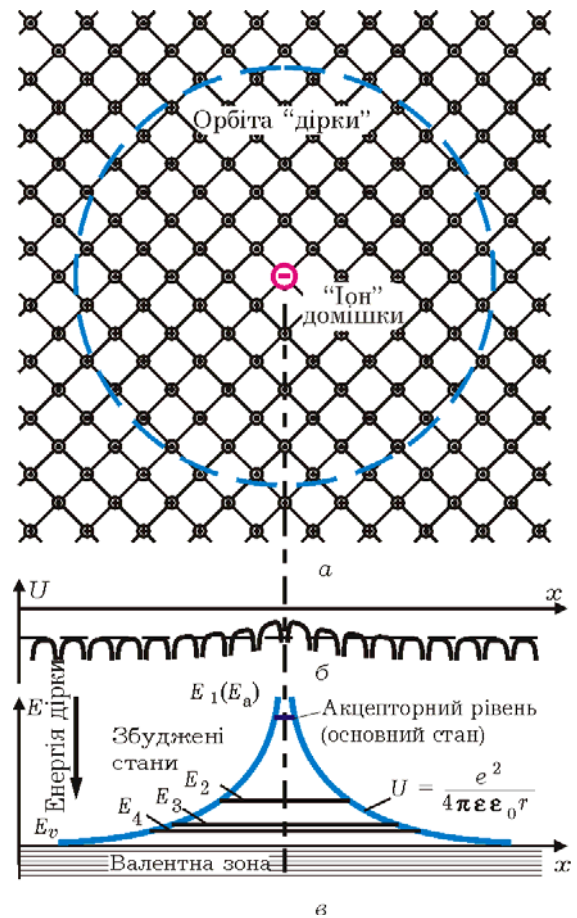


Рис. 1.30. Тривалентний атом домішки у чотиривалентному напівпровіднику «притягує» валентний електрон для заповнення найближчих зв'язків, утворюючи від'ємний «іон». Незаповнений зв'язок або «дірка» знаходиться на відстанях порядку десятка періодів ґратки. Діаграми (б) і (в) показують спотворення поля ґратки «іоном» і можливі енергетичні стани «дірки» в області локального збурення поля

Природньо, що і енергетичний спектр дірки у полі акцепторного атома не відрізнятиметься (окрім позначень) від зображеного на рис. 1.29, в. Електронну енергетичну діаграму можна отримати, лише змінивши напрям відліку енергії.

У результаті можна зробити висновок про те, що вакантні стани відсутнього в акцепторному атомі електрона мають воднеподібний спектр, що включає основний стан E_a і збуджені E_2 , E_3 і так далі (рис. 1.30, в), причому енергія створення від'ємного акцепторного іона $\Delta E_a = E_a - E_v \approx 0,01$ еВ. Значення енергії іонізації акцепторних атомів у германії і кремнії наведені в таблиці. 1.2.

Про просторову локалізацію цієї вакансії можна судити з радіусу орбіти, який, як і в донорному атомі, у багато разів більший за період кристалічної ґратки і складає приблизно декілька десятків ангстрем. Але треба відзначити, що в акцепторних напівпровідниках, на відміну від донорних, рівні збудження у дослідах з поглинання світла не спостерігаються.

1.2.3. Багатозарядні домішки, глибокі рівні

Модель воднеподібного атома задовільно описує електронні стани домішок III і V груп у германії та кремнії. Це видно з таблиці. 1.2 по тому, що результати вимірів енергії іонізації, по-перше, слабо залежать від виду домішки і, по-друге, непогано відповідають обчисленим значенням: $\sim 0,01$ еВ для германію ($m^* \approx 0,2m$ і $\epsilon = 15,8$) та $\sim 0,05$ еВ для кремнію ($m^* \approx 0,5m$ і $\epsilon = 11,8$). Такі домішкові рівні, які на енергетичній діаграмі розташовані поблизу краю відповідних зон, називають *мількими*.

Домішки, для яких енергія іонізації є співмірною з шириною забороненої зони напівпровідника, утворюють *глибокі* домішкові рівні. Для аналізу глибоких енергетичних станів розглянута модель непридатна, оскільки у них радіус орбіти електронів значно менший, ніж для мільких домішок, і такі поняття, як діелектрична проникність і ефективна маса, вже не можуть використовуватися. Навіть у тому випадку, якщо радіус орбіти все ж у декілька разів більший за період ґратки, важко зв'язати глибокий рівень з властивостями електронів в одній з дозволених зон і використовувати ту чи іншу ефективну масу. Поте навіть і для глибоких домішкових станів можна встановити деякі прості закономірності.

Глибокі рівні утворюються домішковими атомами, в яких кількість валентних електронів відрізняється від чотирьох більше, ніж на одиницю. Буде атом донором або акцептором залежить від того, надмірно або недостатньо у нього електронів для встановлення валентних зв'язків з атомами основної ґратки. Проте потрібно мати на увазі, що у цьому випадку домішкові атоми можуть віддавати або приймати більше одного електрона і ставати одно-, дво- чи трикратно зарядженими іонами. Такі домішки іноді називають *багатозарядними*.

Атоми I групи для заповнення валентних зв'язків можуть приймати до трьох електронів і утворювати в забороненій зоні три акцепторні рівні (рис. 1.31) на відстанях від валентної зони, що відповідають різній мірі іонізації домішкових атомів. У домішкових атомів II групи незаповненими є два валентних зв'язки, тому вони утворюють два акцепторні рівні. Атоми VI групи мають, окрім тих, що беруть участь у валентних зв'язках, два «зайві» електрони і є донорами. Природно, що для повторної іонізації атома потрібна більша енергія і відповідний донорний рівень лежить глибше.

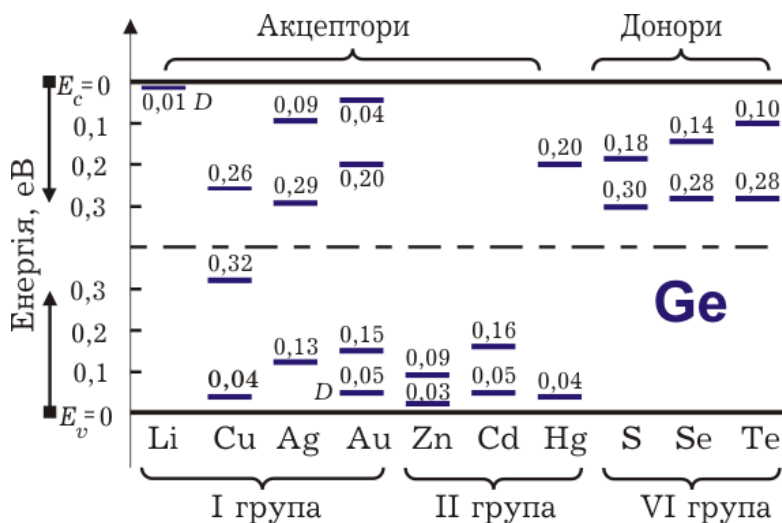


Рис. 1.31. Положення енергетичних рівнів, утворених різними домішками у германії. Відповідні значення енергій відраховуються від краю найближчої дозволеної зони

Наведені міркування є далеко не повними і не можуть описати усієї різноманітності властивостей усіх домішкових атомів. Наприклад, атом золота у германії, окрім трьох акцепторних утворює ще й донорний рівень, який лежить приблизно на 0,05 eV вище за верхній край валентної зони. Це означає, що єдиний валентний електрон атома золота не на рівних з іншими бере участь в утворенні валентних зв'язків і для його переходу в зону провідності потрібна менша енергія, ніж для валентних електронів напівпровідника. Такі домішки, що утворюють і донорні, і акцепторні рівні, називаються *амфотерними*. Розглянемо ще один приклад «нетипової» поведінки домішкового атома. Замість трьох очікуваних, як у міді, глибоких

акцепторних рівнів атом літію утворює один мілкий донорний рівень. Це можна пояснити тим, що через свої малі розміри атом літію розташовується не у вузлах кристалічної ґратки як домішка заміщення, а впроваджується у міжвузля. Тому його валентний електрон, не беручи участі в утворенні валентних зв'язків, виявляється слабо пов'язаним з атомом і подібний у своїй поведінці до зайвого електрона п'ятивалентної домішки.

Відзначимо, що глибокі домішкові стани мають велике значення, наприклад, у розумінні перебігу рекомбінаційних процесів у напівпровідниках.

1.2.4. Вплив концентрації домішки на зонну структуру

Енергетичні електронні стани, утворені окремими домішковими атомами, називають *локальними*, оскільки електрони в цих станах пов'язані з атомами і локалізовані в їх околі. Кількісною мірою величини області локалізації може бути радіус орбіти електрона, який за порядком величин дорівнює одиницям (для глибоких домішок), десяткам (для мілких домішок) і навіть сотням (у деяких з'єднаннях $A^{III}B^V$) періодів кристалічної ґратки. Сукупність локальних рівнів, які виникають за випадкового, але більш-менш рівномірного розподілу по об'єму напівпровідника N домішкових атомів, утворює домішковий рівень (рис. 1.32, а), який залежно від природи домішки може бути донорним або акцепторним.

Таке положення зберігається, поки концентрація домішкових атомів у напівпровіднику незначна. Збільшення її приводить до зближення атомів і за деякої їх концентрації настає положення, коли домішкові атоми починають взаємодіяти між собою. Потенціальна енергія зайвих електронів (для донорів) вже не може бути описана кулонівським полем точкового заряду, і замість рівняння (1.35) у розрахунках енергетичного спектру домішки слід використовувати вираз

$$U(\vec{r}) = -\sum_i^N e^2 / 4\pi\epsilon\epsilon_0 |\vec{r} - \vec{R}_i|, \quad (1.39)$$

де підсумовування проводиться за всіма домішковими атомами, положення яких в ґратках визначається радіус-вектором \vec{R}_i . Доданки цієї суми характеризують взаємодію даного електрона з

«чужими» атомами домішки. Оскільки середня відстань між домішковими атомами порядку $N_d^{-1/3}$, то для $N_d \approx 10^{17} \text{ см}^{-3}$ вона буде співмірною з радіусами орбіт r_m (1.38), а енергія взаємодії - з енергією іонізації домішки. У цьому випадку спектр електронних станів не може бути описаний виразом (1.37), а окремі дискретні стани розщеплюватимуться у зони домішкових станів (рис. 1.32, б) подібно до того, як це мало місце у разі утворення енергетичних зон в кристалі.

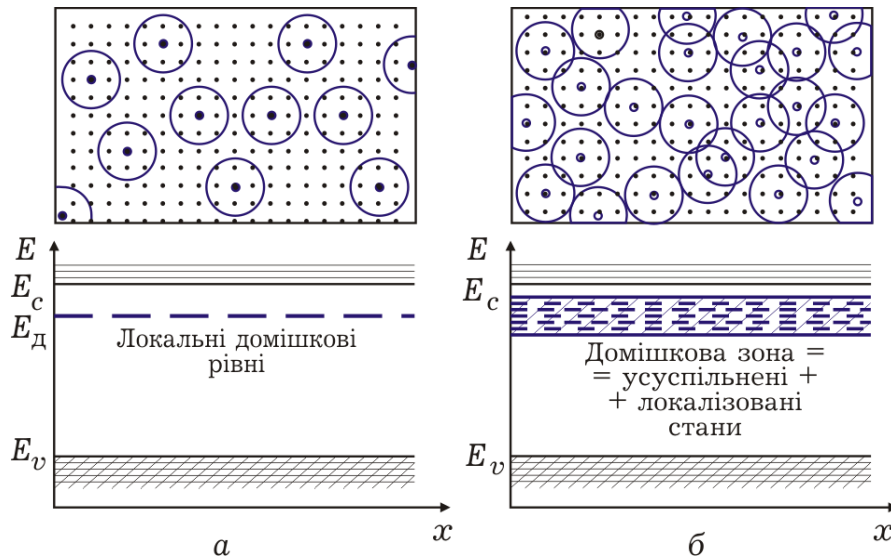


Рис. 1.30. За малої концентрації домішки сукупність локалізованих станів утворює домішковий рівень (а). Взаємодія домішкових атомів у разі їх великих концентрацій призводить до розщеплення локальних станів і утворення енергетичної зони (б)

На рис. 1.33 зображена тільки одна домішкова зона, яка відповідає основному стану E_d . Хвильові функції електронів у зоні перекриваються і перестають бути локалізованими в околі окремих атомів. Результуючі хвильові функції мали б вигляд блохівських функцій, якби розташування домішкових атомів було рівномірним і періодичним, як у ідеальному кристалі. Природно, що ширина домішкової зони залежить від міри перекриття хвильових функцій і зростає зі збільшенням концентрації домішки. При $N_d \approx 10^{19} \text{ см}^{-3}$ домішкова зона перекривається із зоною провідності, утворюючи спільну зону, нижні стани якої заповнені електронами домішкових атомів. Це нагадує зонну структуру металів, а деякі властивості таких напівпровідників, які називаються *виродженими*, близькі до властивостей металів.

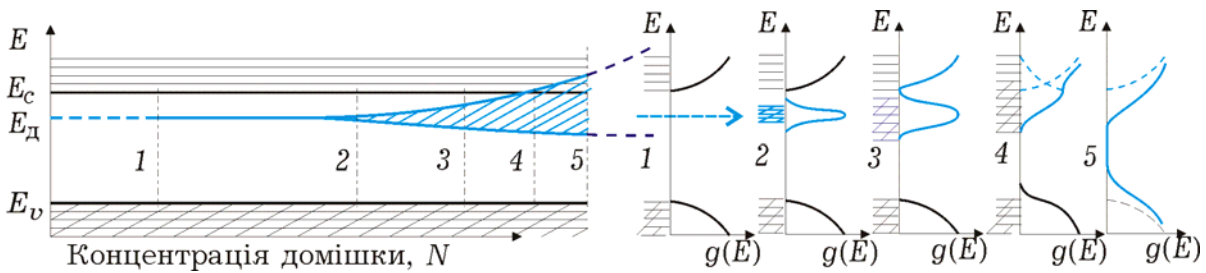


Рис. 1.33. Зміна енергетичного спектру (а) і розподіл густини станів (б) зі збільшенням концентрації домішки. У разі 4 спостерігається виродження напівпровідника. 5-й випадок відповідає аморфній структурі, коли в енергетичних зонах з'являються «хвости» густини станів

З випадковим характером розподілу домішки у кристалі пов'язана одна особливість домішкових зон, яка відрізняє їх від енергетичних зон ідеального кристала. Останні мають чітку межу, оскільки функція розподілу густини енергетичних станів зменшується до нуля на краю зони як $(E - E_m)^{1/2}$. Густина станів у домішковій зоні, яка має вигляд дельта-функції $\delta(E - E_d)$ для дискретних локальних станів, «розмивається» і набирає вигляду дзвоноподібної кривої (рис. 1.33, б). Ширина цієї кривої збільшується із зростанням концентрації домішки, але обов'язково таким чином, що інтеграл від функції розподілу густини станів дорівнює концентрації домішки. У результаті енергетичні зони виродженого напівпровідника мають так звані «хвости» густини станів. Їх поява спостерігається і у валентній зоні навіть тоді, коли напівпровідник донорний. Пояснюється це випадковим характером розподілу домішкових атомів, які помітно порушують періодичність ґратки основної речовини і неминуче викликають зміну структури його енергетичних зон, у тому числі і валентної.

Збільшення концентрації домішки до такої міри, що її кількість стає співмірною з кількістю атомів основної речовини, призводить до ситуації, коли, по-перше, саме поняття домішки втрачає сенс і, по-друге, можливе утворення нерегульованої, тобто *аморфної*, структури. Оскільки висновки про зонний характер енергетичного спектру отримані за умови періодичності кристалічної ґратки, то в аморфних твердих тілах, здавалося б, має бути відсутнім зонний спектр. Проте експерименти підтверджують наявність заборонених і дозволених зон у некристалічних твердих тілах. Відповідальним за це є ближній порядок, тобто таке взаємне розташування і відповідно взаємодія ближніх атомів, яка характерна для кристалів. На відміну від енергетичних зон кристала, зони в аморфних тілах мають згадані

хвости густини станів, обумовлені відсутністю *дального порядку*. Особливістю таких хвостів є те, що вони значною мірою складаються з локалізованих станів, а електрони в них мають незначну рухливість і їх вклад в електропровідність або інші кінетичні явища може виявитися незначним.

1.3. ЕЛЕКТРОННІ СТАНИ НА ПОВЕРХНІ

Енергетичні стани електронів на поверхні твердого тіла повинні відрізнятися від об'ємних станів. Такий висновок можна зробити з того, що розподіл останніх описується зонною теорією, результати якої ґрунтуються на припущенні про трансляційну симетрію кристалічної гратки і можуть використовуватися, строго кажучи, тільки для нескінченного, необмеженого кристала. Для реальних кристалів скінченних розмірів умова такої симетрії приблизно виконується для внутрішньої частини об'єму за порівняно обмеженої трансляції, що надає можливості і у цьому разі з успіхом застосовувати висновки зонної теорії. На поверхні у результаті обриву гратки періодичність потенціалу порушується (рис. 1.34), і питання про стани електрона повинне вирішуватися окремо.

У зв'язку з цим під час розгляду електронних станів на поверхні можна поставити три питання:

1) яка форма потенціального бар'єру на поверхні і якими параметрами він характеризується; 2) які особливості енергетичного спектру електронів у приповерхневому шарі; 3) яка роль сторонніх атомів, які часто у помітній кількості адсорбовані на поверхні.

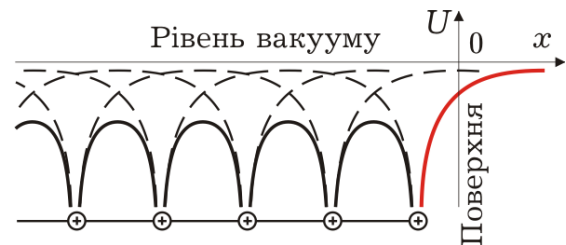


Рис. 1.34. Порушення періодичності потенціальної функції біля краю кристалічної гратки

1.3.1. Макроскопічна енергетична модель поверхні

Потенціальна функція електрона поблизу поверхні твердого тіла легко може бути визначена для макроскопічного розгляду, коли поверхню можна вважати гладкою. У разі виходу електрона за межі твердого тіла на нього діє сила з боку не зкомпенсованого

позитивного заряду (рис. 1.35, а). Обчислення сили взаємодії може бути проведене *методом дзеркального відображення*, суть якого полягає в тому, що розподілений по поверхні заряд замінюється точковим додатним зарядом $+e$, що знаходиться по інший бік від поверхні на такій самій відстані, що і електрон (рис. 1.35, б). Сила, що діє на електрон, відповідно до закону Кулона, дорівнює

$$F(x) = -e^2/4\pi\epsilon_0(2x)^2 = -e^2/16\pi\epsilon_0x^2. \quad (1.40)$$

Потенціальна енергія електрона дорівнює роботі з його переміщення з цієї точки на нескінченність і визначається як інтеграл від сили $F(x)$:

$$U(x) = -e^2/16\pi\epsilon_0x. \quad (1.41)$$

Отримана потенціальна енергія дорівнює 0 для $x \rightarrow \infty$, що відповідає мінімально можливій енергії вільного електрона у вакуумі, і називається *рівнем вакууму*. Функція (1.41) описує форму потенціального бар'єру, що існує поблизу поверхні і перешкоджає виходу електронів з твердого тіла. Робота, що здійснюється електроном з подолання сил цього потенціального бар'єру, називається *роботою виходу*.

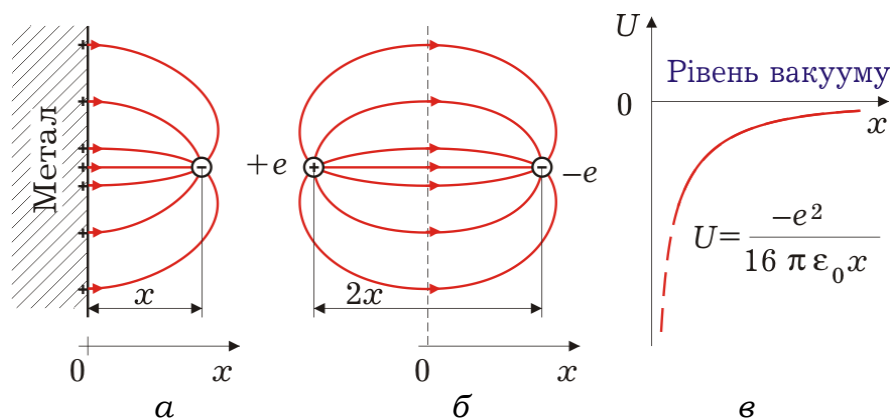


Рис. 1.35. До розрахунку форми потенціального бар'єру на поверхні металу (в). Енергія взаємодії електрона з поверхнею (а) визначається як енергія його взаємодії з дзеркально відображеним щодо поверхні зарядом $+e$ (б)

Введене таким чином поняття роботи виходу не є визначеним, оскільки не вказує на початковий рівень відліку цієї величини, і тому кожен електрон матиме своє «індивідуальне» значення роботи виходу.

Якщо відлічувати її від мінімально можливої енергії вільного електрона в металі, тобто від дна зони провідності E_c , то отримаємо так звану *повну роботу виходу* $U_0 = E_0 - E_c$ (рис. 1.36, а).

Проте в металі електрони навіть за $T = 0$ мають деяку енергію, максимальне значення якої визначається рівнем Фермі (див розд. 3.3). Для подолання бар'єру їм треба виконати меншу роботу, і її значення, відраховане від рівня Фермі, називають *термодинамічною* (іноді ефективною) *роботою виходу* $\Phi = E_0 - E_f$. Природно, що $U_0 > \Phi$ на величину енергії Фермі.

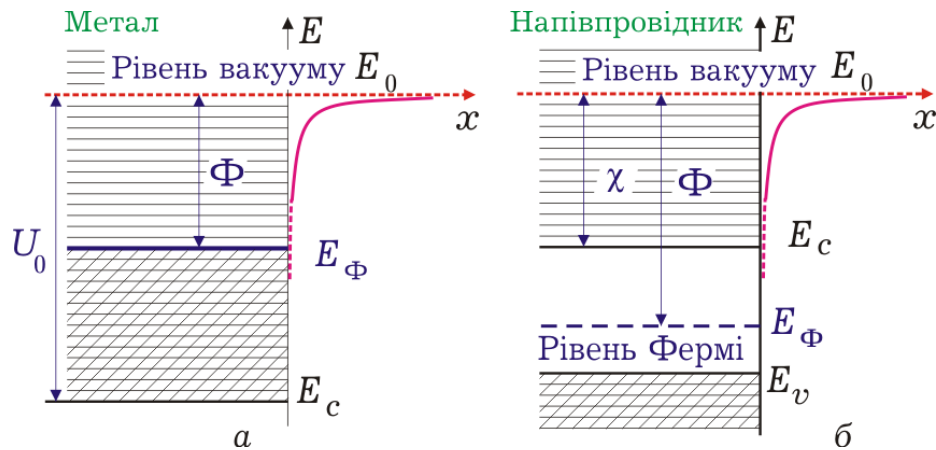


Рис. 1.36. Позначення повної U_0 і термодинамічної Φ робіт виходу, а також енергії електронної спорідненості χ на енергетичних діаграмах металу (а) і напівпровідника (б)

Для напівпровідників поняття термодинамічної роботи виходу зберігає свій сенс, не дивлячись на те, що рівень Фермі може знаходитися у зоні заборонених станів (рис. 1.36, б) і жоден з електронів у разі виходу не виконує такої роботи. Енергію, яка потрібна для переведення електронів з дна зони провідності напівпровідника на рівень вакууму, називають *енергією електронної спорідненості* $\chi = E_0 - E_c$.

Виразом (1.41) можна користуватися для опису тільки верхньої частини потенціального бар'єру на відстані від поверхні, більшої за декілька періодів кристалічної ґратки, коли поле ґратки стає гладким і можливе уявлення про поверхню як про плоску. На близькій відстані отриманий вираз для потенціальної функції, а значить і узгодження рівня вакууму з яким-небудь характерним енергетичним рівнем усередині кристала, тобто обчислення роботи виходу, є складним завданням. Тому зазвичай використовують експериментальні значення роботи виходу.

1.3.2. Таммівські поверхневі стани

Проблема знаходження форми потенціального рельєфу поблизу поверхні виникає і у разі аналізу енергетичного спектру поверхневих станів електрона. Незважаючи на складність кількісних обчислень, загальні міркування тут досить нескладні і корисні. Для електрона з енергією, вищою за рівень вакууму, блохівська хвильова функція для внутрішньої області кристала повинна поступово переходити в хвилю, що біжить і яка описує вільний електрон у вакуумі. Якщо енергія електрона менша за E_0 і відповідає зоні дозволених станів, то його хвильова функція повинна мати вигляд експоненти, що спадає углиб потенціального бар'єру у вакуумі, а поблизу поверхні «зшивається» з хвилею Блоха для об'ємних станів.

Особливий інтерес представляють поверхневі електронні стани (рис. 1.37, а), які потрапляють у заборонену зону об'ємних. З боку вакууму хвильова функція, як і у попередньому випадку, має вигляд експоненти, а з боку кристала подібна до блохівської функції, амплітуда якої експоненціально затухає на глибині в декілька періодів ґратки (рис. 1.37, б).

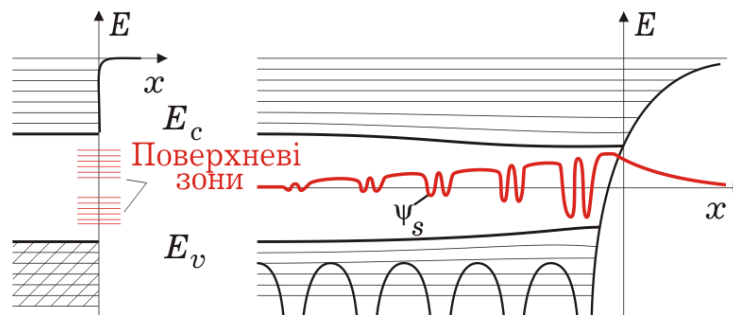


Рис. 1.37. Умовне зображення зон поверхневих станів на енергетичній діаграмі (а) і приблизний вигляд відповідної хвильової функції (б)

Важливою особливістю цих хвильових функцій є те, що завдяки існуванню трансляційної симетрії у напрямках, паралельних поверхні кристала, вони мають відповідно до теореми Блоха вид хвиль $\psi_s(x, y, z) = u(x, y, z)\exp[i(k_y y + k_z z)]$, що біжать, причому амплітуда $u(x, y, z)$ є періодичною з періодом ґратки у напрямках y і z . Ця обставина породжує, як і у тривимірному випадку цілу зону (або зони) поверхневих станів (рис. 1.37, а). Але на відміну від об'ємних ці

зони локалізовані поблизу поверхні і поступово зливаються з об'ємними зонами у міру зменшення поверхневого збурення і перетворення потенціальної функції в періодичну. Таким чином, поверхневі стани є не додатковими, а модифікованими об'ємними. Їх кількість відповідає кількості атомів, що граничать з поверхнею, і за порядком величин складає $10^{14} \dots 10^{15} \text{ см}^{-2}$. Поверхневі стани, пов'язані з обривом ґратки, були уперше досліджені І. Є. Таммом і їх називають таммівськими. Вони завжди присутні на поверхні і грають істотну роль в електронних процесах.

1.3.3. Адсорбовані поверхневі стани

Стани іншого виду виникають за наявності на поверхні адсорбованих атомів. Поведінка цих атомів, їх вплив на стан поверхні залежить від виду атомів і може розглядатись за допомогою такої нескладної моделі. Нехай на поверхні металу знаходиться атом (рис. 1.38, а). Між кристалічною решіткою і атомом існує потенціальний бар'єр, порівнянний за прозорістю з міжатомними бар'єрами кристала. Якщо енергія іонізації атома менша за термодинамічну роботу виходу з металу, тобто валентний рівень атома знаходиться вище за рівень Фермі у металі, то можливий перехід валентного електрона у метал.

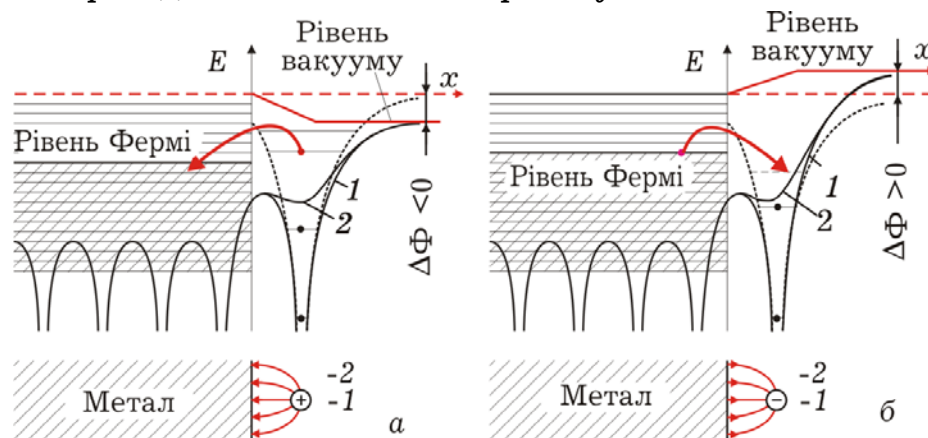


Рис. 1.38. Зміна роботи виходу в разі адсорбції на поверхні металу електропозитивних (а) і електронегативних (б) атомів

У результаті переходу атом стає додатньо зарядженим іоном, а електростатичне поле, що виникає, створює сили, які утримують атом на поверхні (*сили адсорбції*). Іншим не менш важливим наслідком цього електронного обміну є зменшення роботи виходу електронів з металу. Відповідна зміна форми потенціального бар'єру

показана на рис. 1.38, а для перерізів, що проходять через центр атома (крива 1) і поза ним (крива 2). Зниження рівня вакууму, природно, помітне за досить великої кількості адсорбованих атомів. Такі атоми (зазвичай атоми лужних або лужноземельних металів) по відношенню до поверхні є *електропозитивними*.

Цікавою є залежність (рис. 1.39) роботи виходу від міри покриття катода плівкою. Природно, що зі збільшенням кількості адсорбованих атомів робота виходу повинна змінюватися від значення для чистої поверхні (4,5 еВ для вольфраму) до значення роботи виходу речовини плівки (3,4 еВ для торію, 2,5 еВ для барію і 1,9 еВ для цезію). Проте ця залежність не буде монотонною, як можна було очікувати, якби атоми залишалися нейтральними. Іонізація атомів і утворення прискорюючого електричного поля знижує роботу виходу до значень менших, ніж робота виходу покриваючої речовини. Найбільший ефект досягається, коли атоми покривають 60..70% поверхні катода. За більшої міри покриття помітною стає взаємодія адсорбованих атомів між собою, що призводить до зменшення ймовірності їх іонізації і зменшення напруженості електричного поля.

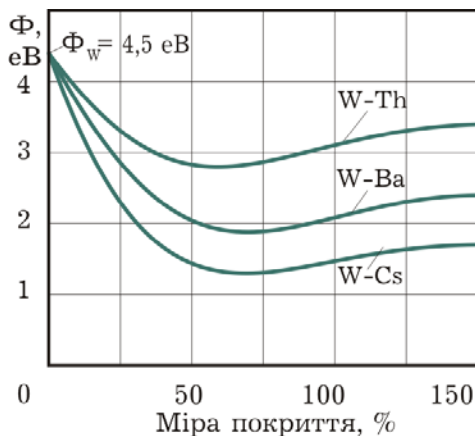


Рис. 1.39. Залежність ефективного роботи виходу вольфраму від ступеня покриття поверхні адсорбованими атомами

Протилежна ситуація виникає для електронегативних атомів, коли робота виходу з металу менша за енергію електронної спорідненості адсорбованих атомів. Під час переходу електрона з металу на розташований нижче вільний рівень атома утворюється від'ємний іон. Електричне поле, що виникає між іоном і поверхнею, створює, по-перше, силу адсорбції, а по-друге, додаткову силу, яка перешкоджає виходу електронів з металу. На енергетичній діаграмі це виглядає як підвищення потенціального бар'єру і збільшення роботи виходу. Типовими електронегативними атомами є, наприклад, атоми кисню, азоту, хлору та ін.

Можливою є також ситуація, коли робота виходу з металу менша за енергію іонізації атома, але більша за його електронну спорідненість. Іон у такому випадку не утворюється, але обмінні

Протилежна ситуація виникає для електронегативних атомів, коли робота виходу з металу менша за енергію електронної спорідненості адсорбованих атомів. Під час переходу електрона з металу на розташований нижче вільний рівень атома утворюється від'ємний іон. Електричне поле, що виникає між іоном і поверхнею, створює, по-перше, силу адсорбції, а по-друге, додаткову силу, яка перешкоджає виходу електронів з металу. На енергетичній діаграмі це виглядає як

сили (типу сил Ван-дер-Ваальса), що виникають за рахунок перекриття хвильових функцій призводять до взаємодії металу і атома та до поляризації останнього. Наявність у адсорбованого атома дипольного електричного моменту також деформує потенціальний бар'єр і впливає на роботу виходу.

Механізм взаємодії адсорбованих атомів з поверхнею напівпровідника принципово такий самий, як і для металів. Але адсорбовані атоми, у залежності від їх спектру поверхневих станів, можуть грати також роль донорів або акцепторів, центрів рекомбінації і так далі, і за достатньої їх концентрації відчутно впливати на електричні властивості поверхні. На рис. 1.40, а зображений приклад ситуації, коли енергія електронної спорідненості напівпровідника менша за енергію іонізації адсорбованого атома. У цьому випадку обмін валентним електроном відбувається тільки під дією якого-небудь зовнішнього збудження, наприклад, тепла. Такі атоми є поверхневими донорами і за їх іонізації на поверхні утворюється електричне поле, що змінює (як і для металів) висоту і форму потенціального бар'єру.

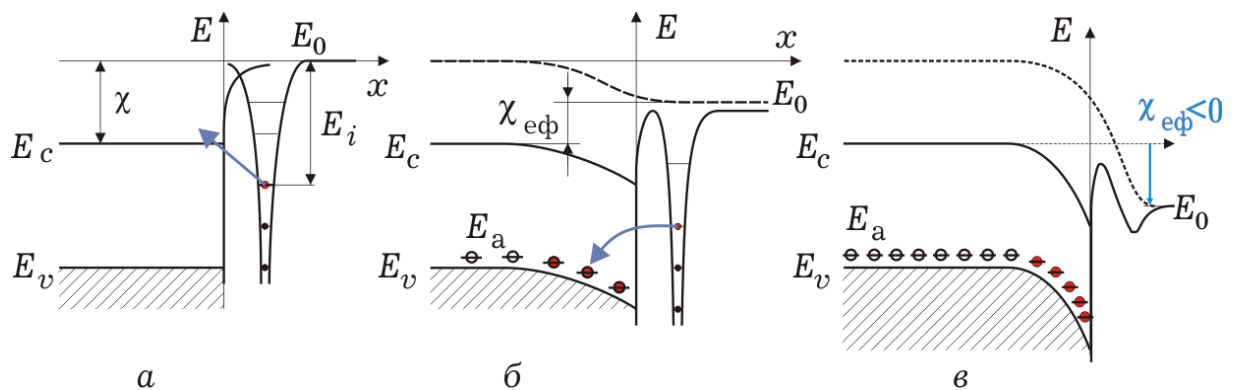


Рис. 1.40. Енергетична діаграма напівпровідника з адсорбованим на поверхні «донорним» атомом (а). Вигин зон біля поверхні напівпровідника у разі адсорбції електропозитивних атомів зменшує ефективну енергію електронної спорідненості (б), яка за певних умов може бути від'ємною (в)

Проте для напівпровідників у порівнянні з металами спостерігаються дві істотні відмінності. По-перше, ймовірність іонізації поверхневих атомів, а значить і сила електричного поля залежить від виду і концентрації домішки в об'ємі. Наприклад, наявність акцепторів (рис. 1.40, б) утворює біля нижнього краю забороненої зони вільні стани, які можуть заповнюватися валентними електронами адсорбованих атомів навіть без зовнішнього збудження. По-друге, поле подвійного електричного шару не закінчується на поверхні, як в

металах, а проникає у приповерхневий шар на деяку відстань. Потенціальна енергія цього поля на енергетичних діаграмах викривлює зони поблизу поверхні вниз для вказаного напрямку поля. Зменшення відстані між рівнем вакууму і дном зони провідності у результаті вигину означає зменшення ефективної енергії електронної спорідненості.

Цей ефект використовується для створення *емітерів з від'ємною електронною спорідненістю* (рис. 1.40, в). У таких катодах на шар сильнолегованого акцепторами матеріалу p^+ -GaAs нанесена моноатомна плівка цезію, яка частково оксидована. Крутий вигин зон виникає за рахунок сильного поля між шаром від'ємно іонізованих акцепторів і шаром додатних іонів цезію. Молекули осиду цезію є диполями, електричне поле яких формує різкий спад потенціального бар'єру, його товщина складає декілька ангстремів. У результаті рівень вакууму розташовуватиметься нижче за дно зони провідності у глибині напівпровідника, що і означає від'ємну електронну спорідненість. Такі фотокатооди мають підвищену чутливість в інфрачервоному діапазоні і використовуються у приладах нічного бачення.

1.4. ЕЛЕКТРОННІ СТАНИ ТА МЕХАНІЗМИ ПРОВІДНОСТІ

1.4.1. Зв'язок зонної структури та провідності

Класифікацію твердих тіл з точки зору електричних властивостей можна проводити за різними ознаками. Простим є поділ речовин за значеннями їх питомої електропровідності. Речовини, які мають $\sigma > 10^6$ См/м, прийнято вважати провідниками, $\sigma < 10^{-8}$ См/м діелектриками, а проміжні значення властиві напівпровідникам. Зрозуміло, що така класифікація твердих тіл за величиною їх електропровідності, що змінюється до того ж більше, ніж на 20 порядків, не може бути універсальною, бо не враховує особливостей різноманітних механізмів провідності, існуючих у твердих тілах.

Інший поширений спосіб класифікації ґрунтується на температурній залежності електропровідності. Відомо, що електропровідність металів із зростанням температури спадає приблизно пропорційно T^{-1} , а напівпровідників - експоненціально зростає і залежно від зовнішніх умов, як-то: температура,

освітленість, тиск та ін., може змінюватися у широких межах. Відомо також, що наявність домішок у металах зазвичай зменшує їх електропровідність, а у напівпровідниках збільшує.

Усі ці факти добре пояснюються зонною теорією, а енергетична зонна структура твердих тіл може бути основою їх класифікації за електричними властивостями. Уся різноманітність зонних структур кристалів, яка виникає внаслідок перекидання їх енергетичних зон, заповнених електронами або вільних, зон для різних квантових станів і різних кристалографічних напрямів, може бути зведена до двох видів, які відрізняються взаємним розташуванням зони, заповненої валентними електронами, і найближчої зони вільних станів (рис. 1.41). Енергетичні зони, які утворені з більш глибоких атомних рівнів, істотно не впливають на електричні властивості і можуть не братися до уваги.

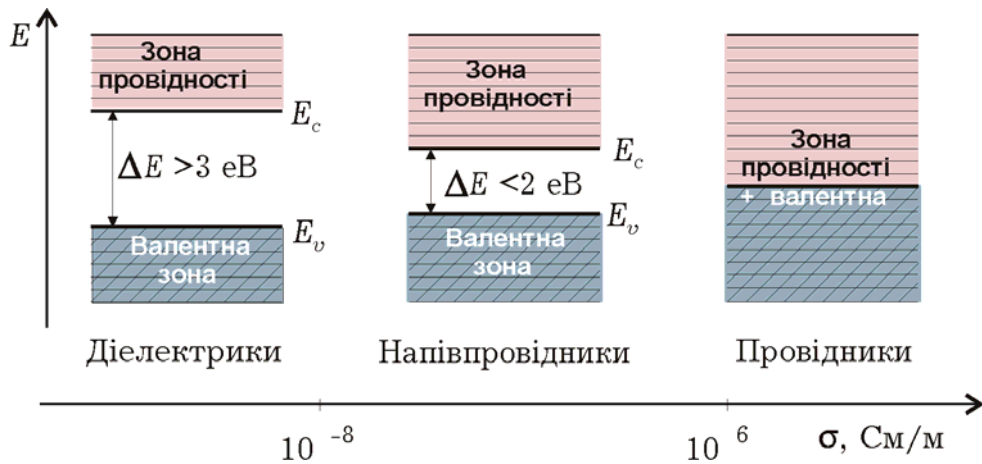


Рис. 1.41. Діаграма з зображенням поділу речовин за значеннями їх питомої електропровідності та її зв'язок зі структурою енергетичних зон

Речовини, у яких заповнені енергетичні стани безпосередньо змикаються із зоною вільних станів, є такими, що проводять електричний струм. Якщо зони вільних і заповнених станів, які називаються відповідно зоною провідності і валентною, розділені забороненою зоною, то така речовина у нормальному стані є непровідною і відноситься до діелектриків або напівпровідників. Відмінність між ними є кількісною і не може бути чіткою. Якщо ширина забороненої зони, тобто енергетична відстань між нижнім краєм зони провідності (дном) E_c і верхнім краєм валентної зони (стелею) E_v перевищує 3 еВ, така речовина є типовим діелектриком. Типовими напівпровідниками вважають речовини із забороненою зоною, меншою за 2 еВ. Проте останнім часом використовуються

широкозонні напівпровідники, наприклад, нітрид галію, карбід кремнію та ін., які мають ширину забороненої зони більшу за 3 eV і які набувають напівпровідникових властивостей за рахунок домішок.

У чому ж причина такої відмінності електричних властивостей кристалів і як це пов'язано з їх зонною структурою? На перший погляд валентні електрони, що є усупільненими, тобто не пов'язаними з окремими атомами, і такими, які описуються хвильовими функціями у вигляді хвиль Блоха, повинні вільно пересуватися уздовж кристала і під дією електричного поля забезпечувати провідність. Річ у тому, що поле прискорює електрони і участь їх у провідності обов'язково пов'язана зі зміною їх енергетичних станів. Якщо ж усі стани у зоні заповнені, то електричне поле не може переводити електрони в інші стани і спрямованого руху у цьому разі не виникає. Виняток становлять дуже сильні електричні поля, під дією яких електрон може бути переведений у зону вільних станів. Але ці випадки розглядатимуться окремо, а звичайні поля з напруженістю навіть в ~ 1 кВ/см на довжині вільного пробігу ~ 1 мкм можуть надати електрону енергію тільки 10^{-1} eV, що звичайно недостатньо для його переходу у зону провідності. У провідних речовинах вільні стани розміщені енергетично настільки близько, що електрони прискорюватимуться під дією навіть дуже слабого електричного поля.

З наведених міркувань виходить, що тільки ті електрони, які розміщені у частково заповнених зонах, можуть брати участь у провідності. Ця ситуація реалізується у металах навіть за нульової температури. Напівпровідники і діелектрики можуть проводити струм, якщо частина електронів перейде з повністю заповненої валентної зони у зону провідності під дією, наприклад, нагрівання, освітлення, сильного електричного поля або якогось іншого зовнішнього чинника.

Таке пояснення зв'язку електричних властивостей з особливостями заповнення енергетичних зон, проведене за допомогою тільки «одновимірних» енергетичних діаграм, не можна вважати досить аргументованим і повним: може скластися уявлення, що на усупільнені електрони у повністю заповненій зоні електричне поле не діє. Для більш глибокого розуміння процесу провідності необхідно, по-перше, з'ясувати вплив електричного поля на зміну не лише енергетичних станів, але і значень хвильового вектору, а розгляд проводити з використанням дисперсійної кривої. По-друге, провідність треба розглядати як колективне явище, обумовлене поведінкою усіх електронів зони.

1.4.2. Електрони у повністю заповненій зоні.

Дія зовнішнього електричного поля на рух електронів у кристалі може бути досліджена за допомогою напівкласичного рівняння (1.16):

$$\hbar d\vec{k} / dt = \vec{F}. \quad (1.42)$$

Якщо направити вісь x назустріч силовим лініям електричного поля, то вектор сили, що діє на електрон ($F = F_x = e\mathcal{E}$), співпадатиме з додатнім напрямом x -складової хвильового вектору. Це дає можливість обмежитися одновимірним рівнянням руху $dk_x/dt = e\mathcal{E}/\hbar$, результат інтегрування якого для однорідного статичного поля має вигляд

$$k_x(t) = k_x(0) + (e/\hbar)\mathcal{E}t \quad (1.43)$$

і показує, що у статичному електричному полі хвильовий вектор з часом лінійно зростає. Насправді цей процес обмежується зіткненнями електрона з дефектами кристала, а (1.43) виконується тільки на проміжку між двома зіткненнями. Але для з'ясування реакції електрона на зовнішнє поле процеси зіткнення поки не братимемо до уваги.

Звертаючись до дисперсійної кривої (рис. 1.42), бачимо, що електричне поле збільшує хвильовий вектор і зміщує електрон, що знаходився в стані k_i , у більш високі енергетичні стани. Рухаючись таким чином вздовж дисперсійної кривої, електрон доходить до краю першої зони Бріллюена (точка 1) і переходить у другу, де його енергія зменшується. Зручно скористатися схемою зведених зон (рис. 1.11, в) і обмежитись тільки першою зоною Бріллюена. Враховуючи, що точки 1 і 1' фізично тотожні, можна вважати, що електрон, досягаючи точки 1, переходить в точку 1' і опускається уздовж лівої гілки дисперсійної кривої. Залишаючись таким чином у першій зоні Бріллюена, електрон здійснює циклічне переміщення по станах дисперсійної кривої.

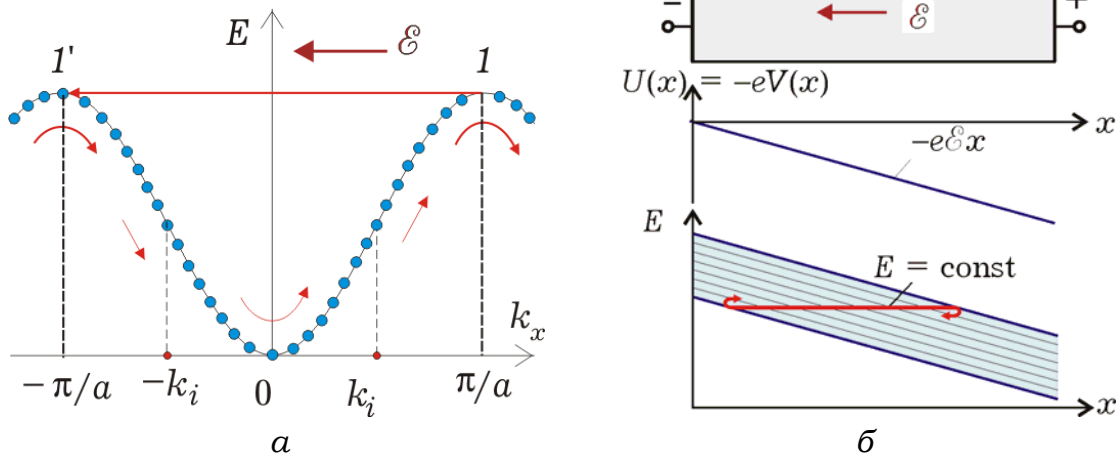


Рис. 1.42. Зовнішнє електричне поле збільшує значення хвильового вектора електронів і вони зміщуються по станах відповідно до закону дисперсії, не порушуючи симетрії заповнення станів повністю заповненої зони (а). Таким переходам між краями зони в електричному полі відповідають просторові коливання (б)

З наведеного опису стає зрозумілим, чому електрони повністю заповненої зони не створюють електричного струму. Якщо усі стани у зоні, які зображено точками на рис. 1.42, а, заповнені електронами і кожен з них реагує на електричне поле, як це описано вище, то постійно усі стани залишаються заповненими. Через симетрію дисперсійної кривої кожному електрону в стані \vec{k}_i відповідає електрон в стані $-\vec{k}_i$, тому векторна сума за усіма N станами зони $\sum_i^N \vec{k}_i = 0$, і переважаючого руху електронів у будь-якому напрямі не відбувається. Саме тому *провідність*, якщо вона існує, *створюється електронами тільки частково заповнених зон*.

У координатному просторі таку поведінку електронів повністю заповненої зони можна зобразити на енергетичній діаграмі (рис. 1.42, б), враховуючи, що електричне поле змінює потенціальну енергію електрона. Для лінійного розподілу потенціала вздовж зразка краї зони спадатимуть за лінійним законом $-e\mathcal{E}x$. Оскільки в статичному електричному полі повна енергія електрона не змінюється, прискорення супроводжується зменшенням потенціальної енергії електрона. Так триватиме, поки електрон не досягне верхнього краю зони (точка 1 на дисперсійній кривій а), де починається зворотний рух і гальмування електрона, яке супроводжується збільшенням його потенціальної енергії

(переміщення по станах дисперсійної кривої вниз з точки 1'). У результаті статичне зовнішнє електричне поле у періодичному полі кристалічної ґратки викликає так звані «блохівські коливання» електрона, не створюючи постійного електричного струму.

1.4.3. Види провідності і носії заряду

У разі часткового заповнення електронами станів у зонах можливі два типові випадки. Перший з них має місце у зоні провідності, де число заповнених станів n зазвичай набагато менше, ніж вільних. У такому разі електрони заповнюють нижні стани (затушовані кружечки на рис. 1.43) симетрично відносно мінімуму дисперсійної кривої. Електричне поле, спрямоване назустріч осі k_x , змінює хвильовий вектор електронів відповідно до (1.43) і вони зміщуються у k -просторі за якийсь час у затемнену на рис. 1.43, б область. Очевидно, що результуючий хвильовий вектор усіх n електронів $\sum_i^n \vec{k}_i$ буде відмінним від нуля, що означає наявність спрямованого перенесення заряду. Величина електричного струму складається з елементарних струмів, які створюються кожним з n електронів зони провідності

$$\vec{j} = (-e) \sum_i^n \vec{v}_i. \quad (1.44)$$

Електрони, розташовані у нижніх станах зони, можна розглядати як *вільні частки з додатною ефективною масою*, тому в межах параболічності закону дисперсії $m^* \vec{v} = \hbar \vec{k}$ і

$$\vec{v}_i = \hbar \vec{k}_i / m^*, \quad (1.45)$$

де m^* має однакове значення для усіх заповнених станів.

Враховуючи це, для повного струму маємо:

$$\vec{j} = (-e\hbar/m^*) \sum_i^n \vec{k}_i. \quad (1.46)$$

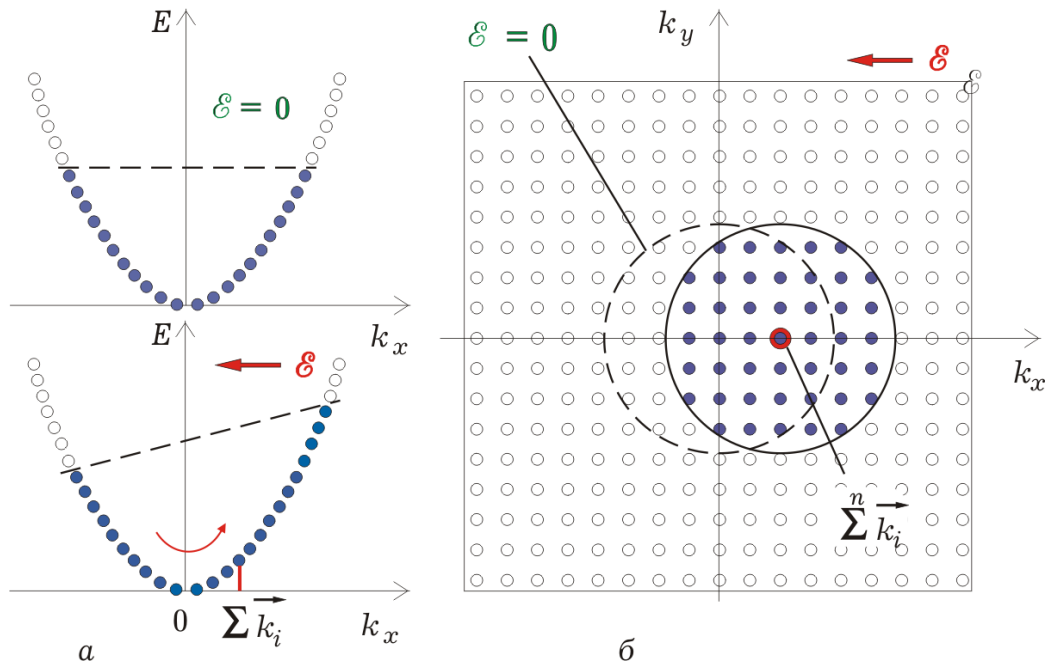


Рис. 1.43. Електрони біля дна частково заповненої зони під дією електричного поля зміщуються по станах так, що їх сумарний хвильовий вектор або сумарна швидкість спрямовані у бік дії сили електричного поля, як і для вільних частинок

Оскільки сумарний хвильовий вектор додатний (рис. 1.43, б) і протилежний до вектора $\vec{\mathcal{E}}$, то електричний струм буде, як завжди, співпадати за напрямом з електричним полем.

Децю інша ситуація спостерігається, наприклад, у валентній зоні, де кількість заповнених станів (n) більша, ніж вільних (p), причому $n + p = N$ - повному числу станів у зоні. У цьому випадку електрони, як і раніше, заповнюють нижні енергетичні стани, так що вільні від електронів стани, позначені на рис. 1.44 світлими кружечками, будуть розміщені симетрично відносно максимуму дисперсійної кривої.

Електричне поле діє на кожен електрон відповідно до (1.43) і зміщує їх, як і у попередньому випадку, у бік більших значень k . Але *колективний* результат виходить зовсім іншим. З форми затемненої області на рис. 1.44, б, яка означає заповнені стани, навіть без детальних розрахунків видно, що сумарний хвильовий вектор усіх електронів $\sum_i^n \vec{k}_i$ спрямований уздовж вектора $\vec{\mathcal{E}}$, тобто *протилежно* зміщенню кожного з електронів! Окрім цієї «незвичності» поведінки електронів також виникають труднощі з

обчисленням струму за формулою (1.46), яку треба для такого випадку записати у вигляді:

$$\vec{j} = -e\hbar \sum_i^n \vec{k}_i / m_i^* \quad (1.46, a)$$

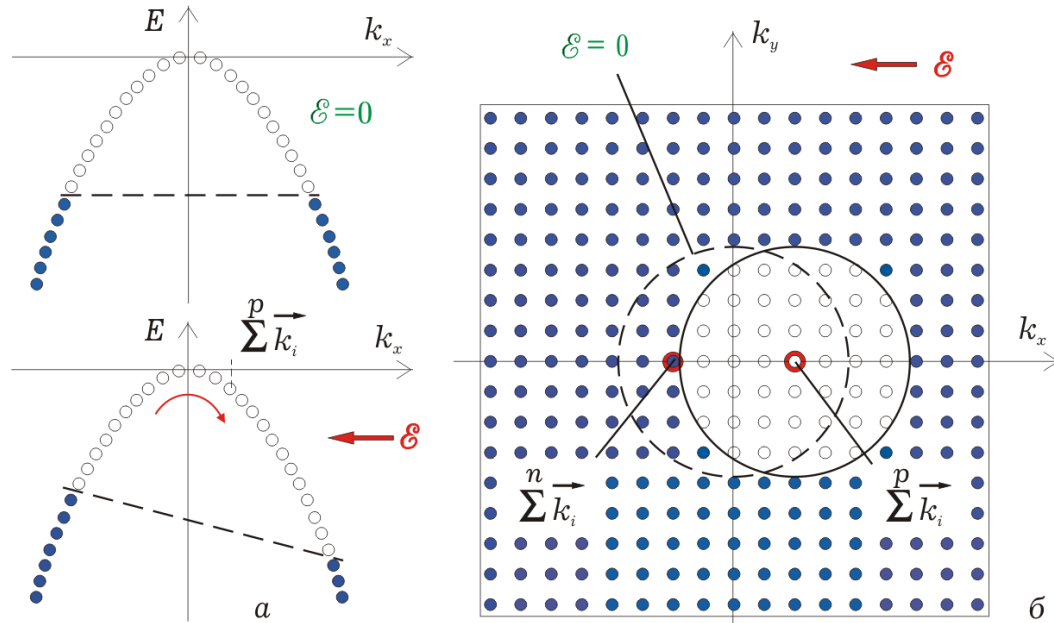


Рис. 1.44. Зміщення електронів майже повністю заповненої зони під дією електричного поля відбувається так, що їх сумарний хвильовий вектор направлений назустріч силі поля

Поняття ефективної маси зручно використовувати лише біля краю енергетичної зони, де дисперсійна крива параболічна, а у формулі (1.46, a) підсумовування потрібно проводити майже за всіма станами зони, де m_i^* має різні значення і метод ефективної маси використовувати немає сенсу. Уникнути цих труднощів можна, розглядаючи поведінку не електронів, а вільних станів, або *дірок*. Формально це здійснюється таким чином. Оскільки $\sum^N \vec{v}_i = \sum^n \vec{v}_i + \sum^p \vec{v}_i = 0$, то $\sum^n \vec{v}_i = -\sum^p \vec{v}_i$, і те ж саме значення струму можна знайти не з формули (1.44), а через суму за *вільними* (дірковими) станами:

$$\vec{j} = e \sum_i^p \vec{v}_i \quad (1.47)$$

Звідси видно, що розглядаючи дірку як деяку умовну частку, ми повинні приписати їй додатній заряд $+e$. Окрім цього, електрон біля

верхнього краю зони має від'ємну ефективну масу і тому, аналогічно до (1.45), матимемо

$$\vec{v}_i = \hbar \vec{k}_i / (-|m^*|), \quad (1.48)$$

де m^* - однакова для усіх вільних станів через квадратичність закону дисперсії. Після підстановки (1.48) у (1.47) бачимо, що повний струм, що створюється електронами майже заповненої зони, можна розрахувати як струм деяких додатніх часток (дірок)

$$\vec{j} = (-e\hbar/|m^*|) \sum_i^p \vec{k}_i. \quad (1.49)$$

Але для цього діркам ще треба присвоїти додатню ефективну масу, рівну за величиною ефективній масі електрона, який «залишив» цю дірку. Напрямок струму (1.49) співпадає з напрямком електричного поля (неначе він створюється додатніми частками), оскільки сумарний хвильовий вектор вільних станів відповідає центру незатемненої області на рис. 1.44, б і $\sum_i^p \vec{k}_i > 0$.

Таким чином, провідність, яка створюється електронами майже порожньої і майже заповненої зон, має істотно різні механізми. Тому для опису цих процесів зручно розглядати *два види квазічасток, або два види рухливих носіїв заряду*: рухливі електрони і рухливі дірки.

Іноді особливості руху дірок у реальному просторі помилково переносять на k -простір. У реальному координатному просторі електрон, рухаючись проти силових ліній електричного поля, зміщується на вільне місце у валентних зв'язках і залишає після себе знову вільне місце, яке може бути зайняте іншим електроном і так далі. Цей рух електронів нагадує «естафетний», і у цьому разі вільні стани (дірки) рухатимуться у напрямі силових ліній поля. У k -просторі дірки під дією поля зміщуються у тому ж напрямі, що і електрони, як це видно з рис. 1.44, б для незатемненої області.

Треба пам'ятати, що дірки є фіктивними частками. Перенесення заряду здійснюється в обох випадках електронами, але у тих випадках, коли стани в енергетичній зоні майже усі заповнені електронами, зручніше (хоча і не обов'язково) електронні процеси розглядати, використовуючи уявлення про дірки. Така незручність пов'язана з дещо незвичайною поведінкою електронів у майже заповненій зоні, а використання дірок приводить механізм провідності до звичних уявлень. Наявність двох механізмів провідності і відповідних носіїв заряду підтверджується

експериментально, наприклад, різними знаками електрорушійної сили, що викликається різними носіями заряду, при спостереженні ефекту Холла.

Контрольні запитання та вправи

1.1. Яким чином взаємодія валентних електронів впливає на їх енергетичний спектр при об'єднанні атомів в кристал?

1.2. У чому полягає головна особливість хвильової функції усупільненого електрона в ідеальній періодичній кристалічній решітці? Які практичні висновки з неї виходять?

1.3. У чому відмінність хвильового вектора електрона в кристалі і вільного електрона? Що таке зона Бріллюена? Скільки і яких станів вона включає?

1.4. Опишіть властивості закону дисперсії електронів у кристалі і його особливості біля краю енергетичних зон. Які типові форми його подання?

1.5. У чому суть методу ефективної маси? Для яких станів його зручно застосовувати?

1.6. Опишіть функції розподілу енергетичних станів біля краю зон у кристалах. Що таке ефективна маса густини станів?

1.7. Опишіть модель утворення напівпровідникової зонної структури у ковалентних кристалах (Si, Ge) і ковалентно-іонних ($A^{III}B^V$).

1.8. Які основні особливості розташування долин і їх позначення в законі дисперсії для найважливіших напівпровідникових матеріалів (Si, Ge, GaAs)?

1.9. Які атоми є типовими донорами? Сформулюйте головну особливість донорних станів, звернувши увагу на енергетичний спектр (енергію іонізації) електронів і область їх просторової локалізації. Чому для опису донорних станів зручно використовувати модель воднеподібного атома? Які типові значення енергії іонізації донорних атомів у найважливіших напівпровідниках?

1.10. Проведіть подібний аналіз для акцепторних атомів, використовуючи симетрію (антисиметрію) моделі порівняно з донорами.

1.11. Які домішки називаються багатозарядними і чому електронні стани в них називають глибокими?

1.12. Як впливає концентрація домішки на енергетичний спектр електронів? Що таке і чому утворюється домішкова зона? За якого порядку концентрації домішки спостерігається виродження напівпровідника?

1.13. Яка фізична природа потенціального бар'єру біля поверхні твердого тіла? Якими параметрами прийнято характеризувати його висоту в металах і напівпровідниках?

1.14. У чому причина відмінності електронних станів в об'ємі кристала і біля його поверхні? Які особливості спектру поверхневих (таммівських) станів?

1.15. У чому фізична причина виникнення сил адсорбції? Які типові прояви електронних станів адсорбованих атомів на поверхневі параметри зонної структури твердих тіл ?

1.16. Поясніть, чому електрони у повністю заповненій зоні не можуть створювати електричний струм, і вкажіть характерний зв'язок зонної структури твердих тіл з їх провідністю. Запишіть напівкласичні рівняння, що описують рух вільних електронів у кристалічній решітці.

1.17. Які типові механізми електронної провідності у частково заповненій зоні провідності і частково вільній валентній зоні? Що таке рухливі носії заряду і які характерні параметри електронів і дірок?

2. ЕЛЕКТРОННИЙ ТРАНСПОРТ

2.1. ПАРАМЕТРИ ПЛАЗМИ

Плазмою зазвичай називають середовище, яке складається із заряджених та нейтральних часток, концентрація яких достатня для значущої ролі взаємодії (співударянь) між ними. Попри відсутність чіткого визначення плазми, існування колективних ефектів, викликаних взаємодією між частками, є чи не найголовнішою ознакою плазми.

Ступінь іонізації плазми, тобто концентрація заряджених часток та її відношення до концентрації нейтральних, може змінюватися на багато порядків. Тому й властивості плазми можуть бути дуже різноманітними. Розглянемо деякі параметри, використовувані для опису властивостей плазми.

Концентрація заряджених часток різного знаку зазвичай в середньому однакова, плазма в цілому є нейтральним середовищем. Про можливі значення концентрації зарядів дає уявлення діаграма на рис. 2.1. Слід звернути увагу на найбільш поширені в електронній техніці плазмові колективи – газорозрядні та твердотільні.

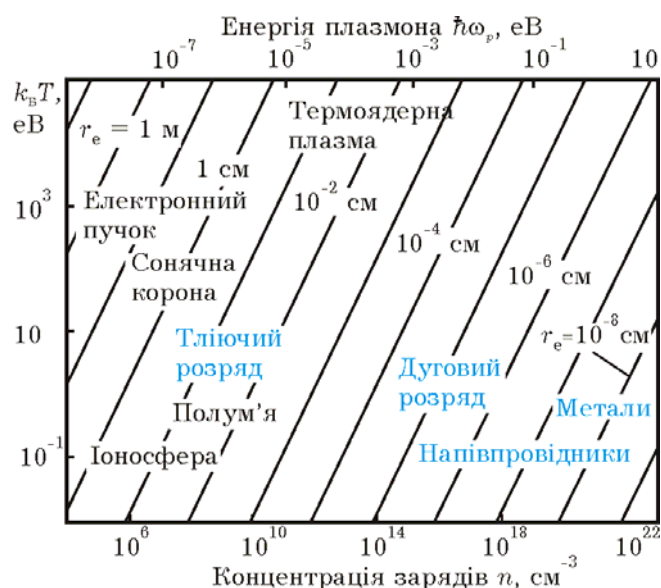


Рис.2.1. Характерні параметри різних видів плазми

Плазмова частота. Хоча плазма в цілому є нейтральною, під дією якихось збурень у ній можливе локальне розділення додатних та від'ємних зарядів, що викликає утворення областей просторового заряду. Електричне поле, яке виникає, намагається повернути заряди у попереднє положення і усунути це розділення. Прискорені полем заряди «проскакують» положення рівноваги і викликають коливання, які й називаються плазмовими. Плазмові коливання характеризуються частотою ω_p , яка залежить від виду заряджених частинок, густини об'ємного заряду та електричного поля, тобто від концентрації.

Найпростіша модель для аналізу плазмових коливань подана на рис. 2.2, де зображено шар у цілому нейтральної плазми, яка складається, наприклад, із додатних іонів кристалічної ґратки та електронного газу. Зміщення шару електронів на відстань Δx відносно положення рівноваги викликає появу подвійного електричного шару, напруженість поля в якому пропорційна концентрації електронів та зміщенню: $\mathcal{E} = en\Delta x / \epsilon\epsilon_0$. Рух електронів у цьому полі можна описати класичним рівнянням

$$m \frac{d^2(\Delta x)}{dt^2} = -e\mathcal{E} = -\frac{e^2 n \Delta x}{\epsilon\epsilon_0}. \quad (2.1)$$

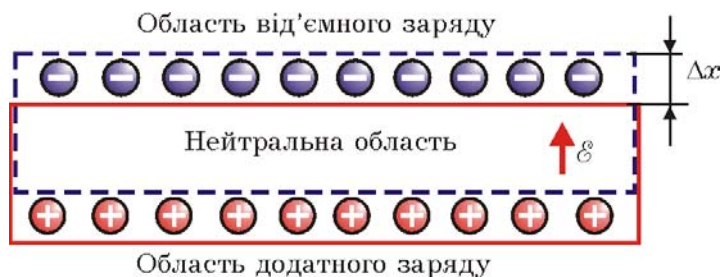


Рис. 2.2. Модель подвійного електричного шару для опису плазмових коливань

Розв'язок цього рівняння показує, що зміщення Δx електронного шару відносно іонного «остову» можна описати гармонічними функціями типу $\cos \omega_p t$ з частотою

$$\omega_p = \sqrt{e^2 n / \epsilon\epsilon_0 m}. \quad (2.2)$$

Цей вираз показує, що плазмова частота визначається концентрацією часток та їх масою. Тому досить часто плазму зручно характеризувати замість концентрації плазмовою частотою ω_p , або енергією кванта плазмових коливань $\hbar\omega_p$, який називають *плазмоном*. Його енергія є мірою електростатичної взаємодії між додатними та від'ємними зарядами.

Цікаво, що енергія плазмонів у металах та високолегованих напівпровідниках має порядок $10^{-2} \dots 1$ еВ, що відповідає плазмовим частотам $10^{13} \dots 10^{15}$ Гц. Це означає, що плазмові коливання можуть збуджуватися електромагнітними коливаннями інфрачервоного та оптичного діапазону.

У плазмі можна спостерігати й більш складні явища, наприклад, поширення у просторі коливальних процесів, або плазмових хвиль. Плазмові коливання та хвилі є типовими проявами колективних явищ, обумовлених взаємодією між двома колективами заряджених часток.

Температура. Через зіткнення плазмові частки знаходяться у постійному хаотичному русі. Мірою середньої кінетичної енергії часток є температура. Оскільки частки, що утворюють плазму, можуть суттєво різнитися своїми параметрами, то й енергія їх руху може бути різною. Відповідно, для газорозрядної плазми розрізняють температуру електронів T_e , іонів T_i та нейтральних часток T , які у загальному випадку можуть бути нерівними одна одній (зазвичай $T < T_i < T_e$). Відмінність електронної та іонної температури викликана великою різницею мас електрона та іона. Для плазми твердого тіла використовують поняття електронної температури T_e , діркової температури T_p та температуру кристалічної ґратки T .

Слід зазначити, що поняття температури можна використовувати лише тоді, коли частки підпорядковуються класичній статистиці Максвелла–Больцмана. Тим не менш іноді формально температурою характеризують колективи, які знаходяться у нерівноважному стані, коли використання рівноважних функцій розподілу є умовним.

Якщо енергія плазмону набагато менша за енергію теплового руху ($\hbar\omega_p \ll k_B T$), квантові явища у плазмі не проявляються і вона підпорядковується законам класичної фізики. За великих концентрацій (наприклад, для електронного газу в металах) квант енергії плазмових коливань вимірюється

одинацями електрон-вольт, і у плазмі суттєвими є квантові ефекти.

Радіус екранування. Частки в плазмі взаємодіють між собою як заряди, напруженість поля яких згідно закону Кулона обернено пропорційна квадрату відстані, або потенціал обернено пропорційний відстані. Проте у плазмі кулонівська взаємодія проявляється не повною мірою, оскільки додатний заряд іона зміщує електрони, які знаходяться в його околі, й їх заряд екранує поле іона. Таке екранування «обрізає» електричні поля і викликає зменшення відстані, на якій ефективно взаємодіють заряджені частки.

Розрахунки для ефекту екранування показують, що поля зарядів будуть спадати швидше, ніж $1/r^2$, а потенціальна енергія електрона у полі екранованого іона змінюється за законом

$$U(r) \sim -\frac{1}{r} e^{-r/r_e}. \quad (2.3)$$

на відміну від закону $-1/r$ для ізольованого іона (рис. 2.3). Тут

$$r_e = \sqrt{\frac{\epsilon\epsilon_0 k_B T}{2e^2 n}} \quad (2.4)$$

величина, яку звать дебаєвим радіусом екранування і яка є одним з характерних параметрів плазми, що визначають ступінь електричної взаємодії між частками.

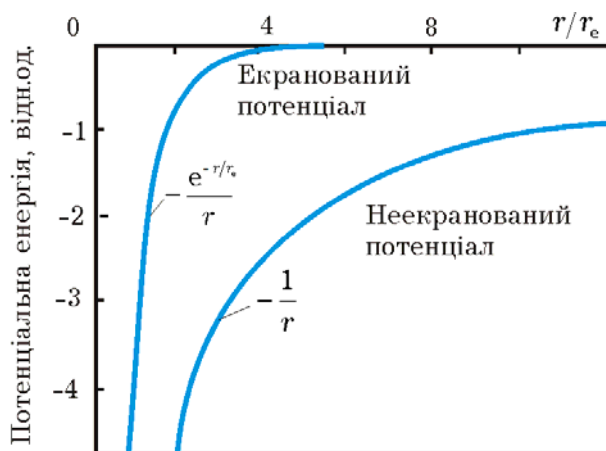


Рис. 2.3. Порівняння потенціальної енергії електрона в екранованому та неекранованому кулонівському полі додатного заряду

Довжина вільного пробігу. Попри те, що плазмові частки не вступають у безпосередній контакт, а їх взаємодія відбувається

на відстані за допомогою відносно далекодіючих електричних полів і не схожа на зіткнення нейтральних, через дуже короткий час цієї взаємодії (порядку $10^{-13} \dots 10^{-15}$ с) також використовують ті ж поняття, що і в кінетичній теорії газів: частота зіткнень, довжина вільного пробігу, ефективний переріз, тощо. За допомогою цих понять оцінюють ефективність співударянь.

Одним з найбільш простих і вживаних параметрів є *середня довжина* вільного пробігу між послідовними зіткненнями \bar{l} . Оскільки у плазмі можливі різноманітні елементарні процеси, які розглядають як різні види зіткнень, прийнято говорити про довжину вільного пробігу по відношенню до тієї чи іншої взаємодії \bar{l}_i . У такому разі середню довжину вільного пробігу знаходять із співвідношення

$$\frac{1}{\bar{l}} = \sum_i \frac{1}{\bar{l}_i}. \quad (2.5)$$

Вона визначається в основному найкоротшим пробігом, тобто тим видом взаємодії, який відбувається найбільш часто.

Ефективний переріз. Зручною величиною для кількісної міри ефективності взаємодії є *переріз зіткнення*, геометричний зміст якого легко сприймається на мові класичної моделі зіткнень нейтральних часток з очевидного (при незалежних одне від одного зіткненнях) твердження про те, що середня довжина вільного пробігу обернено пропорційна концентрації центрів зіткнення N_s :

$$\bar{l} = 1 / N_s S. \quad (2.6)$$

Тут коефіцієнт пропорційності S має розмірність площі, й для електричних часток його називають *ефективним перерізом* співударяння або зіткнення. Якщо довжину пробігу подати через середні значення швидкості та часу вільного пробігу $\bar{l} = \bar{v} \bar{\tau}$, то матимемо обернену пропорційність до *часу вільного пробігу* $\bar{\tau}$

$$S = 1 / N_s \bar{v} \bar{\tau}. \quad (2.7)$$

Іноді розглядають *макроскопічний ефективний переріз* $N_s S = 1 / \bar{l}$, який дорівнює сумі елементарних перерізів усіх

центрів в одиниці об'єму і показує середню кількість взаємодій певного виду на одиниці відстані.

Результатом хаотичного руху часток у рівноважній плазмі є відсутність спрямованого переносу зарядів. Він може з'явитися під дією зовнішнього електричного поля, за наявності градієнта концентрації, а також у результаті перепаду температур у різних місцях плазми. Зіткнення у цих процесах відіграють роль своєрідного «тертя», яке супроводжує спрямований рух зарядів.

Для аналізу процесів струмопереносу під дією електричного поля можна використовувати два статистичних підходи, які різняться способами усереднення параметрів руху. Перший з них передбачає спостереження за рухом одного електрона як завгодно довго у часі, а середня швидкість знаходиться усередненням за час спостереження (*одночасткова модель*). В іншому способі визначають параметри руху для всіх часток колективу у фіксований момент часу, а усереднення проводять за сукупністю (масивом) цих параметрів (*колективна модель*). Головним інструментом дослідження тут є функція розподілу, за якою і проводиться усереднення досліджуваних параметрів

Статистична теорія говорить, що ці способи зазвичай дають однаковий результат. Однак треба мати на увазі, що коректного результату за першого способу можна отримати лише, якщо зберігати умови спостереження незмінними, тобто для *статичного* поля. Друга модель допускає усереднювати масиви, отримані за різних значень електричного поля, тобто функція розподілу може змінюватись у часі, і за її допомогою можна досліджувати *динамічні* процеси

2.2. ДРЕЙФ НОСІЇВ. ОДНОЧАСТКОВА МОДЕЛЬ

Як уже відзначалось, розглянуті у попередньому підрозділі процеси показують лише вплив електричного поля на зміну електронних станів. Для розуміння закономірностей протікання електричного струму у твердому тілі необхідно враховувати ще одне важливе явище, а саме, зіткнення електронів та дірок з дефектами кристалічної ґратки. Ці зіткнення є випадковим процесом, вони хаотизують рух носіїв, після кожного зіткнення

електрони «забувають» свій попередній стан і прискорюються на вільному пробігу до наступного зіткнення.

Під дією електричного поля носії заряду рухаються прискорено і тому відповідно до (1.43) і (1.46) в ідеальному кристалі струм повинен з часом зростати лінійно (за постійної концентрації носіїв). Але насправді після включення поля електричний струм досить швидко встановлюється постійним. Причина цього явища пояснюється зіткненнями електронів і дірок з дефектами кристалічної ґратки, які хаотизують рух. Якщо електричне поле відсутнє (рис. 2.4, а), електрони між зіткненнями рухаються прямолінійно (верхній малюнок), а їх значення швидкості або хвильового вектору постійне (нижній). У момент зіткнення напрям і швидкість їх руху випадково змінюються і точки на нижньому графіку, що позначають хвильовий вектор електронів, «перескакують» у довільні місця. Випадковий характер зіткнень призводить до того, що у рівноважному стані швидкості електронів спрямовані з рівною ймовірністю на всі боки, тому сумарний хвильовий вектор і відповідно середня швидкість електронів дорівнюватимуть нулю, а якийсь переважаючий напрям їх руху буде відсутнім.

Наявність електричного поля (рис. 2.4, б) істотно змінює ситуацію. Як і раніше, напрям швидкості носіїв безпосередньо після кожного зіткнення буде випадковим, але під час вільного пробігу між двома послідовними зіткненнями вони відчуватимуть дію електричного поля і їх траєкторії будуть параболічними. Електрони прискорюватимуться у напрямі сили, під час вільного пробігу їх швидкість у x -напрямку рівномірно зростатиме і точки на нижньому малюнку будуть поступово зміщуватися у бік більших значень $k_x \sim v_x$. Під час зіткнень вони знову набуватимуть випадкових значень у k -просторі. Така поведінка носіїв означає, що на їх хаотичний рух накладатиметься переносний, або *дрейфовий рух*, що є результатом дії електричного поля.

Для кількісного аналізу дрейфу скористаємося простою класичною моделлю, яка ґрунтується на «індивідуальному» розгляді поведінки окремої частки. Розглянемо електрон, який, рухаючись хаотично, за час вільного пробігу τ між двома послідовними зіткненнями в електричному полі з напруженістю \mathcal{E} набуває спрямованого імпульсу $e\mathcal{E}\tau$. Цей спрямований проти поля

імпульс відповідає спрямованій швидкості v , яка дорівнює $e\mathcal{E}\tau/m$.
Усреднюючи її в часі за всією траєкторією спостережуваної частки, отримаємо вираз для дрейфової швидкості:

$$v_{\text{др}} = \frac{e\bar{\tau}}{m} \mathcal{E} = \frac{e}{m} \frac{\bar{l}}{\bar{v}} \mathcal{E}. \quad (2.8)$$

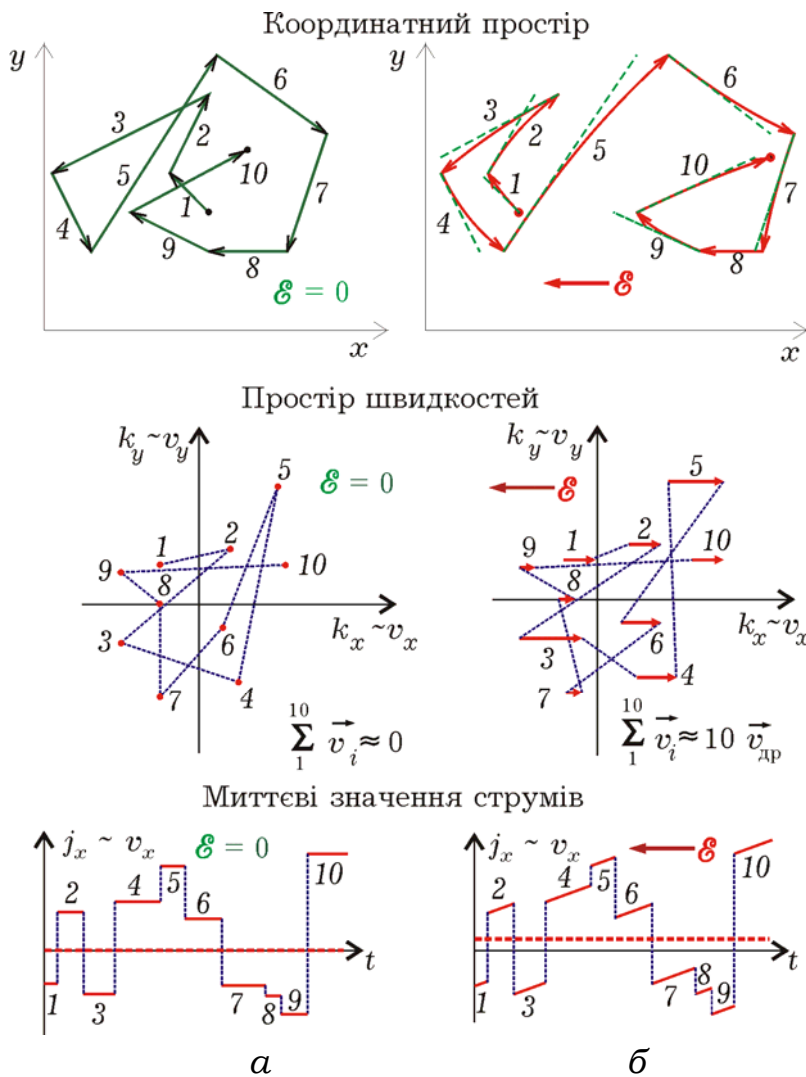


Рис. 2.4. Вплив зіткнень на приблизний вигляд траєкторій електрона і відповідні зміни його швидкості (хвильового вектора) за відсутності електричного поля (а) і в електричному полі для «таких же» зіткнень (б). Числа відповідають номерам вільних пробігів, під час яких створюється електричний струм, середнє значення (штрихова лінія на нижньому рисунку) якого не дорівнює нулю за наявності електричного поля

Використання співвідношення $\bar{\tau} = \bar{l} / \bar{v}$, де \bar{l} - середня довжина вільного пробігу, а \bar{v} - середня швидкість руху між

зіткненнями, строго кажучи, є невірним. Правильно $\bar{\tau} = (\overline{l/v})$, тобто середнє значення відношення двох величин не дорівнює відношенню їх середніх значень. Але для отримання простих основних закономірностей ця похибка не є істотною.

Важливим наслідком (2.8) є факт пропорційності між дрейфовою швидкістю і напруженістю електричного поля. Коефіцієнт пропорційності

$$\mu = e\bar{\tau} / m = e\bar{l} / m\bar{v} \quad (2.9)$$

називається *рухливістю* і має сенс дрейфовій швидкості у полі одиничної напруженості. Розмірність рухливості – м²/В·с, хоча в довідниках її значення частіше наводять в см²/В·с.

Результатом спрямованого дрейфового руху електронів є дрейфовий струм, густина якого пропорційна спрямованому потоку електронів $nv_{др}$ та їх заряду

$$j_{др} = env_{др} = en\mu\mathcal{E}. \quad (2.10)$$

Цей вираз є *законом Ома в диференціальній формі*, а коефіцієнт пропорційності між $j_{др}$ і \mathcal{E} називається *питомою електропровідністю*

$$\sigma = en\mu = e^2 n \bar{\tau} / m \approx e^2 n \bar{l} / m\bar{v}. \quad (2.11)$$

Для напівпровідників треба приймати до уваги дві обставини. По-перше, наявність різних носіїв заряду. Повний струм складається з внесків електронної і діркової складових, тому

$$\sigma = \sigma_n + \sigma_p = en\mu_n + ep\mu_p. \quad (2.12)$$

По-друге, використовуючи метод ефективної маси, ми повинні замінити в (2.11) звичайну масу на ефективну. Але у непрямозонних напівпровідниках, таких як кремній, германій, електрони у зоні провідності займають стани декількох еквівалентних долин, як це показано на рис. 2.5 для кремнію (дві долини, які лежать у напрямі $k_{[010]}$ і зображені на рис. 1.23, не потрапили в площину рисунка). Вклад електронів різних долин в електропровідність не однаковий, оскільки їх ефективна маса у напрямі дрейфу різна і відповідає для двох долин подовжній (m_l^*), а для чотирьох - поперечній (m_t^*) ефективній масі.

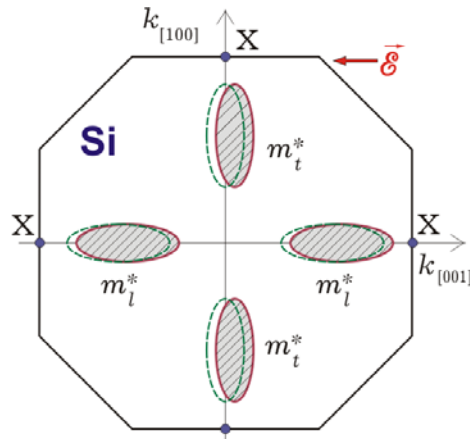


Рис. 2.5. Заповнення станів електронами в Δ -долинах зони провідності кремнію. У разі їх зміщення по станах під дією електричного поля електрони різних долин мають різну ефективну масу

Проте в цілому електропровідність кристала буде ізотропною, а рухливість електронів залежатиме від деякої середньої ефективної маси, яка називається *омічною* і знаходиться зі співвідношення

$$\frac{1}{m_c^*} = \frac{1}{3} \left(\frac{2}{m_t^*} + \frac{1}{m_l^*} \right), \quad (2.13)$$

яке витікає з правила підсумовування вкладів електронів різних долин у загальну електропровідність:

$$\sigma_n = \sum_{i=1}^6 \sigma_{ni} = \sum_{i=1}^6 e^2 \frac{n \bar{\tau}}{6m_i^*} = e^2 n \bar{\tau} / m_c^*. \quad (2.14)$$

Схожа ситуація має місце у валентній зоні: різні гілки дисперсійної кривої поблизу верхнього краю валентної зони відповідають «легким» і «важким» діркам. Використання деякої усередненої величини, званою *омічною* ефективною масою дірок, дає можливість, як і для електронів, оцінювати діркову електропровідність за простою формулою $\sigma_p = e^2 n \tau_p / m_p^*$, подібною до (2.11) для «одночасткової» моделі.

З наведених під час виведення формули (2.8) міркувань видно, що значення рухливості визначається середнім значенням часу

вільного пробігу між зіткненнями. У ідеальній кристалічній ґратці електронні стани описуються хвилями Блоха, що мають у середньому постійну амплітуду уздовж усього кристала. Це означає, що ідеальні ґратки не викликають зіткнень і довжина вільного пробігу носіїв необмежена, а зіткнення викликають лише *дефекти* кристала.

2.3. ДРЕЙФ НОСІЇВ. КОЛЕКТИВНА МОДЕЛЬ

Модель дрейфового руху носіїв під дією зовнішнього електричного поля, яка ґрунтується на усередненні в часі параметрів руху однієї частки і в цілому правильно описує основні закономірності дрейфу в статичному полі, у принципі не може бути придатною для аналізу динамічних процесів. Крім того, провідність розраховують зазвичай з використанням рівноважних значень концентрації носіїв, що можна наближено робити для процесів, які лише незначно порушують термодинамічну рівновагу.

2.3.1. Кінетичне рівняння

Найточніші кількісні моделі електронних процесів за наявності зовнішніх впливів, у тому числі й динамічних, можна створити на основі так званого *кінетичного рівняння*. У загальному випадку воно являє собою інтегро-диференціальне рівняння, за допомогою якого можна знайти функцію розподілу часток, а отже, і всі *середні* величини, що характеризують поведінку колективу часток.

Статистичний стан електронного колективу відображає функція $N(\vec{r}, \vec{k}, t)$, яку можна подати у вигляді добутку функції густини \vec{k} -станів $g(k) = dn_c/dV_k$ і функції $f(E(\vec{k})) = dn/dn_c$, що характеризує ймовірність заповнення станів, причому зв'язок $E(\vec{k})$ визначається законом дисперсії для електронів у певній речовині. Функцію густини станів можна обрати у вигляді (1.24), якщо залишаються в силі основні висновки зонної теорії. Функцію $f(E(\vec{k}))$ також найчастіше беруть у вигляді, подібному до рівноважної функції Фермі – Дірака. У результаті

динамічного впливу відбувається зміна \vec{k} -станів електронів, а їх енергія стає функцією часу – $E(\vec{k}(t))$.

У підсумку поведінку електронного колективу можна описати за допомогою функції розподілу

$$N(\vec{r}, \vec{k}, t) = \frac{d^2 N}{dV_r dV_k} = g(k)f(E(k)). \quad (2.15)$$

Цю функцію нормують таким чином, що її інтегрування за всім геометричним об'ємом V і об'ємом k -простору V_k , дає загальну кількість часток певного виду в певному колективі:

$$\int_{V_r} dV_r \int_{V_k} N(\vec{r}, \vec{k}, t) dV_k = N. \quad (2.16)$$

Надалі обмежимося розглядом одиничного геометричного об'єму, обравши функцію густини станів у вигляді $g(k) = 1/4\pi^3$.

Знання функції розподілу дозволяє знайти основні макроскопічні характеристики колективу. Наприклад, інтегрування її за всім k -простором дає розподіл концентрації часток у геометричному просторі:

$$n(\vec{r}, t) = \int_{V_k} N(\vec{r}, \vec{k}, t) dV_k, \quad (2.17)$$

і, відповідно, розподіл густини об'ємного заряду $\rho(\vec{r}, t) = -en(\vec{r}, t)$.

Усередненням швидкості часток у k -просторі за функцією розподілу можна знайти середню дрейфову швидкість їх спрямованого руху:

$$\vec{v}_{\text{др}}(\vec{r}, t) = \frac{1}{n(\vec{r}, t)} \int_{V_k} \vec{v} N(\vec{r}, \vec{k}, t) dV_k, \quad (2.18)$$

а отже, і розподіл густини струму:

$$\vec{j}(\vec{r}, t) = -en(\vec{r}, t) \vec{v}_{\text{др}}(\vec{r}, t) = -e \int_{V_k} \vec{v} N(\vec{r}, \vec{k}, t) dV_k. \quad (2.19)$$

Подібним чином знаходять локальне середнє значення будь-якої фізичної величини. Тут слід звернути увагу на вибір меж інтегрування. Для електронів у періодичному полі кристалічної ґратки областю неповторюваних значень *квазіхвильового* вектора є зона Бріллюена V_B , причому інтегрувати треба по станах, що відповідають одній енергетичній зоні. Більше того, іноді інтегрують по якій-небудь характерній ділянці дисперсійної кривої, визначаючи концентрацію або середні значення для часток того чи іншого виду. Наприклад, можна знайти концентрацію й параметри електронів окремо у Γ -, L- або X-долині зони провідності, «легких» або «важких» дірок у валентній зоні.

Основою для формулювання *кінетичного рівняння* є твердження, що в замкнутій системі кількість часток зберігається незмінною, тобто, $dN(\vec{r}, \vec{k}, t)/dt = 0$. Розкриваючи тут повну похідну через частинні та з огляду на те, що $d\vec{r}/dt = \vec{v}$, а $d\vec{k}/dt = \vec{F}/\hbar$, отримаємо

$$\frac{\partial N}{\partial t} + \left(\vec{v} \cdot \frac{\partial N}{\partial \vec{r}} \right) + \frac{1}{\hbar} \left(\vec{F} \cdot \frac{\partial N}{\partial \vec{k}} \right) = 0. \quad (2.20)$$

Якщо сили, що діють на електрони, розділити на дві групи: по-перше, сили з боку зовнішніх полів $\vec{F} = -e \left(\vec{\mathcal{E}}(\vec{r}, t) + [\vec{v} \times \vec{B}(\vec{r}, t)] \right)$, які змінюються у просторі порівняно плавно, та, по-друге, сили, які розглядаються як *зіткнення* або *співударяння*. Залишаючи в третьому доданку (2.20) лише силу зовнішніх полів, можна записати

$$\frac{\partial N}{\partial t} + \vec{v} \cdot \frac{\partial N}{\partial \vec{r}} + \frac{\vec{F}}{\hbar} \cdot \frac{\partial N}{\partial \vec{k}} = \left[\frac{\partial N}{\partial t} \right]_c, \quad (2.21)$$

де права частина, називана *інтегралом зіткнень*, символічно враховує їх вплив на функцію розподілу. Оскільки густина станів постійна в k -просторі, то, розділивши всі доданки на $g(k)$, одержимо найпоширенішу форму запису кінетичного рівняння для функції $f(\vec{r}, \vec{k}, t)$

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{r}} + \frac{\vec{F}}{\hbar} \cdot \frac{\partial f}{\partial \vec{k}} = \left[\frac{\partial f}{\partial t} \right]_c. \quad (2.22)$$

Якщо подумки другий і третій доданок перенести у праву частину, то одержимо більш наочну форму запису кінетичного рівняння з огляду на його фізичний зміст: зміна функції розподілу в даній точці можлива за рахунок просторової неоднорідності розподілу (наявність градієнтів концентрації, температури), впливу зовнішніх полів і зіткнень. Кінетичне рівняння є потужним засобом *статистичного опису* поведінки великих колективів часток, тому що одержувана в результаті його розв'язку функція розподілу дозволяє досить детально аналізувати фізичні процеси переносу носіїв в електронних приладах. За своєю природою кінетичне рівняння – класичне, але надаючи тим чи іншим величинам відповідний квантовий зміст, можна одержати так званий *напівкласичний* опис фізичних процесів з урахуванням квантових ефектів.

Права частина отримала таку назву через те, що швидкість зіткнень виражається через інтегральні форми, куди входять ймовірності співударянь різного виду, які в свою чергу знаходяться з квантово-механічного аналізу. Суттєвого спрощення кінетичного рівняння можна досягти, використовуючи *наближення часу релаксації*, згідно з яким права частина вважається пропорційною відхиленню функції розподілу від її рівноважного значення f_0 :

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{r}} + \frac{\vec{F}}{\hbar} \cdot \frac{\partial f}{\partial \vec{k}} = - \frac{f - f_0}{\tau}, \quad (2.23)$$

Коефіцієнт пропорційності τ , що стоїть у знаменнику, має розмірність часу й називається *часом релаксації*. За своїм фізичним змістом й величиною час релаксації близький до середнього часу вільного пробігу в одночастковій моделі дрейфу.

Статична провідність. Як приклад використання наближення часу релаксації розглянемо поведінку функції розподілу електронів в електричному полі. За його відсутності розподіл хвильових векторів і відповідно швидкостей електронів буде статистично рівномірним у всіх напрямках (рис. 2.6, а) і їхня сумарна швидкість дорівнюватиме нулю, а функція розподілу буде рівноважною й симетричною відносно початку координат.

Обчислення густини струму з (2.19) через усереднення швидкості по функції розподілу також дає нульовий результат.

За наявності електричного поля, спрямованого назустріч осі x , електрони прискорюються й всі точки у верхній частині рис. 2.6, б, в, що зображують *миттєві* значення швидкості електронів, будуть із часом рівномірно зміщуватися у напрямку дії сили поля. У момент зіткнення швидкість електронів *стає* випадковою й точки на малюнку «перескакують» у випадкові місця (приклад відповідного переходу позначений стрілкою). Такі процеси приводять до двох протилежних тенденцій: під дією електричного поля функція розподілу зміщується вправо, а зіткнення, хаотизуючи рух електронів, прагнуть повернути функцію розподілу до рівноважного стану. Очевидно, що в *стаціонарному стані* ці дві тенденції врівноважують одна одну за деякої зміщеної функції $N = N_0 + \Delta N$ й $f = f_0 + \Delta f$. Усереднення швидкості по (2.16) дає значення швидкості дрейфу, причому внесок у значення інтеграла дає тільки добавка $\Delta N = g(k) \cdot \Delta f$. Відповідно й величина густини струму визначається також нерівноважною добавкою

$$\vec{j}(\vec{r}) = -e \int \vec{v} \Delta N(\vec{r}, \vec{k}) dV_k = -e \int \vec{v} g(\vec{k}) \Delta f(\vec{r}, \vec{k}) dV_k. \quad (2.24)$$

Для знаходження нерівноважної частини функції розподілу скористаємося рівнянням (2.23), що за умови стаціонарності ($\partial f / \partial t = 0$), просторової однорідності поля ($\mathcal{E} = \text{const}$) і однорідності колективу ($\partial f / \partial \vec{r} = 0$), матиме вигляд

$$\frac{\vec{F}}{\hbar} \cdot \frac{\partial f}{\partial \vec{k}} = -\frac{\Delta f}{\tau}. \quad (2.25)$$

Звідси для $\vec{F} = -e\vec{\mathcal{E}}$ знаходиться значення нерівноважної частини функції розподілу в *стаціонарному стані*

$$\Delta f_{\text{ст}} = \frac{e\tau\vec{\mathcal{E}}}{\hbar} \cdot \frac{\partial f}{\partial \vec{k}}. \quad (2.26)$$

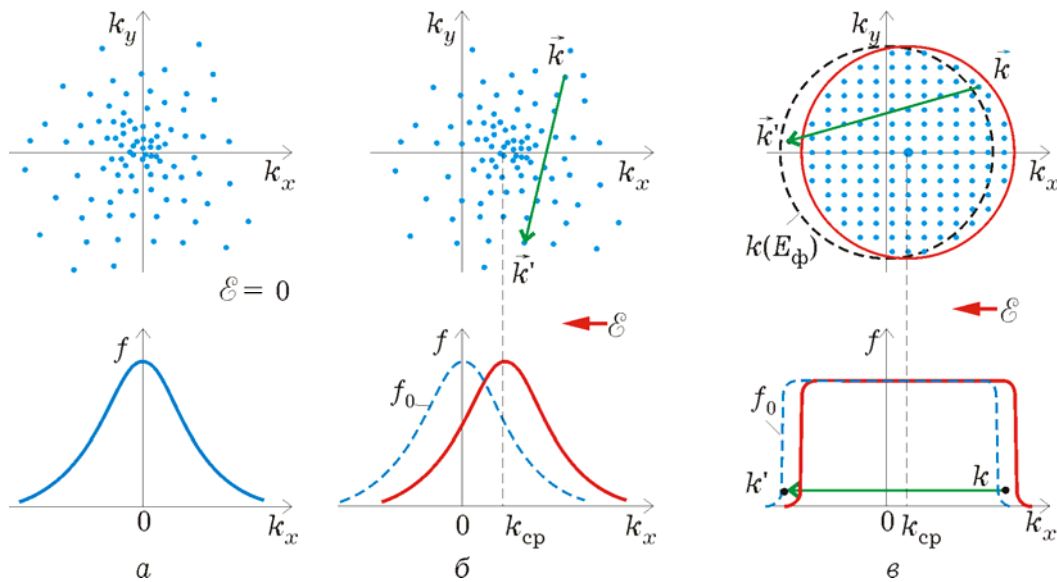


Рис. 2.6. За відсутності електричного поля (а) розподіл хвильових векторів (угорі) спричиняє симетричний характер функції розподілу. В електричному полі (б, в) зсув хвильових векторів у напрямку дії сили й розсіювання електронів у випадкові стани врівноважуються за деякої зміщеної функції розподілу. Треба звернути увагу на те, що у невиродженому колективі (б) розсіювання викликає перехід електрона (стрілка) практично в будь-який стан, а у виродженому (в) – лише в незаповнені стани поблизу енергії Фермі.

Підстановка (2.26) в (2.24), дає, що густина струму у слабкому електричному полі, як і очікувалося, пропорційна напруженості поля:

$$\vec{j} = -\frac{e^2 \vec{\mathcal{E}}}{\hbar} \int \tau \vec{v} g(\vec{k}) \frac{\partial f}{\partial \vec{k}} dV_k. \quad (2.27)$$

Залишаючи поза увагою деякі припущення щодо особливостей перетворення підінтегрального виразу, у підсумку для однодолинної моделі зонної структури отримаємо пряму пропорційність між густиною струму й напруженістю електричного поля, відому як закон Ома

$$\vec{j} = (e^2 n \tau / m^*) \vec{\mathcal{E}} = \sigma \vec{\mathcal{E}}. \quad (2.28)$$

Вираз для питомої електропровідності такий самий, як і для одночасткової моделі, але треба мати на увазі, що сюди входить суттєво інший термін – час релаксації, який, як побачимо далі, визначає інерційність колективних динамічних процесів.

За наявності декількох долин загальний струм складається із внесків електронів кожної долини зі своїми параметрами τ і m^* . Іноді вдається зберегти загальний вигляд виразу (2.28), використовуючи замість m^* деяку усереднену за різними долинами ефективну масу, яку називають *омічною*.

Нестационарна динаміка дрейфу. Омічна залежність між струмом і полем (2.28) отримана для стаціонарного стану. Зі зміною напруженості електричного поля спостерігаються перехідні процеси, які також можна досліджувати за допомогою кінетичного рівняння у формі (2.23). Приймаючи, як і раніше, умову просторової однорідності електричного поля і функції розподілу й заміняючи третій доданок у лівій частині (2.23) на $\Delta f_{\text{ст}}/\tau$ згідно (2.26), одержимо

$$\partial(\Delta f)/\partial t + \Delta f/\tau = \Delta f_{\text{ст}}/\tau. \quad (2.29)$$

Включення електричного поля в момент часу t_0^+ (рис. 2.7) викликає поступовий перехід з рівноважного стану ($\Delta f(t_0^+) = 0$) у стаціонарний стан $\Delta f(\infty) = \Delta f_{\text{ст}}$ відповідно до розв'язку (2.29):

$$\Delta f(t) = \Delta f_{\text{ст}}\{1 - \exp[-(t - t_0^+)/\tau]\}. \quad (2.30)$$

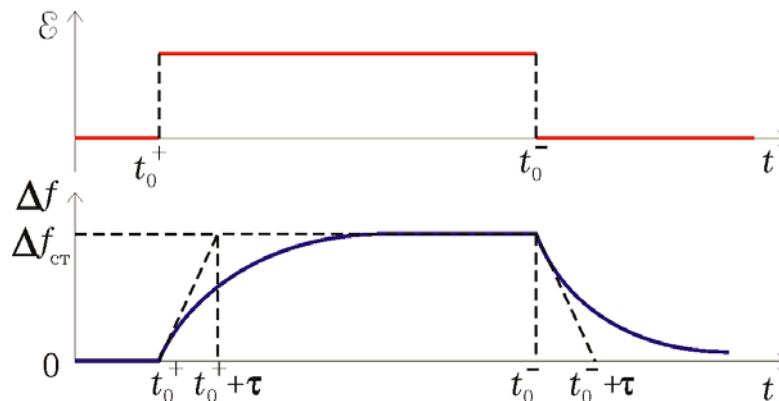


Рис. 2.7. Перехідні процеси зміни нерівноважної частини функції розподілу електронів і густини їхнього струму при включенні й вимиканні електричного поля. Похилі штрихові лінії – дотичні до кривих у моменти t_0^+ й t_0^- .

При вимиканні поля ($\mathcal{E} = 0$) кінетичне рівняння (2.23) для нерівноважної добавки Δf прийме вигляд

$$\partial(\Delta f)/\partial t + \Delta f/\tau = 0. \quad (2.31)$$

Якщо вимикання поля відбулося у стаціонарному стані (рис. 2.7) у момент t_0^- , тобто $\Delta f(t_0^-) = \Delta f_{\text{ст}}$, то розв'язком рівняння (2.31) буде функція

$$\Delta f(t) = \Delta f_{\text{ст}} \exp[-(t - t_0^-)/\tau], \quad (2.42)$$

яка показує, що після вимикання поля відбувається *релаксація*, тобто повернення функції розподілу до рівноважного стану.

Подібно до нерівноважної добавки до функції розподілу змінюються з часом дрейфова швидкість і густина струму. Значення струму, обумовлене законом Ома (2.28), досягається *не миттєво* після включення електричного поля, так само як і зникає після його вимикання. Інерційність цього процесу визначається часом, необхідним для здійснення зіткнень *всіма* частками, у результаті чого повинна відбутися така хаотизація станів, що відповідає рівноважній функції розподілу. Середнє значення цього часу й дорівнює часу релаксації τ . Формально, як це видно з рис. 2.7, за час τ закінчувалися б перехідні процеси, якби їхня швидкість залишалася такою ж, як і спочатку (див. штрихову лінію). У дійсності час установаження стаціонарного або рівноважного стану приймають рівним $(3..5)\tau$. Характерні значення часу релаксації порядку $10^{-12}..10^{-14}$ с приблизно рівні часу вільного пробігу в одночастковій моделі дрейфу.

Слід зазначити одну особливість релаксації функції розподілу, характерну для вироджених колективів (рис. 2.6, в). На відміну від неvirоджених, де ймовірність заповнення станів дуже мала й можливі переходи (рис. 2.6, б) практично в будь-які стани, у вироджених колективах усі стани всередині ізоенергетичної поверхні Фермі заповнені й переходи туди заборонені принципом Паулі. Тому при релаксації розподілу в результаті зіткнень будуть переважати переходи зі станів \vec{k} , близьких до $\vec{k}(E_{\text{ф}})$, де електрони прискорюються електричним полем, у стани \vec{k}' , також близькі до $\vec{k}(E_{\text{ф}})$, але такі, де електрони

гальмуються (рис. 2.6, в). Тому в релаксації беруть участь практично лише електрони в станах поблизу поверхні Фермі й відповідно у виразі (2.28) треба використовувати час релаксації $\tau(E_F)$ саме для цієї групи електронів.

Високочастотна провідність. Якщо електричне поле, що діє на електрони, змінюється за законом $\mathcal{E}(t) = \mathcal{E}_m \exp(i\omega t)$, то у функції розподілу від часу за таким саме законом буде залежати тільки нерівноважна добавка $f(t) = f_0 + \Delta f \exp(i\omega t)$. Природно, що це твердження можна прийняти для слабкого електричного поля, коли $\Delta f \ll f_0$. За тих саме припущень про просторову однорідність поля й колективу часток, що й у статичному полі, кінетичне рівняння приймає вигляд

$$i\omega \Delta f - \frac{e\vec{\mathcal{E}}_m}{\hbar} \frac{\partial f_0}{\partial \vec{k}} = -\frac{\Delta f}{\tau} \quad (2.43)$$

і може бути представлене у такій самій формі, що й (2.25), якщо замінити величину τ на $\tau_\omega = \tau/(1 + i\omega\tau)$. Тому й всі висновки, отримані для постійного поля, можуть бути використані для випадку змінного поля, якщо скористатися цією заміною.

Між струмом і напруженістю поля зберігається пропорційна залежність, однак коефіцієнт пропорційності, тобто питома електропровідність σ_ω , стає комплексною величиною і залежить від частоти за таким саме законом, що й час релаксації τ_ω ,

$$\sigma_\omega = \frac{\sigma}{1 + i\omega\tau} = \frac{\sigma(1 - i\omega\tau)}{1 + \omega^2\tau^2} = \sigma' - i\sigma'' \quad (2.44)$$

Комплексний характер величини σ_ω означає, що між змінами струму й поля є фазове запізнювання на величину $\arctg\omega\tau$. Цей фазовий зсув обумовлений, однак, не інерційністю, пов'язаною з масою часток, а статистичним характером їхніх зіткнень. Проявляється цей ефект на частотах, період коливань яких сумірний з часом релаксації ($\omega\tau$ порядку 1). За менших частот під час напівперіоду поля встигає встановлюватися значення нерівноважної добавки до функції розподілу, як і у статичному полі (2.26). На високих частотах, крім зазначеного фазового зсуву, за час прискорювального напівперіоду поля добавка Δf не встигає

досягти цього стаціонарного значення й величина густини струму зменшується відповідно зменшенню величини питомої електропровідності $|\sigma_\omega| = \sigma / \sqrt{1 + \omega^2 \tau^2}$.

Ще одна цікава особливість пов'язана із уявною частиною σ_ω . Звернемося до рівняння Максвелла $\text{rot } \vec{H} = \partial \vec{D} / \partial t + \vec{j}$, що у комплексній формі може бути представлено у вигляді $\text{rot } \vec{H} = (i\omega \epsilon \epsilon_0 + \sigma' - i\sigma'') \vec{E}$. Звідси легко бачити, що уявна частина електропровідності σ'' грає таку ж роль, як і діелектрична проникність, що може помітно позначатися на оптичних властивостях речовини. Однак, якщо перший доданок являє собою струм зміщення, що має ємнісний характер, то третій доданок, викликаний інерційністю процесів релаксації й пропорційний σ'' , має індуктивний характер.

2.3.2. Релаксаційні рівняння. Часи релаксації

Поряд з кінетичним рівнянням існує ще один зручний і досить потужний інструмент для аналізу колективних ефектів – релаксаційні рівняння. Справа в тому, що кінетичне рівняння дозволяє створювати найбільш точні і детальні моделі електронних процесів, але його розв'язок як інтегро-диференціальне рівняння може викликати значні труднощі. Наближення часу релаксації, істотно спрощуючи кінетичне рівняння, полегшує знаходження функції розподілу, а значить і всіляких середніх статистичних значень та функцій. Але одержувана звідси інформація часто буває надлишковою, якщо потрібно знати лише деякі усереднені величини або їхні просторові або часові розподіли. Найбільш важливими характеристиками колективу зазвичай є концентрація, швидкість та енергія часток. Більш просто й економно визначати ці величини або їхні просторово-часові функції не через функцію розподілу, а з диференціальних рівнянь для цих величин. Такі рівняння можуть бути отримані безпосередньо з кінетичного рівняння, якщо провести *усереднення до його розв'язку*. Для цього всі доданки рівняння (1.12) помножають на відповідну усереднювану величину й інтегрують за областю k -простору, що нас цікавить. У результаті отримуємо диференціальні рівняння в частинних похідних для середніх величин, що залежать тільки

від координат і часу. У літературі такі рівняння зустрічаються у різних формах, що відрізняються використанням тих чи інших, часто досить тонких, припущень, пов'язаних в основному з вибором моделі (гідродинамічна, квазігідродинамічна й ін.) колективу часток. Почасти із цієї причини такі рівняння називають *феноменологічними*, а за своєю суттю вони є рівняннями збереження, які є наслідками кінетичного рівняння.

Не зупиняючись на деталях таких перетворень, які можна знайти у науковій та навчальній літературі, наведемо лише остаточні результати для рівняння концентрації

$$\frac{dn}{dt} = -\frac{\Delta n}{\tau_n}, \quad (2.45)$$

сумарного значення імпульса

$$\frac{d(n\vec{p}_{cp})}{dt} - n\vec{F} = -\frac{n\vec{p}_{cp}}{\tau_p} \quad (2.46)$$

і сумарної енергії

$$\frac{d(nE_{cp})}{dt} - n\vec{F}\vec{v}_{cp} = -\frac{n(E_{cp} - E_0)}{\tau_E}. \quad (2.47)$$

Подібно до того, як у правій частині кінетичного рівняння стоїть нерівноважна добавка до функції розподілу $\Delta N = N - N_0 = g(k)\Delta f$, так і у чисельнику правих частин стоять відхилення від рівноважних значень для концентрації $\Delta n = n - n_0$, середнього імпульсу ($\vec{p}_{cp0} = 0$) та середньої енергії ($E_0 = 3k_B T/2$). Величини у знаменнику звать відповідно *часом релаксації концентрації* (τ_n), *імпульсу* (τ_p) або *енергії* (τ_E).

За умови збереження загальної кількості часток у зоні ($dn/dt = 0$) рівняння (2.46) може бути записано у формі рівняння збереження для середнього *дрейфового* значення імпульсу $\vec{p}(\vec{r}, t)$ (індекс середнього значення опускаємо)

$$d\vec{p}/dt - \vec{F} = -\vec{p}/\tau_p, \quad (2.48)$$

або для *дрейфової* швидкості

$$d\bar{v}/dt - \bar{F}/m^* = -\bar{v}/\tau_p, \quad (2.49)$$

яке має простий фізичний зміст: швидкість спрямованого руху збільшується під дією сили електричного поля, як і для вільних часток згідно закону Ньютона, і зменшується за рахунок зіткнень носіїв, або «сили тертя» $-\bar{p}/\tau_p$. У статичному електричному полі ці сили врівноважуються і

$$\bar{v} = -(e\tau_p/m^*)\bar{\mathcal{E}} = -\mu\bar{\mathcal{E}}, \quad (2.50)$$

тобто *рухливість визначається середнім часом релаксації імпульсу*.

Подібні міркування можна провести і у зв'язку з рівнянням збереження енергії. За незмінної загальної кількості часток рівняння (2.47) для *середньої* енергії у наближенні часу релаксації приймає вигляд (індекс середнього опускаємо)

$$dE/dt - \bar{F}\bar{v} = -(E - E_0)/\tau_E. \quad (2.51)$$

Треба мати на увазі, що ці рівняння отримані усередненням кінетичного рівняння за k -станами, тому тут змінні величини є функціями координат і часу, а у лівих частинах рівнянь стоять *повні похідні*, тобто $d/dt = \partial/\partial t + \bar{v}(\partial/\partial \bar{r})$.

Наведені рівняння дозволяють аналізувати фізичні процеси в плазмових колективах за відомих значень часів релаксації, які в свою чергу потребують теоретичного або експериментального визначення, про що буде мова далі. Попередньо відзначимо, що через деякі особливості зіткнень часи релаксації імпульсу та енергії можуть суттєво відрізнятися, причому зазвичай $\tau_E > \tau_p$.

Їх відмінність визначає деякі особливості перехідних процесів під час вмикання та вимикання електричного поля. Розв'язуючи рівняння збереження імпульсу (2.49) у припущенні постійності часу релаксації можна отримати, що із вмиканням електричного поля у момент часу t_0^+ (рис. 2.8, а) відбувається поступовий перехід електронного колективу з рівноважного стану ($v(t_0^+) = 0$) у стаціонарний ($v(\infty) = \mu\mathcal{E}$) за законом

$$v(t) = \mu \mathcal{E} \{1 - \exp[-(t - t_0^+)/\tau_p]\}. \quad (2.52)$$

Після вимкнення поля ($\mathcal{E} = 0$) в момент t_0^- відбувається релаксація із стаціонарного стану, яка відповідно до розв'язку рівняння (2.49) описується функцією

$$v(t) = \mu \mathcal{E} \exp[-(t - t_0^-)/\tau_p]\}. \quad (2.53)$$

Вона показує (рис. 2.8, б), що після вимкнення поля відбувається повернення в рівноважний стан. Інерційність цих перехідних процесів визначається часом релаксації імпульсу τ_p , який за значенням практично збігається з часом τ , що входить у кінетичне рівняння.

Інша ситуація спостерігається для енергії. Розв'язком рівняння (2.51) при вмиканні поля буде функція

$$E(t) = E_0 + e\tau_E \mu \mathcal{E}^2 \left[1 - \frac{\tau_E}{\tau_E - \tau_p} e^{-(t-t_0^+)/\tau_E} + \frac{\tau_p}{\tau_E - \tau_p} e^{-(t-t_0^+)/\tau_p} \right], \quad (2.54)$$

а при вимиканні

$$E(t) = E_0 + e\tau_E \mu \mathcal{E}^2 \exp[-(t - t_0^-)/\tau_E]. \quad (2.55)$$

Характерна особливість цих функцій (рис. 2.8, в) полягає в тому, що тривалість перехідних процесів визначається в основному часом τ_E , більшим, ніж τ_p . Деякий вплив релаксаційного процесу з часовою константою τ_p помітний після вимкнення поля і проявляється у повільнішому зростанні енергії на початку. Це відображає третій доданок у дужках (2.54), внесок якого зменшується із підсиленням нерівності $\tau_E > \tau_p$. Після вимкнення поля релаксація енергії до рівноважного значення E_0 відбувається за експоненціальним законом (2.55) з характерною часовою константою τ_E .

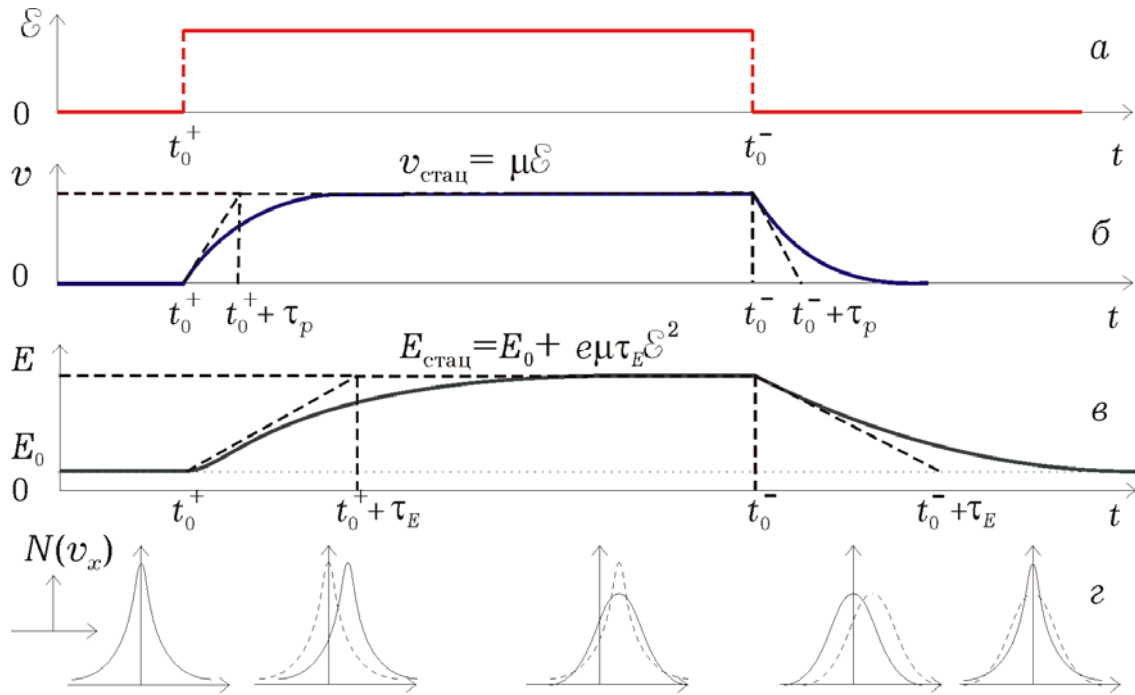


Рис. 2.8. Часові діаграми перехідних процесів під час вмикання та вимикання електричного поля (а). Встановлення стаціонарного стану із вмиканням і релаксація після вимкнення для швидкості (б) та енергії (в) відбувається із суттєво різними швидкостями. Унизу (г) наведено символічне зображення функції розподілу для відповідних моментів часу

Результат таких перехідних процесів – еволюція функції розподілу $N(v_x)$, схематично зображена на рис. 2.8, г. Із вмиканням поля вона за час порядку τ_p з рівноважного стану зміщується так, що середнє значення стаціонарної швидкості встановлюється рівним швидкості дрейфу $v = \mu\mathcal{E}$. Але повністю стаціонарною функція стає лише за час порядку $\tau_E > \tau_p$, коли відбувається її розширення, яке означає збільшення середньої енергії електрона. Після вимкнення поля швидше відбувається релаксація дрейфової швидкості до нульового значення і функція стає симетричною, залишаючись широкою. І лише через час порядку τ_E вона звужується, що означає зменшення енергії до рівноважного значення E_0 і повернення електронного колективу до рівноважного стану.

2.4. РОЗСІЮВАННЯ НОСІЇВ, РУХЛИВІСТЬ

Різноманітність і особливості ефектів переносу носіїв у твердих тілах залежать не тільки від конфігурації й величини зовнішніх полів, але й від фізичних особливостей конкретних механізмів співударянь, що приводять до розсіювання спрямованих потоків носіїв. Ці процеси вивчаються у розділі фізики твердого тіла, що називається *теорією розсіювання*. Тут розглянемо лише характерні особливості деяких найважливіших механізмів розсіювання та їх вплив на рухливість й інші властивості електронних колективів.

Перш за все слід пам'ятати, що поведінку електронів у *ідеальній* кристалічній ґратці описують хвильовими функціями у вигляді хвиль Блоха, які поширюються у кристалі як біжучі хвилі з постійною вздовж кристала амплітудою. Поява відбитих хвиль (тобто розсіювання електронів) можлива лише у разі порушення ідеальної періодичності поля кристала за наявності дефектів ґратки. Порушення періодичності можуть бути викликані двома основними причинами. Перша з них пов'язана із стаціонарними, локалізованими у просторі структурними дефектами – домішками, межами, різними дислокаціями. Друга викликана тепловими коливаннями іонів у вузлах кристалічної ґратки відносно положення рівноваги. Незважаючи на те, що ці коливання являють собою *хвильові* збудження періодичної структури ґратки, які залежать від часу і просторово не локалізовані, розсіювання електронів на них розглядають як зіткнення з деякими квазічастками – *фононами*, кількість яких може змінюватися внаслідок зіткнення. До таких дефектів належать також збудження в енергетичному спектрі електронів (екситони, плазмони), за певних умов помітним стає електрон-електронне та електрон-діркове розсіювання. Але найтипівіші дефекти – це домішкове розсіювання і різні види фононного розсіювання.

Метою теорії розсіювання є квантово-механічний аналіз ймовірностей переходу $w_{kk'}$ з початкового k -стану в стан k' за одиницю часу. Ефективність співударянь часто крім часу релаксації τ у кінетичному рівнянні оцінюють оберненою величиною, яка називається *швидкістю розсіювання* і має сенс середньої кількості актів розсіювання певного виду за одиницю

часу. Її знаходять безпосереднім інтегруванням ймовірності переходів за всіма кінцевими k' -станами

$$\lambda(k) = 1 / \tau(k) = \int w_{kk'} dV_{k'}. \quad (2.56)$$

Саме ці величини використовують для визначення ймовірності того чи іншого виду розсіювання. За наявності різних механізмів їх швидкості розсіювання підсумовують

$$\lambda_{\Sigma} = \sum_i \lambda_i.$$

Це правило виконується, якщо різні акти розсіювання відбуваються незалежно один від одного.

Усереднюючи швидкості розсіювання за функцією розподілу, яку часто обирають максвеллівською з електронною температурою, отримують часи релаксації, середні для всього колективу, для використання в релаксаційних рівняннях.

2.4.1. Фононні спектри

Найбільш визначальними для електронних властивостей напівпровідників, крім зонної структури, є особливості взаємодії електронів з тепловими коливаннями кристалічної решітки.

У багатьох питаннях, пов'язаних з тепловими властивостями твердих тіл, виправдовує себе модель Дебая, відповідно до якої теплове збудження викликає не коливання окремих атомів, випадкові та незалежні одне від одного, а *хвильові* процеси зміщення атомів біля положення рівноваги, які поширюються у кристалічній ґратці як у єдиному цілому. Хвильові коливання ґратки, як і всі інші, квантуються, а кванти цих хвиль – *фонони* – розглядають як квазічастки, що рухаються у кристалі зі швидкістю, яка відповідає груповій швидкості хвильових процесів і в першому наближенні дорівнює швидкості звуку.

Спектр частот фононів визначається періодичністю структури та параметрами ґратки й описується *законом дисперсії*. Як і для електронних хвиль у кристалі, цей закон можна подати як функцію $\omega(\vec{k}_\phi)$ або $E(\vec{k}_\phi)$, де \vec{k}_ϕ – хвильовий (точніше, квазіхвильовий) вектор фононів. Ця функція, як і для електронного закону дисперсії, є парною та періодичною з

періодом $2\pi/a$ (для одновимірного кристала), тому достатньо визначити її на половині зони Бріллюена.

Найпростіша модель поширення коливань в одновимірному ланцюжку зв'язаних пружними силами однакових атомів з періодом a , дає закон дисперсії у вигляді:

$$\omega(k_{\phi}) = \omega_m \sin(k_{\phi}a/2), \quad (2.57)$$

де $\omega_m = 2v_{зв}/a$ – гранична частота фононів (рис. 2.9, а), яка залежить від швидкості поширення звуку в твердому тілі й відповідає найкоротшій довжині хвилі $\lambda_{\min} = 2a$, за якої сусідні атоми коливаються у протифазі. Такі коливання утворюють стоячу хвилю, групова швидкість якої $v_{гр} = d\omega/dk_{\phi}$ дорівнює нулю. Іноді цю граничну частоту характеризують так званою *температурою Дебая* θ_D , яку визначають із рівності $\hbar\omega_m = k_B\theta_D$. Для типових напівпровідників період ґратки має значення порядку 10^{-10} м, швидкість звуку $\sim 10^3$ м/с, тому граничні частоти мають порядок $10^{12} \dots 10^{13}$ Гц і лежать в ультразвуковій області, а температури Дебая становлять кілька сотень кельвінів. Для найнижчих частот закон дисперсії є лінійним ($\omega \approx \omega_m k_{\phi} a/2$), тому фазова і групова швидкості однакові й дорівнюють швидкості звуку. Такі коливання і, відповідно, фонони називають *акустичними*.

Для кристалів, елементарна комірка яких складається з двох атомів, закон дисперсії (рис. 2.9, б, в) має свої особливості – у ньому з'являється так звана *оптична* гілка.

Зрозуміти її походження можна за допомогою відносно простої моделі (рис. 2.9, б). Для лінійного ланцюжка з однакових атомів, розміщених на відстані $a' = a/2$, справджується закон дисперсії (2.57), який, якщо замінити a' на a , дає значення ω_m у точках $k_{\phi} = \pm \pi/a' = \pm 2\pi/a$. Але цей інтервал удвічі перевищує розмір зони Бріллюена ($-\pi/a \leq k_{\phi} \leq \pi/a$) для кристала з періодом a . Тому після зведення закону дисперсії до першої зони Бріллюена, як це роблять для електронних хвиль, він набуває вигляду (рис. 2.9, в), де оптична гілка є продовженням акустичної, зміщеної на відстань $2\pi/a$.

Акустичні коливання поширюються у ланцюжку атомів як у єдиному цілому, а найбільш короткохвильові мають частоту

$\omega_m/\sqrt{2}$ (як у точці $k_\phi = \pi/2a$ на рис. 2.9, а), коли на одному періоді ґратки вкладається половина довжини хвилі і $\lambda = \lambda_{\min} = 2a$. Ця точка вироджена і відповідає також нижній частоті оптичної гілки, де зі збільшенням частоти збільшується і довжина хвилі, а коливання сусідніх атомів однієї комірки протифазні.

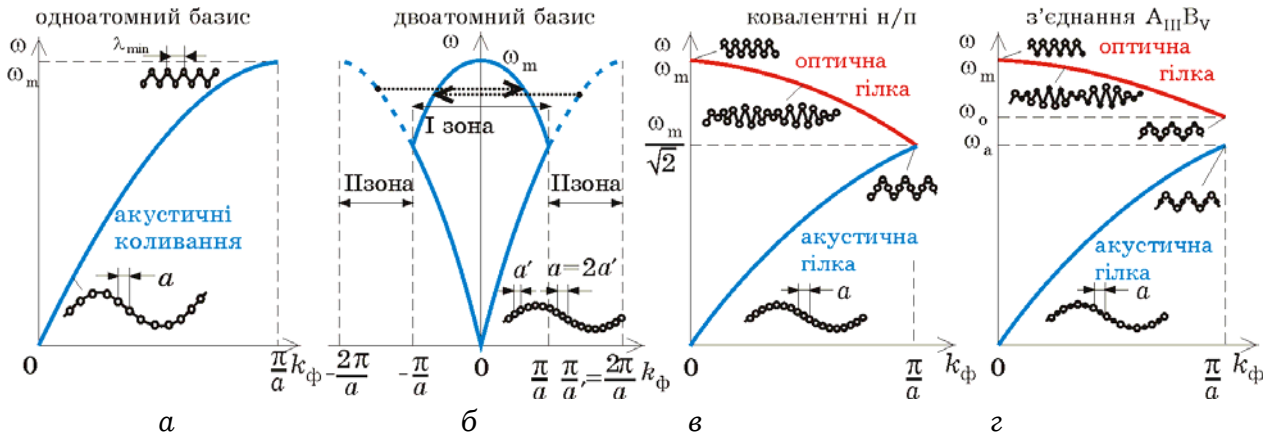


Рис. 2.9. Закони дисперсії хвиль зміщення в одновимірній моделі кристала – ланцюжку атомів: з одноатомним базисом (а), з базисом із двох однакових атомів (б, в) та двох різних атомів (г). Показано (б) утворення оптичної гілки для двоатомного базису зведенням верхньої частини «акустичної» гілки до першої зони Бріллюена (масштаб по горизонтальній осі зменшено). На фрагментах для зручності зображення поздовжні зміщення атомів подано як поперечні

У принципі таку саму картину спостерігають і в ланцюжку з періодом, що складається з двох різнойменних атомів або однойменних, але з різною жорсткістю зв'язків зліва і справа від атома. Через це в точці $k_\phi = \pi/a$ виродження знімається (рис. 2.9, г), бо акустичні й оптичні коливання мають різні частоти, а у спектрі фононів утворюється розрив. Ці найбільш короткохвильові фонони, як акустичні, так і оптичні, іноді називають *міждолинними*.

У розглянутій одновимірній моделі можливі лише поздовжні види (*моди*) коливань, хоча на наведених рисунках їх подано як поперечні. Це зроблено виключно для зручності зображення. У тривимірному кристалі існують поздовжні коливання й дві моди поперечних коливань, причому в кубічному кристалі обом поперечним модам відповідає однаковий закон дисперсії. На рис. 2.10 зображено дисперсійні криві для фононів у найпоширеніших напівпровідникових матеріалах: кремнії,

германії, арсеніді галію. Загальний вигляд цих кривих для поздовжніх мод (LA – акустична, LO – оптична) цілком відповідає наведеній вище моделі, тільки закон дисперсії описують функцією, складнішою за (2.57).

У кремнії та германії, елементарні комірки яких складаються з однакових атомів, гілки поздовжніх LA- та LO-мод вироджені на межі зони Бріллюена, а енергії міждолинних фононів дорівнюють $\sim 0,05$ еВ та $\sim 0,03$ еВ відповідно. Така помітна різниця пов'язана перш за все з різною масою атомів Si і Ge.

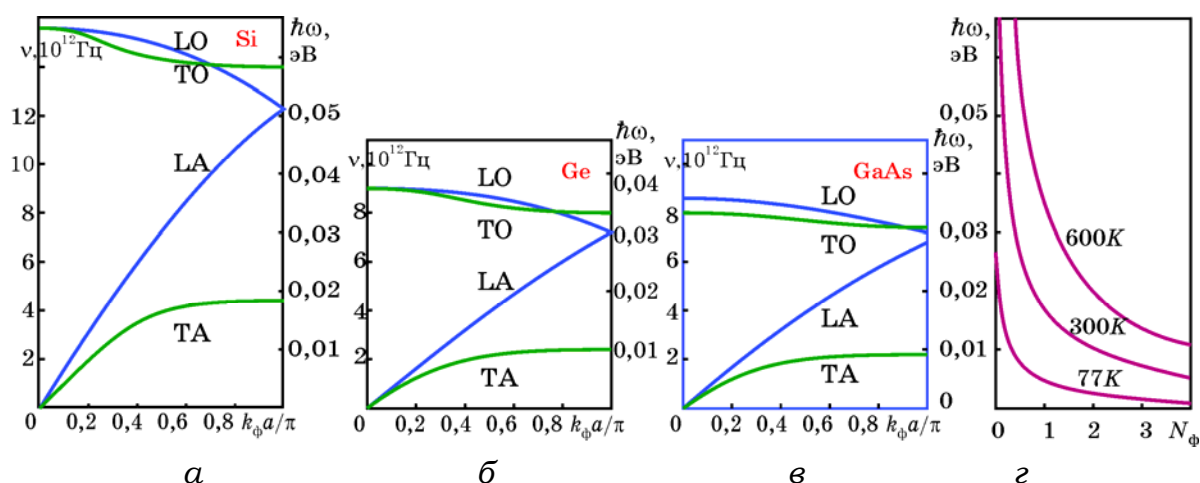


Рис. 2.10. Дисперсійні криві фононів для кремнію (а), германію (б), арсеніду галію (в) для кристалграфічного напрямку[100] (L – поздовжні, Т – поперечні, А – акустичні і О – оптичні моди). На правих шкалах відкладено енергії фононів, на лівих – їх частоти. Поруч (г) подано залежність середньої кількості фононів певної частоти від їх енергії за різних температур

В арсеніді галію виродження немає, але невелика різниця між масами атомів галію та арсену приводить і до невеликої розбіжності частот оптичної та акустичної мод: $\omega_o/\omega_a = \sqrt{M_{As}/M_{Ga}}$. В інших бінарних з'єднаннях А^{III}В^V через більшу різницю атомних мас підґраток це відношення може бути суттєво більшим. Так, наприклад, у нітриді галію гранична частота оптичних коливань перевищує частоту акустичних більш ніж вдвічі ($M_{Ga}/M_N \approx 5$).

На рис. 2.10. зображено також поперечні акустичні (ТА) та оптичні (ТО) гілки. У неполярних кристалах германію і кремнію поперечна та поздовжня оптичні моди вироджені для $k_\phi = 0$. У бінарних з'єднаннях, де оптичні коливання викликають утворення дипольного електричного моменту, ці моди мають різні частоти, а для $k_\phi = 0$ виконується співвідношення

$$\omega_{LO}/\omega_{TO} = \sqrt{\epsilon_{ст}/\epsilon_{\infty}}, \quad (2.58)$$

де $\epsilon_{ст}$ і ϵ_{∞} – відповідно статична і високочастотна діелектрична проникність.

Характерною особливістю фононів, крім спектра, є їх розподіл по енергії. Середню кількість фононів $N_{\phi}(\omega)$ із заданою енергією (частотою) визначають як для безспінових часток статистикою Бозе – Ейнштейна і обчислюють за формулою Планка:

$$N_{\phi}(\omega) = 1/[\exp(\hbar\omega/k_B T) - 1]. \quad (2.59)$$

Залежність кількості фононів від їх енергії (рис. 2.10, з) показує, що у разі $\hbar\omega > k_B T$ кількість фононів $N_{\phi} \ll 1$. Ця умова зазвичай добре виконується для оптичних фононів за кімнатних температур.

Для довгохвильових акустичних фононів з частотами $\hbar\omega \ll k_B T$ можна наближено вважати, що $N_{\phi}(\omega) \approx k_B T / \hbar\omega$, тобто їх кількість пропорційна температурі.

2.4.2. Основні закономірності

Під час взаємодії носія заряду з тепловими коливаннями кристалічної решітки його енергія може як збільшуватися, так і зменшуватися на величину кванта коливань решітки, тобто енергії фонона. У цьому випадку говорять про *поглинання*, коли енергія електрона збільшується за рахунок енергії коливань ґратки, або *збудження* фонона, коли енергія носіїв заряду підвищує енергію теплових коливань. Оскільки багатофононні процеси малоймовірні порівняно з однофононними, то закон збереження енергії для фононного розсіювання матиме вигляд:

$$E' - E = \pm \hbar\omega. \quad (2.60)$$

Крім того, має виконуватися закон збереження імпульсу (квазіімпульсу), який зручніше виражати через хвильові вектори:

$$\vec{k}' - \vec{k} = \pm \vec{k}_\phi. \quad (2.61)$$

У наведених формулах штрихами позначено стани електрона після розсіювання, а знаки « \pm » стосуються відповідно поглинання фонона і його збудження.

Квантовомеханічний аналіз процесів переходу електронів між станами дає вирази для ймовірностей розсіювання з поглинанням фонона електроном і з його збудженням

$$w_{kk'}^\pm = w(\vec{k}_\phi) (N_\phi + 1/2 \mp 1/2) \delta(E' - E \mp \hbar\omega), \quad (2.62)$$

де $w(\vec{k}_\phi)$ – деяка функція, що залежить від конкретного виду фононного розсіювання, δ -функція враховує закон збереження енергії, а верхній знак стосується поглинання фонона, а нижній – його збудження.

Пропорційність імовірності кількості фононів ($N_\phi + 1/2 \pm 1/2$) свідчить про те, що швидкість розсіювання зі збудженням фононів у $(N_\phi + 1)/N_\phi$ разів більша, ніж із їх поглинанням. Однак це правило виконується повною мірою, по-перше, за достатньо великих енергій електрона, бо якщо $E < \hbar\omega$, то розсіювання може відбутися тільки з поглинанням фонона. По-друге, ця відмінність суттєва за звичайних температур для оптичних фононів, коли $\hbar\omega > k_B T$ і $N_\phi \ll 1$. Для акустичних, особливо довгохвильових ($k_\phi \approx 0$) фононів $N_\phi \gg 1$ і швидкості розсіювання з поглинанням і збудженням практично однакові.

Важливою характеристикою процесу зіткнення є *пружність*. У кожному акті співударяння енергія електрона зменшується або збільшується на енергію фонона $\hbar\omega$. Однак у середньому за рахунок переважання процесів збудження енергія електронів зменшуватиметься, а кристалічна гратка нагріватиметься. Середня енергія, що її електрон втрачає за одне зіткнення, може бути визначена як енергія одного фонона, помножена на відношення різниці ймовірностей з випромінюванням і поглинанням до повної ймовірності, тобто їх суми:

$$\Delta E_1 = \hbar\omega \frac{w_{kk'}^- - w_{kk'}^+}{w_{kk'}^- + w_{kk'}^+}. \quad (2.63)$$

Відношення цієї величини до енергії електрона в середньому визначає частину переданої решітці енергії, тобто ступінь пружності фононних співударянь, і називається *коефіцієнтом пружності* ζ . Увівши поняття електронної температури за співвідношенням $E = 3k_B T_e / 2$, з урахуванням виразів (2.62) для ймовірностей розсіювання і (2.59) отримаємо:

$$\zeta = \frac{\Delta E_1}{E} = \frac{\hbar\omega}{E} \frac{1}{2N_\phi + 1} = \frac{2\hbar\omega}{3k_B T_e} \operatorname{th} \frac{\hbar\omega}{2k_B T}. \quad (2.64)$$

За високих температур або малих енергій фононів, коли $\hbar\omega \ll k_B T$ (якщо замінити $\operatorname{th} x \approx x$), коефіцієнт пружності

$$\zeta \approx (\hbar\omega)^2 / 3(k_B)^2 T T_e \ll 1, \quad (2.65)$$

і такі співударяння можна вважати практично пружними.

Для оптичних фононів енергія може бути співмірною з величиною $k_B T$, тому їх взаємодія з електронами є суттєво непружною. Однак зі збільшенням енергії електронів або електронної температури пружність співударянь зростає так, що коефіцієнт пружності зменшується обернено пропорційно T_e .

2.4.3. Види розсіювання

Акустичне розсіювання. У процесі акустичного розсіювання основну роль відіграє взаємодія електрона з поздовжніми коливаннями ґратки, які викликають зміни міжатомних відстаней (рис. 1.19). Хвильові процеси зміни параметра ґратки a приводять до відповідної зміни ширини забороненої зони. Оскільки довжина хвилі акустичних коливань, особливо для малих значень k_ϕ , а саме вони найбільш ефективно взаємодіють з електронами, помітно більша за міжатомну відстань, то тут можна використати макроскопічну модель викривлення енергетичних зон (рис. 2.11). У цьому випадку говорять про локальну зміну енергії електрона в зоні провідності, пропорційну поздовжній деформації кристала, тобто відносній зміні міжатомної відстані.

$$\Delta E = D_a (\Delta a/a). \quad (2.66)$$

Коефіцієнт пропорційності D_a (розмірність еВ) називають *константою акустичного деформаційного потенціалу*, а саме розсіювання у науковій літературі – розсіюванням на акустичному деформаційному потенціалі. Воно пов'язане з виникненням локальних електричних полів, що з'являються за рахунок зміни взаємодії між атомами у результаті поздовжніх деформацій решітки і викликають зміни траєкторії руху електрона.

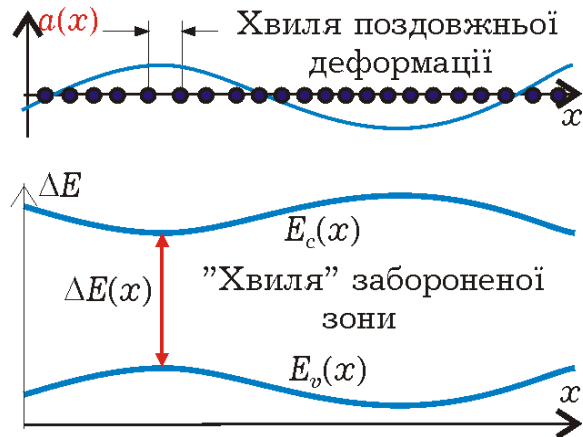


Рис.2.11. Зміна міжатомної відстані викликає зміну ширини забороненої зони і відповідно енергії електрона

Поперечні акустичні фонони пов'язані з деформаціями зсуву атомних шарів і практично не змінюють міжатомних відстаней, але можуть викликати п'єзоелектричний ефект. Однак *п'єзоелектричне розсіювання* у типових випадках набагато слабше за розсіювання на поздовжніх акустичних фононах.

Із сумісного розв'язку рівнянь балансу енергії та імпульсу слідує, що у розсіюванні беруть участь переважно довгохвильові акустичні фонони, енергія яких помітно менша за енергію електронів, що дорівнює тепловій – $3k_B T/2$. Тому поглинання або випромінювання фонона майже не змінює енергію електрона і *акустичне розсіювання можна вважати практично пружним* ($\zeta \ll 1$).

Для дрейфових процесів переносу носіїв часто вважають, що функція розподілу фононів (2.59) залишається рівноважною, а за умови $\hbar\omega \ll k_B T$ кількість фононів $N_\phi + 1 \approx N_\phi \approx k_B T / \hbar\omega \gg 1$ пропорційна температурі. Цей факт свідчить, що ймовірності розсіювання з випромінюванням і поглинанням фонона за цих умов практично однакові. Це дозволяє відносно просто обчислити інтеграл (2.56) з урахуванням (2.62) і отримати вираз для швидкості акустичного розсіювання:

$$\lambda_a = \frac{(2m^*)^{3/2} D_a^2}{2\pi\rho v_{зв}^2 \hbar^4} k_B T \sqrt{E}, \quad (2.67)$$

де m^* – ефективна маса густини станів; D_a – деформаційний акустичний потенціал; $v_{зв}$ – швидкість звуку, що відповідає нахилу акустичної гілки поздовжніх коливань для $k_{\phi} \approx 0$; ρ – густина речовини.

Характерно, що час релаксації $\tau = 1/\lambda$ зменшується зі зростанням енергії електрона обернено пропорційно $E^{1/2}$ або його швидкості. У вироджених електронних колективах, наприклад, у металах, енергія електронів визначається енергією Фермі (див. розд. 3) і практично не залежить від температури, тому питома електропровідність $\sigma = en\mu$ спадає зі зростанням температури як $1/T$ через зменшення рухливості $\mu = e\tau/m^*$. У невироджених напівпровідниках у слабких електричних полях в енергії електрона переважає внесок теплового руху, $E = 3k_B T/2$ і рухливість змінюється як $\sim T^{-3/2}$.

Оптичне розсіювання. Для оптичних коливань кристалічної решітки характерний пологий хід закону дисперсії (див. рис. 2.10), тому частоту оптичних фононів ω_0 часто беруть незалежною від хвильового вектора і рівною $\hbar\omega_0 = k_B\theta_D$. Їхня енергія співмірна з тепловою енергією носіїв, тому розсіювання на оптичних фононах вважають *непружним*. Як випливає з (2.60), поглинання фонона можливе навіть за найнижчих енергій (температур), а для збудження фонона енергія електрона має бути не нижчою за його енергію: $E \geq \hbar\omega_0$. З іншого боку, пропорційність імовірності розсіювання кількості фононів $(N_{\phi} + 1/2 \pm 1/2)$ приводить до різних швидкостей розсіювання: імовірність збудження може бути суттєво більшою за ймовірність поглинання, якщо $N_{\phi} \ll 1$ (звичайно, якщо енергії електрона достатньо для випромінювання фонона).

Взаємодія з оптичними фононами в атомарних напівпровідниках із суто ковалентними (гомеополлярними) зв'язками і в бінарних з'єднаннях з помітною часткою полярних (гетерополярних) зв'язків має різні механізми.

В атомарних напівпровідниках протифазні поздовжні коливання атомів двох підґраток, характерні для оптичної гілки, не приводять до зміни ширини забороненої зони, як для

акустичних коливань, але змінюють потенціальну енергію взаємодії між сусідніми атомами. За цих умов змінюється й енергія електрона, причому пропорційно зміщенню: $\Delta E = D_0 \Delta x$. Коефіцієнт пропорційності D_0 у цьому випадку має розмірність еВ/см і називається *константою оптичного деформаційного потенціалу*, а саме розсіювання називають розсіюванням на оптичному деформаційному потенціалі або *неполярним оптичним розсіюванням*. Швидкість такого розсіювання розраховують за формулою

$$\lambda_{\circ}^{\pm} = \frac{(m^*)^{3/2} D_0^2}{\sqrt{2\pi} \rho \hbar^3 \omega_0} (N_{\phi} + 1/2 \mp 1/2) \operatorname{Re} \sqrt{E \pm \hbar \omega_0}. \quad (2.69)$$

Тут, як і раніше, верхній знак стосується поглинання фонона, а нижній – його збудження. Оператор Re дає нульову швидкість збудження, якщо $E \leq \hbar \omega_0$, а в разі поглинання фонона він втрачає сенс (рис. 2.12, а).

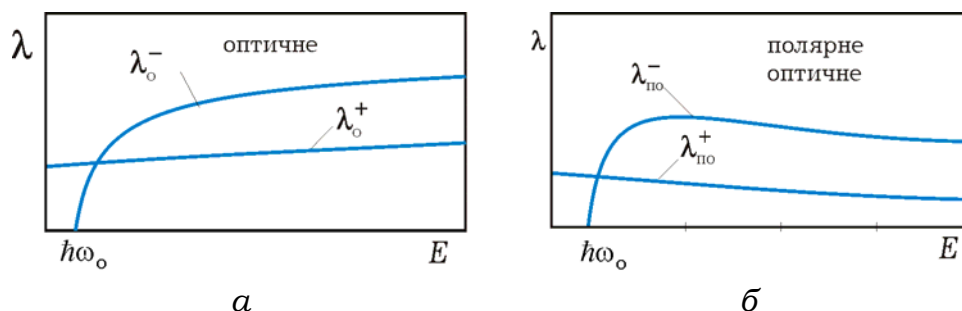


Рис. 2.12. Характерні залежності швидкостей розсіювання від енергії електрона для розсіювання на оптичному деформаційному потенціалі (а) та полярного оптичного (б)

У бінарних напівпровідникових сполуках, таких як арсенід галію, зустрічні поздовжні коливання підрешіток атомів різної полярності приводять до утворення поляризаційних електричних полів, які спричиняють так зване *полярне оптичне розсіювання*. Його швидкість можна визначити за формулою

$$\lambda_{\text{по}}^{\pm} = \frac{e^2 \sqrt{m^*} \omega_0 (N_{\phi} + 1/2 \mp 1/2)}{4\sqrt{2\pi} \epsilon_0 \hbar \sqrt{E}} \left(\frac{1}{\epsilon_{\infty}} - \frac{1}{\epsilon_{\text{ст}}} \right) \ln \left| \frac{\sqrt{E} + \operatorname{Re} \sqrt{E \pm \hbar \omega_0}}{\sqrt{E} - \operatorname{Re} \sqrt{E \pm \hbar \omega_0}} \right|. \quad (2.69)$$

Основною особливістю полярного оптичного розсіювання є залежність його швидкості від енергії електрона: $\lambda_{\text{по}} \sim E^{-1/2}$, якщо не брати до уваги відносно повільну логарифмічну залежність у (2.69). Із збільшенням енергії електрона зменшується ймовірність розсіювання, що сприяє накопиченню енергії та ще меншому її розсіюванню. Така ситуація спостерігається, якщо полярне оптичне розсіювання переважає, і приводить до нестійкого процесу, який називають *полярним оптичним убіганням* і який викликає інтенсивне розігрівання електронного газу (розд. 4).

Міждолинне розсіювання. Результатом міждолинного розсіювання є перехід електрона з однієї долини дисперсійної кривої в іншу. Для атомарних напівпровідників (германій, кремній) він зазвичай відбувається між еквівалентними долинами, наприклад, для кремнію (рис. 2.13, а) з Δ -долини у напрямку [100] також в Δ -долини еквівалентних напрямків [010], [001] тощо. В електричному полі таке розсіювання може привести до перерозподілу заселеності долин електронами, яка в рівноважному стані однакова. Для бінарних з'єднань $A^{III}B^V$ більш характерним є перехід між нееквівалентними долинами. Наприклад, для GaAs типовим є розсіювання (рис. 2.13, б) із центральної Γ -долини в бокові L-долини, розміщені в напрямках [111], або навіть у X-долини, що лежать ще вище.

Практично всі міждолинні переходи супроводжуються змінами хвильового вектора, співмірними з розмірами зони Бріллюена, і тому можливі лише за участі фононів з великими значеннями хвильових векторів. Такі фонони називають *міждолинними*, і відповідають вони ділянкам дисперсійних фононних кривих біля меж зони. Їх енергію переважно обирають середньою між крайніми значеннями енергій акустичних і оптичних фононів: біля $\approx 0,05$ еВ для Si і $\approx 0,03$ еВ для Ge і GaAs (див. рис. 2.10). Ці значення співмірні з тепловою енергією, тому розсіювання з поглинанням чи збудженням міждолинного фонона буде суттєво непружним.

Для переходів між нееквівалентними долинами слід враховувати зміну початку відліку енергії. Якщо електрон розсіюється з i -долини в j -ту, то закон збереження енергії (2.60) набуде вигляду

$$E_j - E_i = \pm \hbar \omega_{ij} - \Delta_{ij}, \quad (2.70)$$

де $\hbar \omega_{ij}$ – енергія міждолинного фонона; $\Delta_{ij} = E_{cj} - E_{ci}$ – міждолинна відстань (з урахуванням знака) від дна j -долини до дна i -долини. Відлік значень енергій E_i та E_j починають від мінімумів відповідних долин.

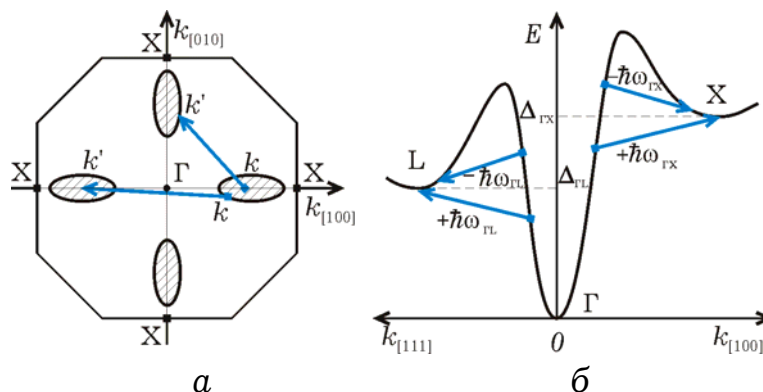


Рис. 2.13. Приклади переходів під час міждолинного розсіювання між еквівалентними долинами в кремнії (а) та між нееквівалентними долинами в арсеніді галію (б)

Швидкості міждолинного розсіювання розраховують за формулою, подібною за структурою до (2.69) для неполярного оптичного розсіювання:

$$\lambda_{\text{мд}}^{\pm} = \frac{Z_j (m_j^*)^{3/2} D_{ij}^2}{\sqrt{2\pi} \rho \hbar^3 \omega_{ij}} (N_{\text{ф}} + 1/2 \mp 1/2) \text{Re} \sqrt{E_i \pm \hbar \omega_{ij} - \Delta_{ij}}, \quad (2.71)$$

де Z_j – кількість долин, у які може відбутися розсіювання; m_j^* – ефективна маса густини станів у j -долині; D_{ij} – константа міждолинної взаємодії, що має значення порядку $10^8 \dots 10^9$ eV/cm; $N_{\text{ф}}$ – кількість фононів з енергією $\hbar \omega_{ij}$, яку визначають за (2.59). Оператор Re робить неможливим розсіювання за недостатньої енергії електрона. Якщо розсіювання відбувається між еквівалентними долинами, то замість Z_j потрібно підставляти число, на одиницю менше їх кількості, а $\Delta_{ij} = 0$. Для долин, що лежать на межі зони Бріллюена, їх кількість треба зменшити вдвічі.

Кількісне уявлення про інтенсивність різних видів розсіювання можна дістати з рис. 2.14, де зображено залежності

швидкостей розсіювання від енергії електронів у Γ - та L-долинах арсеніду галію. Слід звернути увагу на пороговий характер кривих для оптичних і міждолинних видів розсіювання, які потребують переходу електрона у вищий енергетичний стан.

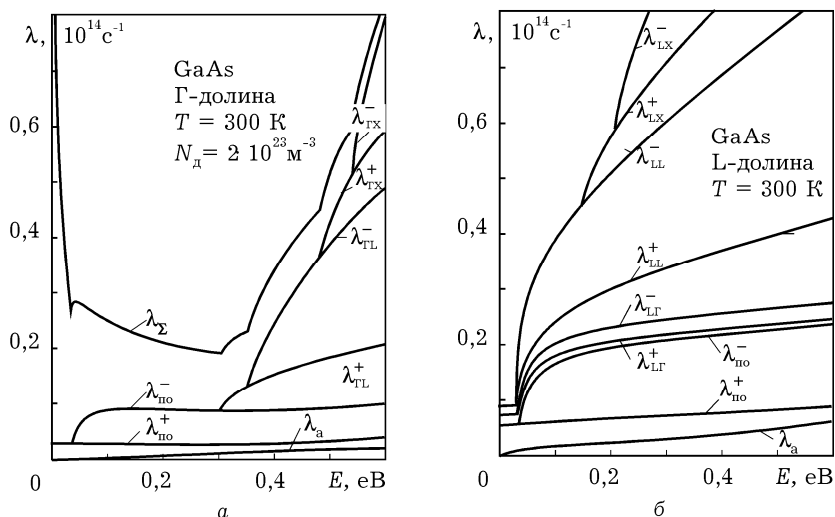


Рис. 2.14. Діаграма залежностей швидкості розсіювання для різних видів від енергії електрона в Γ -долині (а) та L-долині (б). Графік для кожного виду підсумовується з попередніми. Верхня крива λ_{Σ} сумарної швидкості включає також домішкове розсіювання

Домішкове розсіювання. У типових випадках домішки, уведені в напівпровідник, є іонізованими, і викликане ними розсіювання означає відхилення траєкторії електрона електростатичним полем іона. Подібну задачу розглядав Резерфорд, коли досліджував розсіювання α -часток. Електрон у кулонівському полі домішкового іона має гіперболічні траєкторії (рис. 2.15). Оскільки маса іона домішки на 3...4 порядки перевищує масу електрона, то передавання енергії в результаті їх взаємодії майже не відбувається, і таке співударяння буде практично *пружним*. Відповідно не змінюється значення швидкості, а тільки її напрямок, і $|\vec{k}'| = |\vec{k}|$. Кут розсіювання θ між векторами \vec{k}' і \vec{k} тим більший, чим менші значення швидкості електрона або дірки і так звана *прицільна відстань* b (рис. 2.15).

Для носіїв у твердому тілі модель Резерфорда дає невірний результат за великих прицільних відстаней: для малих кутів розсіювання інтеграл (2.56) розбігається, що пов'язано з повільністю спадання кулонівського потенціалу. Цей недолік можна усунути, враховуючи поля, створювані іншими іонами домішки і носіями заряду. Їхні поля тією чи іншою мірою

екранують поле іона домішки і «обрізають» кулонівський потенціал, послаблюючи його дію перш за все на великих відстанях (рис. 2.16).

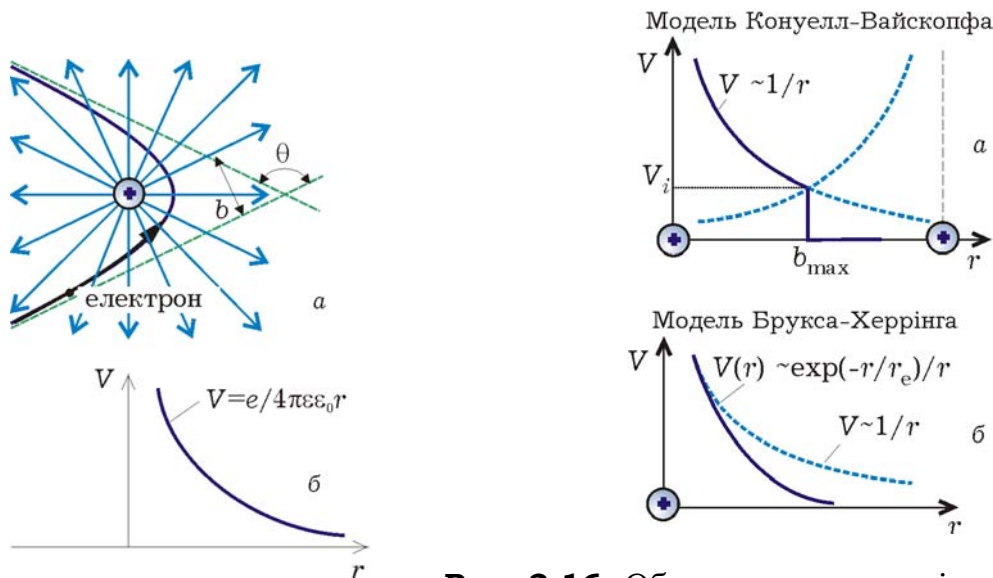


Рис. 2.15. Траєкторія руху електрона у кулонівському полі іона донорної домішки

Рис. 2.16. Обмеження потенціальної функції іона домішки у разі екранування розсіювального поля: *a* – сусідніми іонами домішки (модель Конуелл-Вайскопфа), *б* – рухливими носіями заряду (модель Брукса-Херрінга)

У моделі Конуелл – Вайскопфа область дії поля іона домішки обмежують половиною середньої відстані між іонами (рис. 2.16, *a*), тому максимальна прицільна відстань

$$b_{\max} = a/2 = 1/2 \sqrt[3]{N_{\text{д}}}. \quad (2.72)$$

Цей вираз впливає з умови, що одному атому домішки в середньому відповідає кубічний об'єм з ребром *a*. Іноді в літературі можна зустріти дещо інший вираз, який отримано, якщо вважати, що цей елементарний об'єм є сферою з радіусом *a*. У підсумку врахування обмеження розсіювального поля полями сусідніх атомів домішки дає для швидкості розсіювання

$$\lambda_{\text{д}} = \frac{v}{b_{\max}} \left(\frac{U_i}{2E} \right)^2 \ln \left(1 + \frac{2E}{U_i} \right), \quad (2.73)$$

де $E = m^*v^2/2$ – кінетична, а $U_i = e^2/4\pi\epsilon\epsilon_0b_{\max}$ – потенціальна енергія електрона в полі іона на відстані b_{\max} .

У моделі Брукса-Херрінга далекодія розсіювального поля іона домішки обмежується його екрануванням полями носіїв заряду. У такому разі розподіл потенціалу в околі донорного іона (рис. 2.16, б) описують виразом

$$V(r) = (e/4\pi\epsilon\epsilon_0r) \exp(-r/r_e), \quad (2.74)$$

де $r_e = (\epsilon\epsilon_0 k_B T/e^2n)^{1/2}$ – радіус екранування; n – концентрація екранувальних зарядів (електронів), а у формулі (2.73) змінюється останній співмножник:

$$\lambda_d = \frac{v}{b_{\max}} \left(\frac{U_i}{2E} \right)^2 \left[\ln(1+\eta) - \frac{\eta}{1+\eta} \right]. \quad (2.75)$$

Тут $\eta = (2kr_e)^2$, а для квадратичного закону дисперсії, коли $E = \hbar^2 k^2/2m^*$, $\eta = 8m^* E r_e^2 / \hbar^2$.

Слід зазначити, що вираз (2.74) придатний для максвеллівського розподілу. Але іноді його використовують, якщо середня енергія хаотичного руху електронів помітно перевищує величину $k_B T$, підставляючи у вираз для радіуса екранування електронну температуру T_e .

Обидві моделі підтверджують відомий результат про температурну залежність рухливості. Якщо вважати, що енергія носіїв відповідає тепловій ($E \sim k_B T$, $v \sim T^{1/2}$), і знехтувати повільною логарифмічною залежністю у (2.73), то отримаємо, що швидкість домішкового розсіювання пропорційна концентрації домішки N_d і обернено пропорційна $T^{3/2}$. Модель Конуелл-Вайскопфа за дуже низьких температур дає значно завищений результат (формально $\lambda \rightarrow \infty$ при $T \rightarrow 0$), якщо не враховувати ефекти виморожування домішки (див. розд. 3).

Розсіювання на сплавному потенціалі. Останнім часом широко використовують тверді розчини бінарних сполук $A^{III}B^V$, які утворюють три- та чотирикомпонентні напівпровідникові матеріали, наприклад $Al_xGa_{1-x}As$, $In_xGa_{1-x}As_yP_{1-y}$, тощо. Їх можна розглядати як сплави бінарних з'єднань $A^{III}B^V$, у яких навіть за ідеальної будови кристалічної ґратки розміщення різних атомів

однієї групи має випадковий характер. У таких з'єднаннях може бути суттєвим розсіювання на сплавному потенціалі, викликане хаотичними відхиленнями потенціалу кристалічної ґратки від ідеальної періодичної функції.

Його інтенсивність залежатиме від упорядкованості кристалічної ґратки. За абсолютної упорядкованості, якби атоми різних груп були рівномірно розміщені у вузлах кристалічної решітки, утворюючи надґратку, відмінність між ними можна було б врахувати в ефективній масі і розсіювання на сплаві було б відсутнім. З іншого боку, дослідження міри неупорядкованості кристалічної ґратки є досить складним, тому розглядають абсолютно впорядковане розташування атомів, але ймовірності їхнього перебування у конкретних вузлах решітки, що призначені для атомів різних груп, дорівнюють x і $1-x$ відповідно.

Швидкість розсіювання на сплавному потенціалі (ймовірність переходів) можна розрахувати за співвідношенням:

$$\lambda(E) = \frac{3\pi\sqrt{2}(m^*)^{3/2}}{16\hbar^4} x(1-x)\Omega|\Delta U|^2 \sqrt{E} \quad (2.76)$$

де E – енергія електронів, x – стехіометричний індекс сплаву, $\Omega = a(x)^3 / 4$ – об'єм елементарної комірки з радіусом, що дорівнює половині відстані до найближчих сусідів, а $\Delta U(x)$ – сплавний потенціал сполуки.

Співвідношення (2.76) демонструє, що час релаксації імпульсу для розсіювання на сплавному потенціалі залежить від величини $\Delta U(x)$, а вибір цього параметру викликає деякі складнощі. У різних наукових джерелах як $\Delta U(x)$ пропонують використовувати різницю енергій електронного ΔU_{EA} споріднення бінарних сполук, що входять до складу сплаву, або як різницю екранованих кулонівських потенціалів атомів ΔU_{EN} . Іноді застосовують різницю величин заборонених зон компонентів розчину – ΔU_{EG} . Значення ΔU_{EG} є більш надійним порівняно з ΔU_{EA} та ΔU_{EN} через наявну неоднозначність відповідних значень енергій електронного споріднення та ковалентних радіусів у літературі.

Серед інших видів розсіювання, які можуть відігравати помітну роль за певних умов, слід назвати *електрон-електронне розсіювання*, яке відбувається за високих концентрацій домішки. Під час міжелектронної взаємодії повний імпульс і енергія колективу залишаються незмінними, лише перерозподіляючись між частками, хаотизуючи їх рух та згладжуючи функцію розподілу.

З позицій теорії розсіювання можна розглядати й ті взаємодії, що викликають міжзонні переходи: генерацію носіїв під дією світла, співударянь, сильних електричних полів тощо, та їх рекомбінацію з випромінюванням квантів світла, фононів. Однак час релаксації для цих процесів зазвичай на декілька порядків більший, ніж для внутрізонних, тому генераційно-рекомбінаційні явища відносять до повільних і розглядають окремо.

2.4.4. Рухливість електронів

Результати теорії розсіювання застосовують, мабуть, найчастіше у питанні про розрахунок рухливості. Саме особливості різних механізмів розсіювання та деталей зонної структури визначають характерні електричні властивості тих чи інших напівпровідників та динаміку електронних процесів.

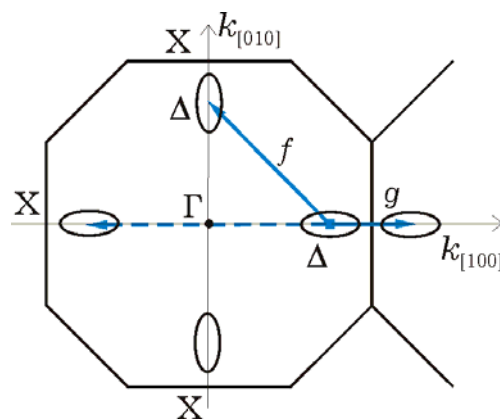
Нагадаємо, що рухливість $\mu = e\tau_p/m_c^*$ визначається середнім часом релаксації імпульсу τ_p і омічною ефективною масою m_c^* . Середній час релаксації імпульсу найчастіше знаходять через суму обернених часів релаксації для окремих механізмів розсіювання, хоча точність цього правила, яке ґрунтується на принципі незалежності одне від одного різних видів зіткнень, іноді піддають сумніву. Питома вага певного виду розсіювання визначає вплив зовнішніх чинників на рухливість. У цьому підрозділі йдеться про рухливість у слабких електричних полях, коли температура електронного газу T_e дорівнює температурі кристалічної ґратки T .

Кремній. З фононних механізмів для кремнію характерні акустичне та неполярне оптичне розсіювання. Акустичне розсіювання, строго кажучи, є анізотропним, бо електрони взаємодіють як з поздовжніми, так і з поперечними коливаннями кристалічної ґратки. Але переважаючою є взаємодія з поздовжніми коливаннями. Тому для розрахунку часів релаксації

для акустичного розсіювання можна з достатньою точністю використовувати ізотропне наближення, але зі швидкістю розповсюдження *поздовжніх* коливань $v_{зв} = v_l = 9,0 \cdot 10^3$ м/с і відповідним деформаційним потенціалом $D_a = 6$ еВ. Діелектрична проникність і густина речовини відповідно дорівнюють $\epsilon = 11,9$ і $\rho = 2,33$ г/см³. Слід також мати на увазі, що для розрахунку швидкості розсіювання необхідно використовувати значення ефективної маси густини станів, яку визначають через компоненти тензора ефективної маси ($m_l^* = 0,98m$, $m_t^* = 0,19m$) з урахуванням кількості еквівалентних долин ($M = 6$), що відповідають дну зони провідності: $m_d^* = 6^{2/3}(m_t^{*2} m_l^*)^{1/3} = 1,08m$. Для розрахунку рухливості треба використовувати омічну ефективну масу $m_c^* = 3m_t^* m_l^* / (2m_l^* + m_t^*) = 0,26m$.

Оптичне розсіювання в кремнії є неполярним і відбувається на оптичному деформаційному потенціалі. За наявності різних долин таке розсіювання розглядають фактично як еквівалентне міждолинне. Оскільки нижні долини кремнію розміщені в зоні Бріллюена в точках Δ , яким відповідає значення $k = 0,85(2\pi/a)$, то можливі два типи переходів (рис. 2.17).

Рис. 2.17. Види переходів під час оптичного розсіювання в кремнії: f – переходи між перпендикулярними долинами, g – переходи між співвісними долинами (штрихова лінія), які зручніше розглядати, як переходи в сусідню зону Бріллюена



Переходи між співвісними долинами (штрихова лінія) з урахуванням періодичності закону дисперсії зручніше розглядати як так звані g -переходи (суцільна лінія) у сусідню зону Бріллюена зі зміною хвильового вектора на $\Delta k = 0,29(2\pi/a)$. Інший можливий тип розсіювання викликає так звані f -переходи (рис. 2.17) між долинами, що лежать на перпендикулярних вісях. Розглянувши відповідні фононні дисперсійні криві побачимо, що такі переходи можуть відбуватися за участю різних типів фононів з різними енергіями. Цим трьома видами розсіювання

відповідають різні оптичні деформаційні потенціали. Для спрощення розрахунків без великої втрати точності всі три види розсіювання можна розглядати як один з деякими усередненими еквівалентними параметрами. В ефективну масу густини станів при цьому не включають кількість долин $m_{d1}^* = (m_t^* m_l^*)^{1/3}$, але час релаксації помножують на кількість долин, у які відбувається розсіювання, тобто на 5. Слід також мати на увазі, що оптичні розсіювання відбуваються переважно зі збудженням фонона.

Для аналізу домішкового розсіювання придатний вираз (2.73). Основною проблемою при цьому є вибір того чи іншого наближення. Нагадаємо, що в моделі Брукса-Херрінга фактором, що обмежує далекодію поля іона домішки, вважається екранування кулонівського поля рухливими носіями заряду. У моделі Конуелл-Вайскопфа таким фактором є наявність сусідніх домішкових атомів. Певним критерієм для вибору тієї чи іншої моделі може слугувати співвідношення між середньою міждомішковою відстанню та радіусом екранування. Так, якщо концентрація домішки $N_d = 10^{21} \text{ м}^{-3}$ і температура $T = 300 \text{ К}$, то радіус екранування $r_e = 0,13 \text{ мкм}$, а максимальна прицільна відстань $b_m = 0,05 \text{ мкм}$. Співвідношення $b_m < r_e$ говорить про те, що сильнішим фактором є обмеження розсіювання сусідніми іонами домішки, а більш придатною – модель Конуелл-Вайскопфа. Наближення Брукса-Херрінга обґрунтоване у випадку суттєвого зростання ролі екранування рухливими носіями, що спостерігається із підвищенням концентрації домішки і (або) за низьких температур. За дуже низьких температур слід також враховувати «виморожування» домішки.

Результати розрахунків часів релаксації, які пояснюють внесок різних механізмів розсіювання у підсумкове значення рухливості, наведено на рис. 2.18, а. Переважаючим, крім області низьких температур, є розсіювання на фононах.

Вплив оптичних фононів виявляється лише із підвищенням температури до $\approx 100 \text{ К}$. За низьких температур зростає роль домішкового розсіювання, причому для найнижчих температур – на нейтральних домішкових атомах.

Підсумком такого аналізу є залежності рухливості від температури (рис. 2.18, б) за різних концентрацій домішки. У логарифмічних координатах температурна залежність рухливості для найнижчих температур спрямлюється (якщо немає домішки) і її нахил відповідає степеневій залежності $T^{-3/2}$, характерній для

акустичного розсіювання. Із підвищенням температури ця лінійність зникає, що означає переважання оптичного розсіювання. На найкрутішій ділянці (при температурі близько 200 К) нахил відповідає залежності $\mu \sim T^{-2,4}$, яка добре узгоджується з експериментальною залежністю, поданою на тому самому рисунку.

Наявність домішки знижує рухливість тим більше, чим вища її концентрація. Це сильніше виявляється зі зменшенням температури, коли фононні механізми розсіювання слабшають. Помітне відхилення від експериментальних значень (рис. 2.19) спостерігається лише для великих концентрацій домішки, коли починається виродження електронного газу.

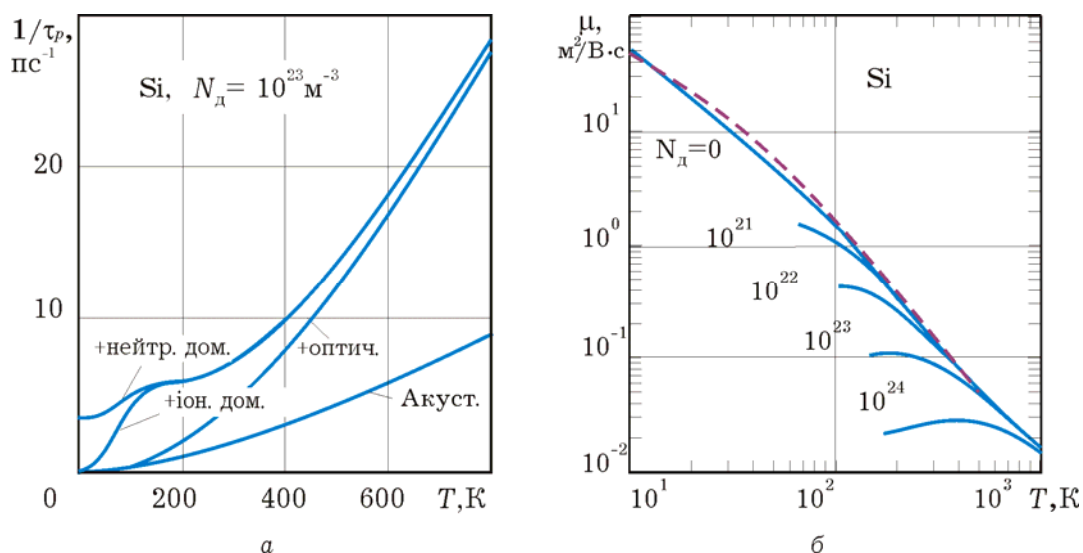


Рис. 2.18. Температурні залежності обернених часів релаксації імпульсу за фіксованої концентрації домішки (а) та рухливості за різних концентрацій домішки (б) для кремнію. Штрихова лінія – експериментальна. Графіки обернених часів релаксації для кожного виду, як і на рис. 2.14, підсумовуються з попередніми

Германій. Для германію характерні ті самі механізми розсіювання, що й для кремнію. Акустичне розсіювання тут також переважає в області низьких температур (якщо немає домішки). Для розрахунку часу акустичної релаксації може бути використане значення акустичного деформаційного потенціалу $D_a = 8 \text{ eV}$, швидкості звуку $v_{зв} = 5,4 \cdot 10^3 \text{ м/с}$, густини $\rho = 5,3 \text{ г/см}^3$, діелектричної проникності $\epsilon = 16,0$ та ефективної маси густини станів $m_d^* = 0,56m$.

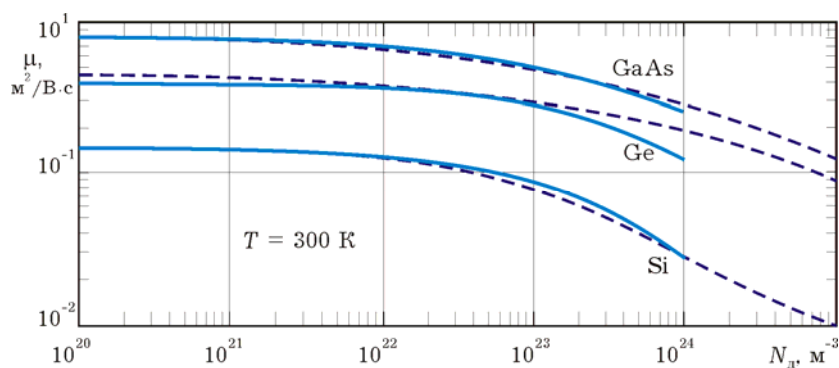


Рис. 2.19. Концентраційні залежності рухливості для кремнію, германію та арсеніду галію за кімнатної температури. Штрихові лінії – експериментальні

Оптичне розсіювання відбувається на оптичному деформаційному потенціалі і приводить до переходів електронів в еквівалентні долини. Основна відмінність від кремнію пов'язана з особливістю закону дисперсії. Як відомо, дно зони провідності для германію знаходиться в Λ -напрямку на межі зони Бріллюена в точках L (рис. 1.24). Це відповідає наявності восьми половин або чотирьох повних еквівалентних долин із сильною анізотропією ефективної маси $m_t^*/m_l^* = 0,08/1,58$. Тому, як і в кремнії, можливі різні типи електронних переходів між долинами, але в середньому їх можна змодельювати як одне міждолинне еквівалентне розсіювання на оптичному деформаційному потенціалі для $D_0 = 5,5 \cdot 10^8$ eV/cm за участі оптичних фононів з енергією $\hbar\omega_0 = 0,034$ eV. Для аналізу домішкового розсіювання тут так само, як і в кремнії, перевагу має модель Конуелл-Вайскопфа.

Результати розрахунку обернених часів релаксації, що показують внесок різних механізмів розсіювання, зображено на рис. 2.20, а. За найнижчих температур домішкове розсіювання переважає як на нейтральних, так і на іонізованих атомах домішки. Воно залежить від концентрації домішки і слабшає зі зменшенням N_d і збільшенням температури. Це наочно видно на рис. 2.20, б, де зображено температурну залежність рухливості.

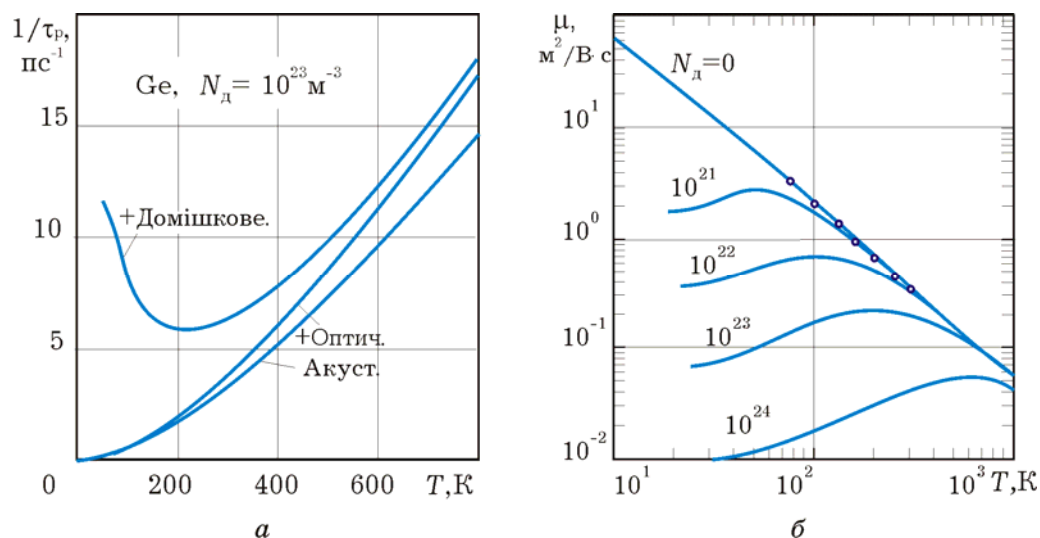


Рис. 2.20. Такі самі залежності, як і на рис. 2.18, для германію. Експериментальні значення для $N_d = 0$ позначено точками

Зі зростанням температури зростає і роль фононних механізмів розсіювання. Але порівняно із кремнієм оптичне розсіювання тут має менше значення. Це проявляється у похилішому спаданні рухливості з максимальним нахилом кривої $\mu(T)$ у логарифмічних координатах, який відповідає степеневій залежності $\sim T^{-1,6}$.

Як і для кремнію, але ще помітніше, обчислені для $T = 300$ К (рис. 2.19) значення $\mu(N_d)$ розходяться з експериментальними, якщо концентрації домішки більші за 10^{23} м^{-3} . Можливо, це пов'язано з меншою енергією іонізації атомів домішки в германії і, відповідно, із ще більшим концентраційним виродженням електронів у зоні провідності. Також невиправдано високою виявилась роль розсіювання на нейтральних домішках, що особливо помітно за високих концентрацій домішки.

Арсенід галію. Одна з головних особливостей розсіювання в GaAs пов'язана з помітним внеском іонного зв'язку між атомами Ga і As. Це приводить до того, що з фононних механізмів основну роль відіграє полярне розсіювання за участю оптичних фононів з енергією $\hbar\omega_0 = 0,035$ еВ. Результати розрахунку швидкості цього виду розсіювання (2.14) і відповідного часу релаксації (2.26) суттєво залежать від точності задання близьких значень статичної і високочастотної діелектричних проникностей $\epsilon_{ст}$ і ϵ_∞ , які зв'язані виразом (2.2) із значеннями частот ω_{LO} і ω_{TO} відповідно поздовжньої і поперечної мод фононного спектра (рис.

2.17, в) при $k_{\text{ф}} = 0$. На рис. 2.21 наведено результати для значення $\omega_{\text{ЛО}}/\omega_{\text{ТО}} = (\epsilon_{\text{ст}}/\epsilon_{\infty})^{1/2} = 1,15$. Цей вид розсіювання виявляється за температур вищих ~ 100 К і для кімнатної температури є переважаючим (якщо немає домішкового розсіювання).

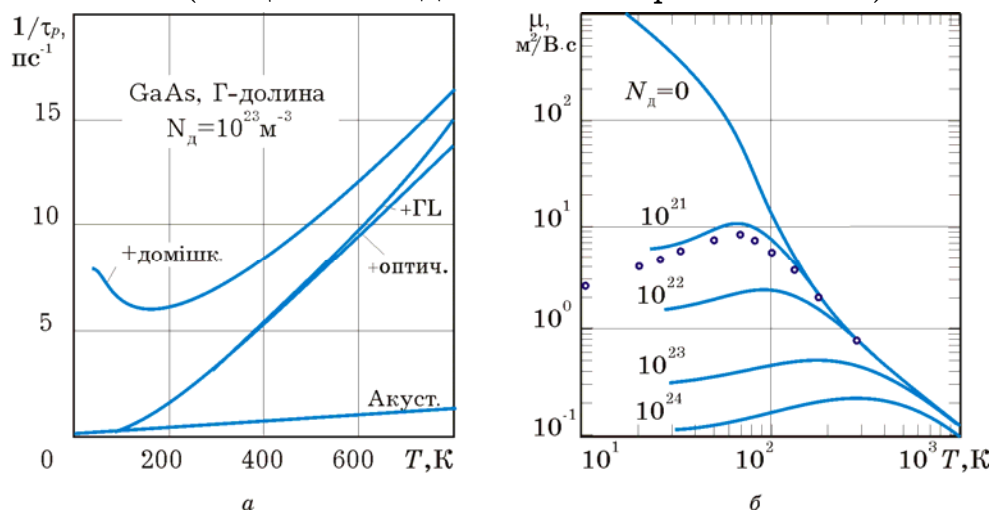


Рис. 2.21. Такі самі залежності, як і на рис. 2.18, для арсеніду галію. Точками позначено експериментальні значення для $N_{\text{д}} = 10^{21} \text{ м}^{-3}$

Друга суттєва відмінність пов'язана з особливістю закону дисперсії GaAs і зі структурою його долин. Крім нижньої Г-долини біля дна зони провідності з ізотропною ефективною масою $m_{\Gamma}^* = 0,063m$ існують і більш високі: вісім L-долин (як для германію) на межі зони Бріллюена ($M_{\text{L}} = 4$) на відстані вздовж осі енергії $\Delta_{\text{ГЛ}} = 0,33$ еВ із сильно анізотропною ефективною масою $m_t^*/m_l^* = 0,075/1,9$, а також на відстані $\Delta_{\text{ГХ}} = 0,55$ еВ шість X-долин, як для кремнію, але не в точках Δ , а на межі зони Бріллюена ($M_{\text{X}} = 3$), з ефективними масами $m_t^*/m_l^* = 0,23/1,3$. Наявність L-долин починає виявлятися за температур вищих ~ 600 К: взаємодія електронів з оптичними фононами з енергією $\hbar\omega_{ij} = 0,026$ еВ приводить до міждолинного розсіювання, яке спричиняє їх переходи із станів Г-долини у стани L-долин. Для розрахунку часів релаксації цього виду розсіювання можна скористатися співвідношенням (2.32) з константами міждолинної взаємодії $D_{ij} = 10^9$ еВ/см і густиною речовини $\rho = 5,32$ г/см³.

Акустичне розсіювання в GaAs дуже слабке і проявляється лише якщо немає конкуруючих видів розсіювання: полярного оптичного і домішкового. Це видно з рис. 2.21, б, де подано

залежності рухливості від температури, для $N_d = 0$: в області температур, нижчих за ~ 70 К. У логарифмічних координатах залежність $\mu(T)$ має вигляд прямої з нахилом $-1,5$, характерним саме для акустичного розсіювання. Для вищих температурах (до ≈ 600 К) спостерігається зменшення рухливості за рахунок зростання ролі розсіювання на полярних оптичних фонах. За ще вищих температур включаються механізми розсіювання в L- та X-долини, що приводить до ще крутішого спадання кривої $\mu(T)$.

На тлі цих механізмів домішкове розсіювання виявляється тим сильніше, чим вища концентрація домішки, і відповідно зростає значення температури, за якої рухливість максимальна. Тому, як і для Si та Ge, залежність рухливості від концентрації домішки (рис. 2.19) спадає тим швидше, чим вища N_d . Наведені результати розрахунків близькі до експериментальних, але розходяться з ними в області високих концентрацій, за яких починається виродження напівпровідника..

Нітрид галію. Аналогічна картина спостерігається і для нітриду галію. З фононних механізмів для GaN характерні акустичне і полярне оптичне розсіювання. Акустичне розсіювання зростає з підвищенням температури, але його внесок проявляється тільки за відсутності конкуруючих домішкового та оптичного розсіювань. Для розрахунку часу релаксації акустичного розсіювання були використані значення акустичного деформаційного потенціалу $D_a = 10,1$ еВ, швидкості звуку $v_{zv} = 4,57 \cdot 10^3$ м/с, густини $\rho = 6,15$ г/см³, діелектричної проникності $\epsilon = 7,5$ та ефективні маси густини станів відповідно до кожної долини.

У нітриді галію так само як і в інших сполуках переважаючим є полярне оптичне, яке виявляється із підвищенням температури до ≈ 120 К. Для розрахунків часів релаксації обирають значення статичної і високочастотної діелектричної проникності 9,6 та 5,5 відповідно, енергії полярних оптичних фононів 65 меВ.

Вклад міждолинних механізмів з енергією міждолинних фононів біля 59 меВ та константою міждолинного зв'язку $1,1 \cdot 10^9$ еВ/см є незначним навіть за високих температур. За низьких температур зростає роль домішкового розсіювання, причому для найнижчих температур домішкове розсіювання

переважає як на нейтральних, так і на іонізованих атомах домішки.

Роль оптичних механізмів розсіювання порівняно з GaAs в GaN має більше значення, що виявляється у більш пологому спаданні кривої $\mu(T)$. Підсумком такого аналізу є залежності на рис. 2.22.

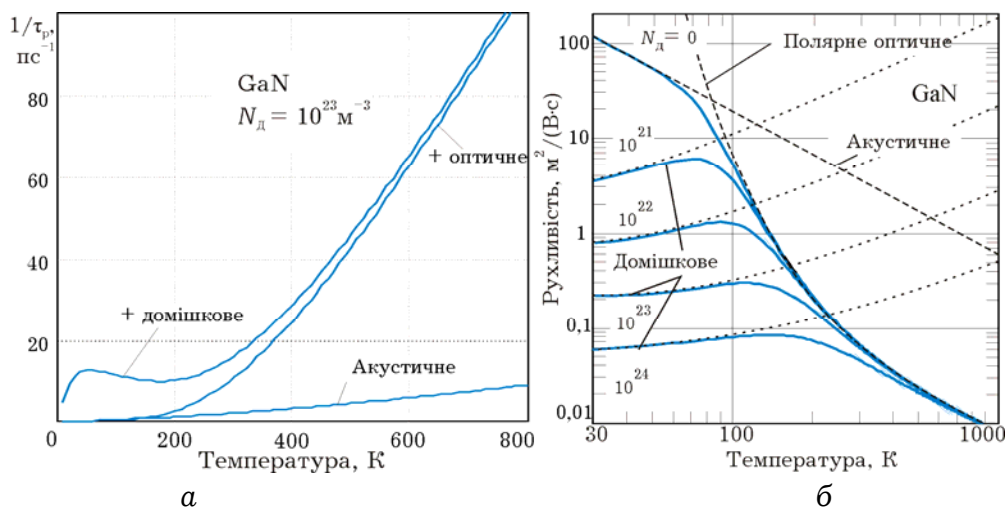


Рис. 2.22. Температурна залежність обернених часів релаксації імпульсу за концентрації домішки $N_d = 10^{23} \text{ м}^{-3}$ (а) та рухливості за різних концентрацій домішки (б) для нітриду галію. Криві для кожного виду розсіювання підсумовуються з попередніми

Наведені результати розрахунків близькі до відомих експериментальних, але розходяться з ними в області високих концентрацій, за яких починається виродження напівпровідника.

2.5. ДИФУЗИЯ НОСІВ ЗАРЯДУ

У багатьох приладах електронної техніки використовується ще один, дифузійний, спосіб створення спрямованого перенесення носіїв заряду, який проявляється в неоднорідних середовищах. Дифузією називається спрямований рух потоку часток, який виникає за наявності градієнта концентрації під дією їх хаотичного руху. Саме хаотичний рух є «рушійною силою» дифузії, оскільки тільки він призводить до того, що зустрічні потоки часток (що рухаються з рівною ймовірністю на всі боки) будуть різні за наявності перепаду концентрації: переважаючим

буде перенесення часток з області з більшою концентрацією в область з її меншим значенням.

Дифузія заряджених часток підпорядковується тим саме законам, що і дифузія нейтральних, і описується законами Фіка. Перший з них характеризує швидкість спрямованого дифузійного руху \vec{v}_D і має вигляд:

$$\vec{v}_D = -D \frac{1}{n} \text{grad } n, \quad (2.77, a)$$

або у більш зручній формі:

$$n\vec{v}_D = -D \text{grad } n. \quad (2.77, б)$$

Оскільки ліва частина є густиною потоку дифундуючих часток, тобто їх кількістю, яка проходить через одиницю поперечного перерізу за одиницю часу, то сенс першого закону дифузії зводиться до того, що густина потоку дифузії пропорційна градієнту концентрації. Коефіцієнт пропорційності називається коефіцієнтом дифузії і відповідно до кінетичної теорії газів дорівнює:

$$D = \bar{l} \cdot v_{\text{хаот}} / 3. \quad (2.78)$$

Знак «-» в (2.77) означає, що дифузійний потік спрямований у бік зменшення концентрації, тобто протилежно вектору градієнта, який спрямований у бік збільшення концентрації.

Нестаціонарний процес дифузії описується другим законом Фіка, який має вигляд

$$\frac{\partial n}{\partial t} = D \nabla^2 n, \quad (2.79)$$

і зв'язує швидкість зміни концентрації в часі з її просторовим розподілом $n(\vec{r}, t)$.

Не дивлячись на те, що дифузія заряджених часток описується такими ж законами, що і нейтральних, вона має ряд характерних особливостей. Передусім потрібно відмітити, що дифундуючі частинки переносять заряд і створюють *дифузійний струм* $\vec{j}_D = -en\vec{v}_D$, величина якого пропорційна потоку часток та їх заряду. За наявності частинок різного виду (наприклад, електронів або дірок у напівпровіднику) повний струм має різні

складові, які з урахуванням напрямку (рис. 2.23) виражаються співвідношеннями:

$$\begin{aligned} \vec{j}_{Dn} &= eD_n \text{grad } n, \\ \vec{j}_{Dp} &= -eD_p \text{grad } p. \end{aligned} \quad (2.80)$$

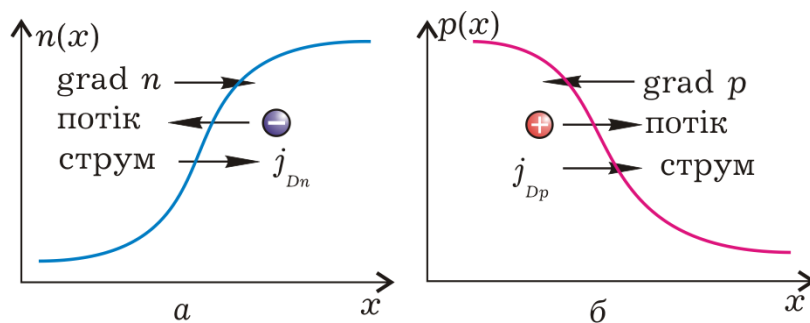


Рис. 2.23. Напрямки дифузійних потоків і електричних струмів електронів (а) і дірок (б) відносно напрямків градієнта їх концентрації

Друга особливість дифузії пов'язана з перерозподілом зарядів у результаті дифузійного перенесення. Якщо нейтралізації дифундуючих зарядів не відбувається, то виникає просторовий заряд, який створює гальмівне електричне поле, або на мові енергетичних моделей - потенціальний бар'єр, який стримує процес дифузії (рис. 2.24). Це явище можна також описати за допомогою дрейфового струму, викликаного електричним полем. Нехай дифузійний потік електронів (рис. 2.24, а) спрямований справа наліво. У результаті дифузійного перенесення зарядів ліворуч виникає надлишок електронів, а справа - недолік, і відповідно створюється просторовий заряд у вигляді подвійного електричного шару, поле якого спрямоване, як вказано на рисунку. Сила цього поля перешкоджає дифузійному перенесенню, повертаючи назад ті з електронів, енергії яких недостатньо для подолання гальмівної дії поля. Це створює дрейфовий струм, напрям якого протилежний до дифузійного. Аналогічна ситуація виникає і у разі розгляду дифузії дірок (рис. 2.24, б). У результаті повний струм визначається сумою дрейфової і дифузійної складових: $\vec{j} = \vec{j}_{др} + \vec{j}_D$. Вирази для повного струму електронів і дірок відповідно мають вигляд:

$$\vec{j}_n = en\mu_n \vec{E} + eD_n \text{grad } n, \quad (2.81, a)$$

$$\vec{j}_p = e\mu_p\vec{\mathcal{E}} - eD_p \text{grad } p. \quad (2.81, б)$$

Ці вирази придатні також і у тому випадку, коли електричне поле створюється не внутрішнім перерозподілом зарядів, а зовнішнім джерелом, або їх спільною дією.

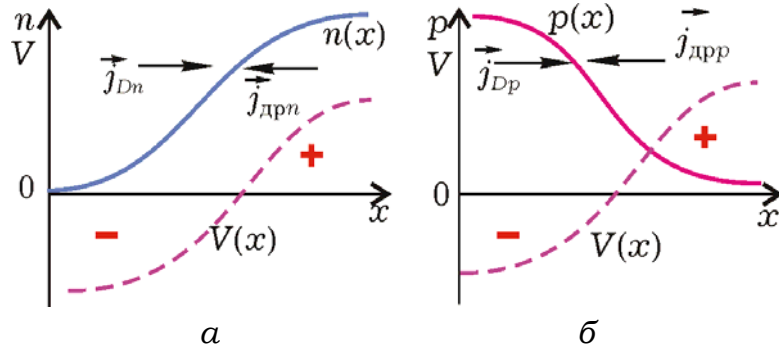


Рис. 2.24. Напрямки дифузійних і внутрішніх дрейфовий струмів електронів (а) і дірок (б). Створюване дифузійним перерозподілом зарядів електричне поле викликає дрейфові струми, спрямовані назустріч дифузійним

Ще одна особливість дифузії заряджених часток проявляється, якщо в описаній перед цим ситуації настає рівновага між зустрічними дифузійними і дрейфовими процесами і їх повний струм дорівнює нулю. Рівність та взаємна компенсація струмів $j_{др}$ і j_D визначає також наявність зв'язку між параметрами цих процесів, тобто між рухливістю і коефіцієнтом дифузії. Встановити цей зв'язок, наприклад, для електронів можна, прирівнявши повний струм в (2.81, а) нулю. Розглядаючи для спрощення одновимірний випадок, і замінюючи $|\text{grad } n| = dn/dx$, а $\mathcal{E} = -dV/dx$, отримаємо:

$$D_n \frac{dn}{dx} = \mu_n n \frac{dV}{dx},$$

звідки

$$\frac{dn}{n} = \frac{\mu_n}{D_n} dV. \quad (2.82)$$

Інтегрування цього рівняння дає закон зміни концентрації електронів у полі з потенціальним рельєфом $U(x) = -eV(x)$:

$$n(x) = n(0) \exp[-\mu_n U(x)/eD_n], \quad (2.83)$$

де $n(0)$ - концентрація в точці з нульовим потенціалом.

Цей вираз співпадає з розподілом Больцмана $n \sim \exp(-E/k_B T)$, якщо прийняти $eD_n = \mu_n k_B T$. Аналогічне співвідношення може бути отримане і для дірок, а враховуючи те, що кожен колектив може характеризуватися своєю температурою, запишемо:

$$D_n = \mu_n k_B T_n / e, \quad (2.84, a)$$

$$D_p = \mu_p k_B T_p / e. \quad (2.84, б)$$

Ці вирази, які називаються *співвідношеннями Ейнштейна*, встановлюють зв'язок між коефіцієнтом дифузії заряджених частинок і їх рухливостю. Слід також пам'ятати, що вони придатні для невірджених колективів (див. гл. 3), що знаходяться у рівноважному стані, а температура, яка входить в них, підкреслює теплову природу механізму дифузії.

Відмічені особливості описують процес дифузії частинок одного типу, так звану монополярну дифузію. Ще одна особливість проявляється в середовищі, яке містить рухливі носії різного знаку, і полягає в тому, що їх дифузія не може протікати незалежно одна від одної. Це явище називається *амбіполярною дифузією*. У цьому випадку носії різного типу, які зазвичай мають різну швидкість дифузії, утворюють просторовий заряд у вигляді подвійного електричного шару. Локальне електричне поле спрямоване таким чином, що гальмуватиме дифузію «швидких» часток (здебільшого електронів) і прискорюватиме «повільні». У стаціонарному стані сумарні швидкості дифузії різних часток вирівнюються, і повний струм буде відсутній, тобто $\vec{j}_n + \vec{j}_p = 0$. Якщо градієнти концентрацій носіїв не дуже відрізняються один від одного ($\text{grad } n \approx \text{grad } p$), то, виключаючи з (2.23, а) і (2.23, б) внутрішнє електричне поле, можна отримати

$$\vec{j}_n = -\vec{j}_p \approx e \frac{\sigma_p D_n + \sigma_n D_p}{\sigma_n + \sigma_p} \text{grad } n. \quad (2.85)$$

Порівнюючи цей вираз з (2.80), неважко побачити, що колективи різних часток дифундуватимуть як одне ціле, а дріб в (2.85) є коефіцієнтом амбіполярної дифузії

$$D_A = (\sigma_p D_n + \sigma_n D_p) / (\sigma_n + \sigma_p). \quad (2.86)$$

Корисно відзначити одну особливість амбіполярної дифузії, яка проявляється, коли переважає концентрація носіїв якогось типу. Наприклад, в донорному напівпровіднику $n \gg p$, тому $\sigma_n \gg \sigma_p$, і якщо D_n одного порядку з D_p , то $D_A \approx D_p$. В акцепторному напівпровіднику $p \gg n$ і $D_A \approx D_n$. В обох випадках швидкість амбіполярної дифузії визначається швидкістю дифузії неосновних носіїв.

2.6. РІВНЯННЯ НЕПЕРЕРВНОСТІ

Одним з рівнянь, які використовують для моделювання процесів в електроніці і відносять до фундаментальних, є рівняння неперервності, яке у класичній формі зазвичай має вигляд

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{j} = 0. \quad (2.87)$$

Це рівняння виражає закон збереження заряду: зміна густини просторового заряду у даній точці можлива лише за рахунок розбіжності струму, а народження та зникнення зарядів неможливе. На перший погляд може здатися дивним, що це рівняння, не зважаючи на його фундаментальність, не враховує процесів генерації та рекомбінації. Але ці процеси означають утворення або зникнення *рухливих носіїв* заряду, тобто їх перехід зі зв'язаного стану у вільний або навпаки, так що нейтральність середовища в цілому не порушується і зміни просторового заряду не відбувається. Явища генерації та рекомбінації мають бути враховані в окремих рівняннях неперервності для носіїв додатних та від'ємних зарядів. Оскільки для електронів $\rho_n = -en$, а для дірок $\rho_p = ep$, то для них рівняння (2.87) можна «розшарувати» на два:

$$\begin{aligned} \frac{\partial n}{\partial t} &= \frac{1}{e} \operatorname{div} \vec{j}_n + g_n - r_n, \\ \frac{\partial p}{\partial t} &= -\frac{1}{e} \operatorname{div} \vec{j}_p + g_p - r_p. \end{aligned} \quad (2.88)$$

У цих рівняннях доданки g та r визначають швидкості генерації та рекомбінації відповідно, тобто кількість рухливих носіїв, які з'являються або зникають в одиниці об'єму за одиницю часу.

У такій формі ці рівняння неперервності разом з виразами (2.81) для повного струму, рівнянням Пуассона

$$\operatorname{div} \vec{\mathcal{E}} = \rho / \epsilon \epsilon_0 \quad (2.89)$$

та рівняннями, які описують генераційно-рекомбінаційні процеси, є замкненою системою рівнянь для моделювання процесів у багатьох приладах твердотільної електроніки. Необхідно зазначити, що величина ρ у рівнянні Пуассона для напівпровідників визначається не тільки рухливими носіями, але й іонізованими домішковими атомами.

Контрольні запитання та вправи

2.1. На прикладі поведінки одного електрона в координатному і в k -просторі опишіть дрейфовий рух (однчасткова модель) в електричному полі, звернувши увагу на роль зіткнень. Що таке рухливість електронів і питома електропровідність речовини і від яких параметрів вони залежать?

2.3. Поясніть суть і ціль методу кінетичного рівняння. Які макроскопічні характеристики можуть бути визначені із знайденої функції розподілу? Чому цей метод вважається за напівкласичний?

2.4. Поясніть структуру і зміст кожного доданку в кінетичному рівнянні. Що таке інтеграл зіткнень і як він перетворюється при використанні наближення часу релаксації?

2.5. Яким чином кінетичного рівняння можна отримати закон Ома? Зверніть увагу на припущення, за яких це можливо. Що таке омична ефективна маса?

2.6. опишіть характер перехідних дрейфових процесів, що виникають в однорідному провідному середовищі при включенні-виключенні електричного поля. Від чого залежить швидкість цих релаксаційних процесів?

2.7. Яким чином змінюється провідність з підвищенням частоти? Від чого залежить порядок значень частот, за яких

зменшення провідності стає помітним? Що означає комплексний характер питомої електропровідності? Який граничний фазовий зсув між струмом і напругою може бути отриманий?

2.8. Поясніть вигляд релаксаційних рівнянь для концентрації, дрейфового імпульсу і енергії. У чому зміст і відмінність відповідних часів релаксації?

2.9. Що таке розсіювання, яка його роль у процесах переносу носіїв і якими параметрами можна характеризувати його інтенсивність?

2.10. У чому полягають основні особливості і механізми домішкового розсіювання? Як швидкість розсіювання (час релаксації, рухливість) залежать від температури і концентрації домішки?

2.11. Які види (моди) коливань кристалічної решітки характерні для ковалентних (Ge, Si) і ковалентно-іонних (GaAs) кристалів? Який порядок енергій мають відповідні фонони і від чого залежить їх кількість?

2.12. Які загальні закономірності фононного розсіювання? Від чого залежить ступінь пружності взаємодії електронів фононами?

2.13. Перерахуйте види фононного розсіювання, дайте їх коротку характеристику і порівняйте їх швидкості та часи релаксації, звернувши увагу на відмінність часів релаксації імпульсу і енергії.

2.14. Які види і результат міждолинного розсіювання? Чи можна його віднести до фононних видів розсіювання і в чому його специфіка? Що таке міждолинні фонони?

2.15. Опишіть типовий вигляд залежності рухливості напівпровідників від температури та вкажіть переважаючий вид розсіювання за різних температур.

2.16. Дайте визначення дифузії як макроскопічному процесу і вкажіть його внутрішню фізичну причину («рушійну силу»). Які особливості дифузії заряджених частинок? Які величини зв'язує співвідношення Ейнштейна, і за яких умов воно застосовне?

3. РІВНОВАЖНИЙ СТАН НОСІЇВ

Функціонування будь-якого електронного приладу пов'язане з процесами взаємодії рухливих носіїв заряду з електромагнітними полями в міжелектродному просторі. Для отримання (генерації) достатньої кількості рухливих носіїв заряду використовується яке-небудь зовнішнє джерело, енергія якого витрачається на розрив електричних зв'язків, що існують між зарядами в речовині.

3.1. ОСНОВНІ ПОЛОЖЕННЯ СТАТИСТИКИ ЕЛЕКТРОНІВ

Кристал є колективом, що складається з величезної кількості мікрочасток. За постійної температури і відсутності іншого зовнішнього впливу колектив електронів твердого тіла знаходиться у термодинамічній рівновазі з кристалічною ґраткою. За цих умов безперервно відбувається обмін енергією між електронами і кристалічною ґраткою. Електрони, поглинаючи теплову енергію ґратки (фонони), переходять у збуджений стан з більш високою енергією. Одночасно інші електрони, які знайшли незайняті стани з нижчою енергією, переходять до них, збуджуючи коливання кристалічної ґратки. У стані термодинамічної рівноваги ці зустрічні процеси врівноважують один одного, і у будь-який момент отримана і передана колективами часток енергії в середньому однакові та постійні у часі. Такий стан колективу часток називається *рівноважним*.

Для опису властивостей великого колективу немає сенсу, та і практично неможливо, знати детальну поведінку кожної частки окремо. Поведінка колективу, як цілого, підпорядковується статистичним закономірностям, основною особливістю яких є їх ймовірнісний характер. Ймовірнісний стан колективу часток може бути описаний статистичною функцією розподілу

$$N(E) = dn/dE, \quad (3.1)$$

яка відповідає енергетичній густині заповнення частками станів поблизу цього значення енергії E . Функцію $N(E)$ зручно представити

як добуток густини можливих станів $g(E) = dn_c/dE$ на середнє число часток $f(E) = dn/dn_c$ у цьому стані, тобто

$$N(E) = g(E) \cdot f(E). \quad (3.2)$$

Зручність такого подання полягає у тому, що задача пошуку функції розподілу розпадається на дві: *квантова задача* визначення розподілу енергетичних станів і *статистична задача* знаходження ймовірності заповнення цих станів частками.

Для опису колективу електронів у кристалі за умови термодинамічної рівноваги використовується *статистика Фермі-Дірака*. Відповідно до цієї статистики за низьких температур електрони кристала займають енергетичні стани з мінімально можливою енергією відповідно до принципу Паулі, а їх розподіл по станах описується функцією розподілу $f(E, T)$, яка визначає ймовірність того, що енергетичний стан E за температури T зайнятий електроном. Ця функція має вигляд

$$f(E, T) = 1 / \left(\exp \frac{E - E_{\phi}}{k_B T} + 1 \right), \quad (3.3)$$

де $E_{\phi}(T)$ - параметр функції розподілу, званий *електрохімічним потенціалом* або *енергією Фермі*, яка також може залежати від температури. З виразу (3.3) виходить, що $f(E_{\phi}) = 1/2$ за будь-якої температури; тому іноді користуються формальним сенсом величини E_{ϕ} , яка відповідає такому енергетичному стану, ймовірність заповнення якого електронами дорівнює $1/2$.

Графік функції $f(E)$ показаний на рис. 3.1. Для $T = 0$ з ймовірністю 1 електронами заповнені усі енергетичні стани нижче величини E_{ϕ} і немає жодного електрона з енергією, вищою за E_{ϕ} . У разі підвищення температури частина електронів набуває теплової енергії і повна їх енергія може стати вищою за рівень Фермі. Ступінчастість функції $f(E)$ «розмивається», і тим сильніше, чим вища температура. Проте теплове збудження впливає тільки на ті електрони, енергія яких трохи відрізняється від E_{ϕ} (на величину порядку $k_B T$ в обидві сторони).

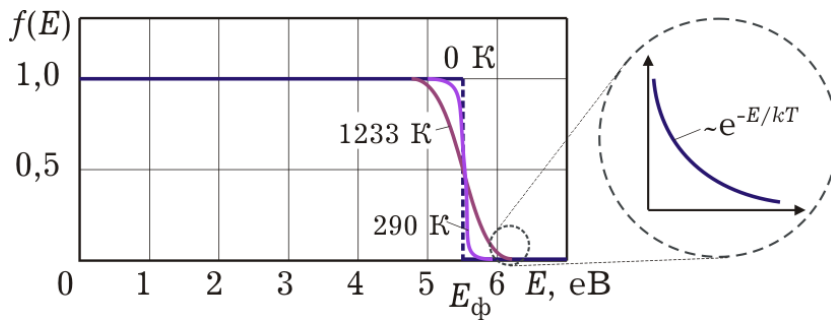


Рис. 3.1. Розподіл Фермі-Дірака для електронів провідності у сріблі за нульової температури, кімнатної температури і температури плавлення

Знання функції розподілу дозволяє визначити основні статистичні параметри колективу електронів. Наприклад, інтегруючи функцію розподілу у деякому інтервалі енергій, можна знайти концентрацію електронів, які мають енергію у цьому інтервалі:

$$n = \int_{E_1}^{E_2} N(E) dE. \quad (3.4)$$

Інтеграл від добутку енергії на функцію розподілу дорівнює повній енергії усіх електронів, енергія яких лежить у межах інтегрування

$$E_{\Sigma} = \int_{E_1}^{E_2} EN(E) dE. \quad (3.5)$$

Відношення (3.5) до (3.4) визначає середню енергію одного електрона

$$\bar{E} = E_{\Sigma} / n = \int_{E_1}^{E_2} EN(E) dE / \int_{E_1}^{E_2} N(E) dE. \quad (3.6)$$

Обчислення вказаних інтегралів у явному виді можливе тільки в окремих випадках, тому зазвичай доводиться використовувати наближені методи. Проте вирішення подібного виду задач істотно спрощується, якщо дана область енергій лежить вище за рівень Фермі хоч би на декілька одиниць $k_B T$. Це пов'язано з тим, що за умови $(E - E_{\phi}) > k_B T$ можна знехтувати одиницею у знаменнику формули (3.3) і функція розподілу Фермі-Дірака набуває вигляду

$$f_{\text{Ф-Д}}(E) \approx \exp[(E_{\phi} - E)/k_B T] = A \exp(-E/k_B T) = f_{\text{М-В}}(E). \quad (3.7)$$

Вираз (3.7) характерний для розподілу Максвелла-Больцмана, який описує поведінку класичних часток. Той факт, що за великих енергій функція $f(E) \ll 1$ означає, що кількість станів у багато разів більша, ніж кількість електронів, що заповнюють їх. Електрони слабо взаємодіють між собою і тому особливості, пов'язані з квантовою природою часток (зокрема, з їх спіном), у розподілі по станах не проявляються. Саме тому такий колектив часток може бути описаний класичним розподілом Максвелла-Больцмана і називається *невиродженням*. Якщо вказане наближення використовувати не можна, то колектив часток називається *виродженням*, і тоді він описується квантовою статистикою Фермі-Дірака.

З розподілом електронів по станах тісно пов'язаний розподіл дірок. Оскільки будь-який стан E може бути зайнятим або електроном, або діркою, то $f_n(E) + f_p(E) = 1$. Звідси функція розподілу дірок матиме вигляд

$$f_p(E) = 1 - f_n(E) = 1 / \{ \exp[(E_{\text{ф}} - E) / k_B T] + 1 \}. \quad (3.8)$$

Оскільки енергія дірок за знаком протилежна до енергії електронів, то збільшення енергії електронів відповідає переміщенню дірок вниз на електронній енергетичній діаграмі. Якщо змінити знак енергії в (3.8), то функції розподілу для електронів і дірок за формою будуть тотожні (рис. 3.2).

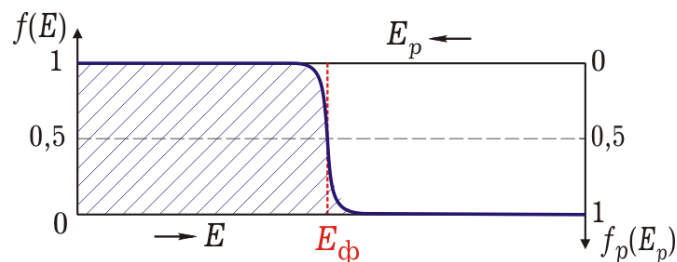


Рис. 3.2. Функції Фермі-Дірака для електронів (заштрихована) і дірок доповнюють одна одну до одиниці і збігаються із заміною напрямку відліку енергії дірок на зворотній

Корисно також мати на увазі одну цікаву особливість функції Фермі-Дірака $f(\varepsilon)$

$$\begin{aligned} \frac{df(\varepsilon)}{d\varepsilon} &= \frac{d}{d\varepsilon} \left(\frac{1}{1 + e^\varepsilon} \right) = \frac{e^\varepsilon}{(1 + e^\varepsilon)^2} = \frac{1}{1 + e^\varepsilon} \cdot \frac{1}{1 + e^{-\varepsilon}} = \\ &= f(\varepsilon) f_p(\varepsilon) = f(\varepsilon) [1 - f(\varepsilon)] \end{aligned} \quad (3.9)$$

Використовуючи цю властивість, можна інтеграл, які мають вид $\int f(E)[1-f(E)]dE = \int f(E)f(-E)dE$ зводити до виду $\int df(E)$.

3.2. РОЗРАХУНОК КОНЦЕНТРАЦІЇ РІВНОВАЖНИХ НОСІЇВ

3.2.1. Електрони у зоні провідності

Розрахувати концентрацію електронів у зоні провідності можна відповідно до формули (3.4), знаючи функцію розподілу густини електронів $N(E) = g(E) \cdot f(E)$ по станах у зоні провідності. Густина станів $g(E)$ поблизу дна зони провідності описується параболічною залежністю (рис. 3.3, а), отриманою у розділі 1.1.5

$$g(E) = 4\pi(2m_d^*/h^2)^{3/2}(E-E_c)^{1/2}. \quad (3.10)$$

Вигляд функції розподілу Фермі-Дірака (рис. 3.3, б) визначається рівнем Фермі, положення якого підлягає уточненню, і температурою. Концентрація електронів у зоні провідності дорівнює площі під кривою $N(E)$ на рис. 3.3, в і може бути обчислена як інтеграл (3.4) у межах від нижнього краю зони провідності до верхнього. Проте, для зручності інтегрування доцільно розширити верхню межу до нескінченності, що майже не позначиться на точності обчислень, бо підінтегральна функція під впливом співмножника $f(E)$ дуже швидко спадає зі зростанням енергії. Тому, підставляючи в (3.4) вирази (3.10) і (3.3), отримаємо

$$n = \int_{E_c}^{\infty} g(E)f(E)dE = \int_{E_c}^{\infty} \frac{4\pi(2m_d^*/h^2)^{3/2}(E-E_c)^{1/2}}{\exp[(E-E_f)/k_B T] + 1} dE. \quad (3.11)$$

Проведемо заміну змінних

$$(E-E_c)/k_B T = \varepsilon, \quad (E_f-E_c)/k_B T = \xi, \quad (3.12)$$

де ε і ξ - приведені до одиниць $k_B T$ енергія електрона у зоні провідності і енергія Фермі, відраховані від дна зони провідності. Підставляючи (3.12) в (3.11) і вводячи величину

$$N_c = 2 (2\pi m_d^* k_B T/h^2)^{3/2} = 4,83 \cdot 10^{21} (m_d^*/m)^{3/2} T^{3/2} \text{ (М}^{-3}\text{)}, \quad (3.13)$$

яка іноді називається *ефективною густиною станів у зоні провідності*, отримаємо

$$n = N_c \Phi_{1/2}(\xi). \quad (3.14)$$

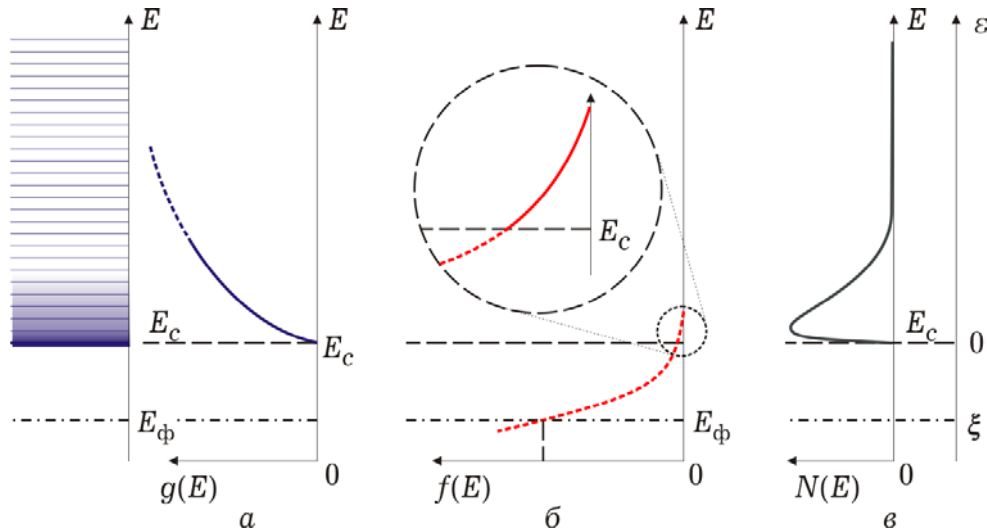


Рис. 3.3. Функції розподілу густини станів (а), ймовірності їх заповнення (б) і енергетичної густини електронів (в) біля дна зони провідності напівпровідника. Масштаб функції $N(E)$ збільшено порівняно з $g(E)$

Тут

$$\Phi_{1/2}(\xi) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{\exp(\varepsilon - \xi) + 1} \quad (3.15)$$

є так званий *інтеграл Фермі* порядку $1/2$. Цей інтеграл не обчислюється аналітично, а задається у вигляді таблиць або графіків. У додатку до гл. 3 представлена його залежність від аргументу ξ . Іноді зручно використовувати асимптотичні вирази для цього інтеграла:

$$\Phi_{1/2}(\xi) = \begin{cases} e^{\xi}, & -\infty < \xi < -1, \quad \text{а)} \\ e^{\xi} / (0,27 e^{\xi} + 1), & -1 < \xi < 5, \quad \text{б)} \\ 4\xi^{3/2} / 3\sqrt{\pi}, & 5 < \xi < \infty. \quad \text{в)} \end{cases} \quad (3.16)$$

Формула (3.16, а) придатна для невідроджених напівпровідників, а (3.16, в) для сильно вироджених напівпровідників і металів. Формулу (3.16, б) можна використовувати для розрахунків як невідроджених колективів, де вона дає вищу точність, ніж (3.16, а), так і за слабого виродження. Таким чином, вираз (3.14) є

універсальним і придатним для розрахунку концентрації електронів у зоні провідності як металів, так і напівпровідників.

Сумарна енергія усіх електронів у зоні провідності (відрахована від нижнього краю зони) може бути знайдена за формулою (3.5), яка у разі підстановки (3.12) набуває вигляду

$$E_{\Sigma} = \int_{E_c}^{\infty} (E - E_c) g(E) f(E) dE = (3 k_B T / 2) N_c \Phi_{3/2}(\xi). \quad (3.17)$$

Тут

$$\Phi_{3/2}(\xi) = \frac{4}{3\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{3/2} d\varepsilon}{\exp(\varepsilon - \xi) + 1}$$

є інтегралом Фермі порядку 3/2, який для невироджених колективів ($\xi < -1$) обчислюється аналітично і дорівнює e^{ξ} . Це означає, що середня енергія (3.6) електрона у зоні провідності

$$\bar{E} = 3k_B T / 2, \quad (3.18)$$

має теплову природу і обчислюється так саме, як і для молекул ідеального газу. Часто таке ж співвідношення використовують і для нерівноважних станів, використовуючи так звану *електронну температуру* T_e , яка може помітно відрізнитися від температури кристалічної ґратки T .

Багатодолинні напівпровідники, наприклад, GaAs, InP, GaSb (рис. 1.26), мають у зоні провідності декілька мінімумів дисперсійної кривої. Практично важливе значення мають випадки, коли верхні підзони (зазвичай L - або X- долини) знаходяться від абсолютного мінімуму зони провідності E_c (Γ -долини) на відстані у декілька десятих долей електронвольта. Для розрахунку концентрацій електронів у підзонах, використовуючи цей же метод для дводолинних моделей зонної структури, можна отримати

$$n = n_{\Gamma} + n_L = N_{c\Gamma} \Phi_{1/2}(\xi) + N_{cL} \Phi_{1/2}(\xi - \Delta_{\Gamma L} / k_B T), \quad (3.19)$$

де $N_{c\Gamma} = 2(2\pi m_{d\Gamma}^* k_B T / h^2)^{3/2}$, $N_{cL} = 2(2\pi m_{dL}^* k_B T / h^2)^{3/2}$ - ефективна густина станів, а $m_{d\Gamma}^*$ і m_{dL}^* - ефективні маси густини станів в Γ - і L-долинах відповідно.

Цікава особливість спостерігається у тих матеріалах, де $m_{d\Gamma}^* < m_{dL}^*$ і відповідно $N_{c\Gamma} < N_{cL}$. Завдяки більшій густині станів концентрація електронів у верхній долині за деяких умов може бути вищою, ніж у нижній. Заселеність верхньої долини зручно характеризувати функцією $b_L(T) = n_L/n = (n_L/n_\Gamma)/(1 + n_L/n_\Gamma)$, де для невідродженого напівпровідника

$$n_L/n_\Gamma = (m_{dL}^*/m_{d\Gamma}^*)^{3/2} \exp(-\Delta_{\Gamma L}/k_B T). \quad (3.20)$$

На рис. 3.4 представлена залежність $b_L(T)$ для GaAs ($m_{d\Gamma}^* = 0,067m$, $m_{dL}^* = 0,56m$, $\Delta_{\Gamma L} = 0,3$ eВ), звідки видно, що за кімнатної температури практично всі електрони знаходяться у нижній долині, за $T > 1500$ К, переважають електрони, що заповнюють L-долину. Треба також мати на увазі, що такі температури зазвичай досягаються не нагріванням матеріалу напівпровідника, а «розігріванням» електронного газу, наприклад, у сильному електричному полі. У цьому випадку в (3.20) треба використовувати електронну температуру T_e .

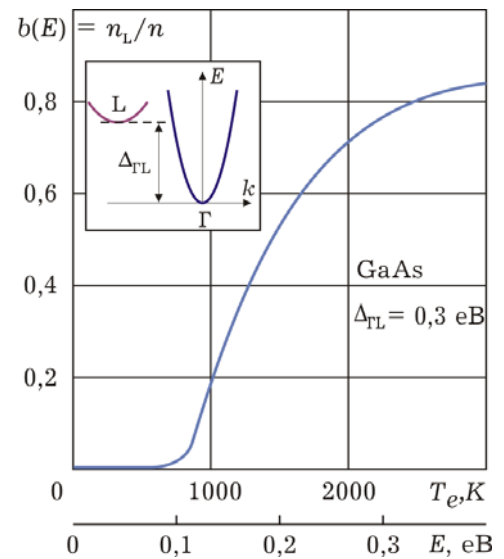


Рис. 3.4. Залежність відносної заселеності L-долини в GaAs від електронної температури або середньої енергії електронів

3.2.2. Дірки у валентній зоні

Метод розрахунку концентрації дірок у валентній зоні напівпровідника по суті такий самий, як і для електронів у зоні провідності. Густина станів поблизу стелі валентної зони описується функцією (рис. 3.5, а)

$$g(E) = 4\pi(2m_{pd}^*/h^2)^{3/2} (E_v - E)^{1/2}. \quad (3.21)$$

Ймовірність заповнення станів електронами у валентній зоні описується функцією Фермі-Дірака (рис. 3.5, б). Там же на вставці зображений фрагмент цієї функції для дірок $f_p(E) = 1 - f(E)$. Інтегрування функції N (рис. 3.5, в) у межах усієї валентної зони не має сенсу, оскільки функція густини станів (3.21) визначена тільки для верхніх станів зони. Зручніше розглядати вільні стани, де густина розподілу дірок зображена на рис. 3.5, в заштрихованою частиною, або на рис. 3.5, г у збільшеному масштабі.

Концентрація дірок у валентній зоні дорівнює інтегралу від функції розподілу $N_p(E) = g(E) f_p(E)$ по всіх станах валентної зони:

$$p = \int_{-\infty}^{E_v} g(E) f_p(E) dE. \quad (3.22)$$

Тут, подібно до попереднього інтеграла, нижня межа інтегрування розширена до нескінченності, що практично не впливає на результат. За такої умови інтеграл (3.22) за структурою співпадає з (3.11) з точністю до напрямку відліку енергії. Повна тотожність досягається підстановками, аналогічними (3.12):

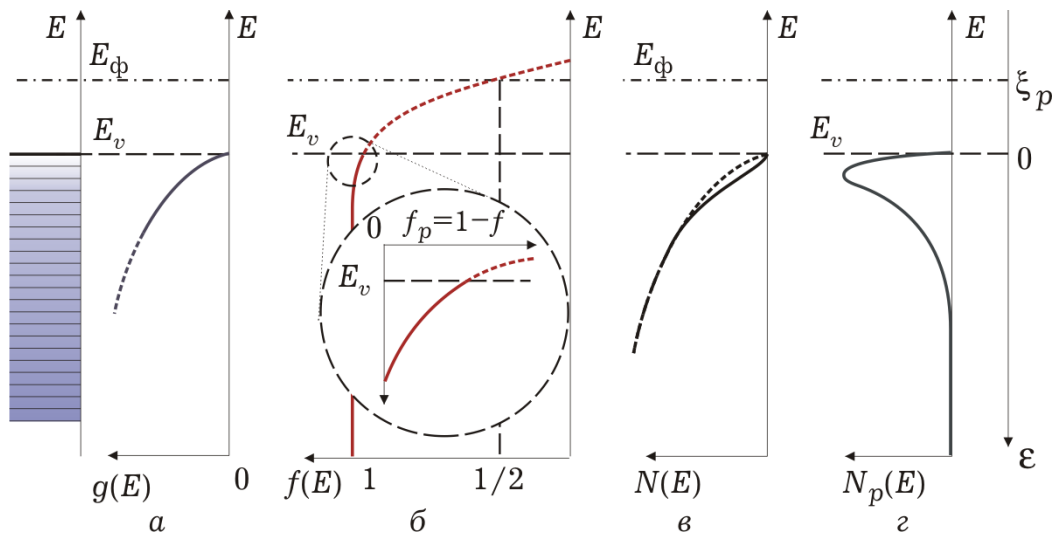


Рис. 3.5. Функції розподілу густини станів (а), ймовірності їх заповнення (б) і енергетичної густини електронів (в) і дірок (г) біля стелі валентної зони напівпровідника. Масштаб функції $N_p(E)$ збільшено порівняно з $N(E)$

$$(E_v - E)/k_B T = \varepsilon, (E_v - E_\phi)/k_B T = \xi_p, \quad (3.23)$$

з якими результат інтегрування набуває вигляду

$$p = N_v \Phi_{1/2}(\xi_p), \quad (3.24)$$

де

$$N_v = 2(2\pi m_{pd}^* k_B T / h^2)^{3/2} = 4,83 \cdot 10^{21} (m_{pd}^* / m)^{3/2} T^{3/2} \text{ (м}^{-3}\text{)} \quad (3.25)$$

- ефективна густина станів у валентній зоні, а інтеграл Фермі $\Phi_{1/2}(\xi_p)$ відрізняється від (3.15), тим що тут ξ_p є рівнем Фермі, відрахованим від стелі валентної зони і вираженим в одиницях $k_B T$. Асимптотичні вирази для інтеграла Фермі $\Phi_{1/2}(\xi_p)$, аналогічні відповідним виразам (3.16) для інтеграла $\Phi_{1/2}(\xi)$.

Отримані співвідношення (3.14) і (3.24) для розрахунку концентрації рухливих носіїв у наближенні ефективної маси є універсальними, бо у процесі їх виводу приймалося тільки одне слабе припущення про розширення меж інтегрування. Найбільш поширеним є наближення (3.16, а), згідно з яким

$$n = N_c \exp [(E_\phi - E_c) / k_B T], \quad (3.26)$$

$$p = N_v \exp [(E_v - E_\phi) / k_B T]. \quad (3.27)$$

Ці вирази придатні для невідроджених напівпровідників, оскільки отримані за умови $\xi, \xi_p < -1$, яка означає, що рівень Фермі віддалений від краю відповідної дозволеної зони більше, ніж на $k_B T$. У цьому разі в інтегралі Фермі можна знехтувати одиницею у знаменнику, що фактично означає перехід до статистики Максвелла-Больцмана.

З отриманих виразів виходять два корисні висновки. Узявши відношення останніх виразів, враховуючи (3.13) і (3.25), отримаємо

$$p/n = (m_{pd}^* / m_d^*)^{3/2} \exp\{2 [(E_c + E_v) / 2 - E_\phi] / k_B T\}. \quad (3.28)$$

Звідси бачимо, що співвідношення між концентраціями рухливих електронів і дірок залежить від відношення їх ефективних мас і особливо сильно від положення рівня Фермі відносно середини забороненої зони $(E_c + E_v) / 2$. Концентрація рухливих носіїв більша в тій зоні, ближче до якої розташований рівень Фермі. Останнє твердження придатне для рівних ефективних мас електронів і дірок, але вони відрізняються один від одного не більше, ніж на порядок, тоді як за рахунок навіть невеликого зміщення рівня Фермі можна отримати зміну концентрацій на декілька порядків. Цей перерозподіл концентрацій рухливих носіїв між зонами може бути здійснений введенням у напівпровідник домішок, оскільки

положення рівня Фермі істотно залежить від виду і концентрації домішки, що вводиться. Проте добуток концентрацій рухливих електронів і дірок не залежить від положення рівня Фермі і за даної температури залежить тільки від ширини $\Delta E = E_c - E_v$ забороненої зони напівпровідника і ефективних мас електронів і дірок :

$$n p = N_c N_v \exp(-\Delta E/k_B T) = n_i p_i = n_n p_n = n_p p_p. \quad (3.29)$$

Першу рівність у цьому співвідношенні, яке безпосередньо виходить з множення (3.26) на (3.27), називають *законом діючих мас*. Продовження цієї рівності означає, що добуток концентрацій носіїв не залежить від виду і концентрації домішок (індекси i , n і p означають приналежність до власного, донорного чи акцепторного напівпровідника відповідно). Тому введення у напівпровідник, наприклад, донорних домішок збільшує у зоні провідності кількість електронів і автоматично відповідно до (3.29) зменшує кількість дірок у валентній зоні.

Закон діючих мас придатний для невідроджених напівпровідників, що знаходяться у стані термодинамічної рівноваги. Іноді (3.26) і (3.27) можна використовувати також у тих випадках, коли окрім теплового існує і деякий інший спосіб збудження носіїв, що порушує рівновагу. У цих випадках статистика Фермі-Дірака непридатна, але вирази (3.26) і (3.27) можна використовувати, якщо замість рівня Фермі формально ввести так звані *квазірівні* E_{fn} і E_{fp} . Оскільки для електронів і дірок вони, як правило, різні, то

$$np = n_0 p_0 \exp [(E_{fn} - E_{fp}) / k_B T]. \quad (3.30)$$

Тут через n_0 і p_0 позначені рівноважні концентрації електронів і дірок, розраховані відповідно до (3.26) і (3.27). Останній вираз показує, що закон діючих мас не виконується для нерівноважних носіїв заряду. Крім того, корисно відзначити, що у разі, коли нерівноважні концентрації електронів і дірок перевищують їх рівноважні значення, то $E_{fn} > E_f > E_{fp}$. Це спостерігається, якщо зовнішнє збудження, викликаючи порушення термодинамічної рівноваги, призводить до появи *надлишкових носіїв*. Якщо ж порушення рівноваги інтенсифікує процеси рекомбінації, то нерівноважні концентрації зменшуються і електронний квазірівень Фермі знижується, а дірковий підвищується відносно E_f .

3.2.3. Електрони у кристалах низької розмірності

Для одновимірних і двовимірних кристалів спостерігається усупільнення електронів і відповідно зонний спектр лише у тих напрямках, де виконується умова трансляційної симетрії. Для інших напрямків електронні стани залишаються локалізованими і зберігається дискретний спектр. Обчислення кількості електронів у відповідних зонах може бути проведене таким саме методом, що і для об'ємних (3D) кристалів. Але добуток функції (1.27) або (1.29) для густини енергетичних станів на функцію Фермі-Дірака дає енергетичну густину електронів у розрахунку на одиницю площі у двовимірних кристалах і на одиницю довжини в одновимірних. Тому інтегрування цієї густини по усій дозволеній зоні дозволяє розраховувати значення відповідних концентрацій. Так, в 2D-кристалах інтегрування

$$\int_{E_i}^{\infty} g_i(E) f(E) dE = n_{is} \quad (3.31)$$

визначає значення *поверхневої концентрації* електронів у двовимірній зоні, що відповідає кожному дискретному стану E_i . Підстановка виразів (1.31) для енергетичної густини двовимірних станів і (3.3) для функції Фермі-Дірака, а також заміна змінних $(E - E_i)/k_B T = \varepsilon$, $(E_F - E_i)/k_B T = \xi$, подібна до (3.12), приводить вираз (3.31) до форми

$$n_{is} = (4\pi m_{xy}^* k_B T / h^2) \int_0^{\infty} \frac{d\varepsilon}{\exp(\varepsilon - \xi) + 1}. \quad (3.32)$$

Такий інтеграл називається інтегралом Фермі нульового порядку $\Phi_0(\xi)$. Хоча його значення і табульовані, як і для подібних інтегралів інших порядків (див. додаток до гл. 3), він може бути обчислений у явному вигляді

$$\Phi_0(\xi) = \int_0^{\infty} \frac{d\varepsilon}{\exp(\varepsilon - \xi) + 1} = \ln \frac{\exp \xi}{\exp(\varepsilon - \xi) + 1} \Big|_0^{\infty} = \ln(1 + e^{\xi}). \quad (3.33)$$

Позначаючи $N_{c2} = 4\pi m_{xy}^* k_B T / h^2$, отримуємо вираз для поверхневої концентрації

$$n_{is} = N_{c2} \ln\{1 + \exp[(E_F - E_i)/k_B T]\}. \quad (3.34)$$

За малих значень аргументу ($\xi < -1$) інтеграл $\Phi_0(\xi)$, як і подібні (рис. Д. 3.1), апроксимується функцією e^ξ . Тому для невиродженого колективу у даній зоні зручніше користуватися виразом

$$n_{is} = N_{c2} \exp[(E_\phi - E_i)/k_B T]. \quad (3.35)$$

В *одновимірних* (1D) кристалах концентрація електронів n_l на одиницю довжини у зоні, що лежить над станом E_{ij} , квантованим по двох напрямках, визначається інтегруванням функції розподілу по усій зоні:

$$\int_{E_{ij}}^{\infty} g_{ij}(E) f(E) dE = n_{ijl}. \quad (3.36)$$

Підстановка одновимірної густини станів (1.27) і подібна заміна змінних $(E - E_{ij})/k_B T = \varepsilon$, $(E_\phi - E_{ij})/k_B T = \xi$ призводять до виразу

$$n_{ijl} = N_{c1} \Phi_{-1/2}(\xi), \quad (3.37)$$

де

$$\Phi_{-1/2}(\xi) = \frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{-1/2} d\varepsilon}{\exp(\varepsilon - \xi) + 1} \quad (3.38)$$

є інтеграл Фермі порядку $-1/2$ (див. додаток до гл. 3), а $N_{c1} = (2\pi m_x^* k_B T / h^2)^{1/2}$ - так звана ефективна густина станів в одновимірній зоні. Як і для інших розмірностей, для невироджених електронів у цій зоні ($\xi < -1$) інтеграл $\Phi_{-1/2}(\xi)$ апроксимується функцією e^ξ і для розрахунку концентрації зручніше користуватися виразом

$$n_{ijl} = N_{c1} \exp[(E_\phi - E_{ij})/k_B T]. \quad (3.39)$$

Часто цікавим є питання про середню енергію електронів у зонах кристалів низької розмірності. Використовуючи для цього вираз (3.6), після підстановки відповідних функцій розподілу неважко отримати для кристала n -ї розмірності

$$\bar{E} = \frac{n}{2} k_B T \frac{\Phi_{n/2}(\xi)}{\Phi_{n/2-1}(\xi)}. \quad (3.40)$$

Для невироджених напівпровідників інтеграли Фермі дорівнюють e^{ξ} і середня енергія визначається тільки тепловим рухом і дорівнює $nk_B T/2$, по $k_B T/2$ на кожен ступінь свободи руху електронів уздовж кристала, що знаходиться відповідно до статистики Максвелла-Больцмана.

3.2.4. Заповнення локалізованих станів

Загальні принципи статистики Фермі-Дірака, які використовуються для опису заповнення неперервних або квазінеперервних електронних станів у дозволених зонах, можуть бути використані і для локалізованих станів з дискретним енергетичним спектром. Строго кажучи, у цьому випадку розподіл Фермі-Дірака непридатний, оскільки він отриманий за умови, що кожен стан може бути зайнятий тільки двома електронами з протилежними спінами. Але, наприклад, у донорному енергетичному стані, утвореному п'ятивалентним домішковим атомом, може знаходитися тільки один електрон з будь-яким напрямом спіна, тому такий стан вважається двократно виродженим. Подібна ситуація має місце і для інших домішок, у тому числі багатозарядних, і може бути описана за допомогою статистичного розподілу Гіббса, який враховує кратність виродження локалізованих станів. Проте для розуміння основних механізмів генерації носіїв заряду обмежимося простішою функцією Фермі-Дірака з невеликою втратою точності у кількісних розрахунках.

Відповідно до фізичного сенсу функції $f(E)$ ймовірність заповнення електроном якогось локалізованого енергетичного стану E_t дорівнює значенню функції для цього значення енергії:

$$f(E_t) = 1/\{\exp [(E_t - E_{\phi})/k_B T] + 1\}. \quad (3.41)$$

Якщо концентрація таких центрів дорівнює N_t , то з них заповненими будуть

$$N_t^n = N_t f(E_t) = N_t/\{\exp [(E_t - E_{\phi})/k_B T] + 1\}, \quad (3.42)$$

а інші $(N_t - N_t^n)$ - вільними від електронів

$$N_t^p = N_t - N_t^n = N_t [1 - f(E_t)] = N_t/\{\exp [(E_{\phi} - E_t)/k_B T] + 1\}. \quad (3.43)$$

У тому випадку, коли локалізовані стани утворені п'ятивалентною донорною домішкою з концентрацією N_d , то (3.43) визначає кількість додатньо іонізованих донорів

$$N_d^+ = N_d / \{ \exp [(E_f - E_d) / k_B T] + 1 \}, \quad (3.44)$$

а інші домішкові атоми утримуватимуть свої п'яті «зайві» електрони і залишатимуться нейтральними. Якщо ж у напівпровіднику присутні тривалентні акцепторні атоми з концентрацією N_a , то відповідно до (3.42) від'ємно зарядженими з них будуть

$$N_a^- = N_a / \{ \exp [(E_a - E_f) / k_B T] + 1 \}, \quad (3.45)$$

а інші ($N_a - N_a^-$) залишатимуться нейтральними.

У таких випадках треба пам'ятати, що іонізація донорних атомів супроводжується збудженням «зайвого» електрона у зону провідності, а іонізація акцепторів - збудженням валентного електрона у вакантний стан і відповідно утворенням вакансії в системі валентних зв'язків, тобто дірки. Це справедливо за наявності у напівпровіднику тільки одного виду домішки.

На закінчення відзначимо, що отримані вирази для розрахунку концентрацій рухливих носіїв і іонізованих атомів у межах прийнятих припущень є універсальними. Але скористатися ними у тій чи іншій конкретній ситуації можна, тільки заздалегідь розрахувавши положення рівня Фермі. У цьому і полягає задача визначення статистичних властивостей рухливих носіїв заряду для даного металу або напівпровідника. Знайти положення рівня Фермі можна, якщо на отримані вирази накласти деяку додаткову умову. Для металів такою умовою є відоме значення концентрації електронів у зоні провідності. Для напівпровідників воно формулюється як *рівняння електронейтральності*

$$n + N_a^- = p + N_d^+ \quad (3.46)$$

суть якого полягає у тому, що процеси генерації і рекомбінації рухливих носіїв не порушують електричної нейтральності речовини. Ця фізично проста умова справедлива для рівноважних станів за відсутності електричних полів. Надалі ми побачимо, що наявність поля може викликати *локальне* порушення електронейтральності речовини за рахунок просторового перерозподілу рухливих носіїв.

3.3. ПРОВІДНІСТЬ МЕТАЛІВ

У металевих кристалах усі усупільнені електрони енергетично розташовані у зоні провідності, а їх концентрація визначається концентрацією атомів і кількістю усупільнених валентних електронів від кожного атома. Розподіл електронів по енергетичних станах зони провідності описується статистикою Фермі-Дірака, тому концентрацію електронів можна визначити за формулою (3.11). Вона значно спрощується за нульової температури, оскільки для $T = 0$ електронами зайняті усі стани тільки до рівня Фермі, тобто $f(E) = 1$ для $E < E_{\text{Ф}}$ і $f(E) = 0$ для $E > E_{\text{Ф}}$. Тому формулу (3.11), замінюючи верхню межу інтегрування на $E_{\text{Ф}}(0)$, можна записати у вигляді

$$n = \int_0^{\infty} g(E)f(E)dE = \int_0^{E_{\text{Ф}}(0)} g(E) \cdot 1 \cdot dE = \frac{8\pi}{3} \left[2m^* E_{\text{Ф}}(0) / h^2 \right]^{3/2}.$$

Звідси за відомої концентрації електронів n знаходимо енергію Фермі для нульової температури

$$E_{\text{Ф}}(0) = \frac{h^2}{2m^*} \left(\frac{3n}{8\pi} \right)^{2/3}. \quad (3.47)$$

Це є найбільше значення енергії, яку може мати електрон у металі за $T = 0$. Такий саме результат для $E_{\text{Ф}}(0)$ можна отримати з (3.14), якщо використовувати асимптотичний вираз для інтеграла Фермі (3.16, в).

Оцінюючи енергію Фермі, треба мати на увазі, що для типових металів ефективна маса електронів близька до звичайної, тому

$$E_{\text{Ф}}(0) \text{ (eV)} = 0,365 \cdot 10^{-18} n^{2/3} \text{ (м}^{-3}\text{)}, \quad (3.47, a)$$

що для властивих металам концентрацій електронів близько $10^{28} \dots 10^{29} \text{ м}^{-3}$ складає декілька електронвольт.

Для температур, відмінних від нуля, наближений вираз для рівня Фермі $E_{\text{Ф}}(T)$ можна отримати, скориставшись розкладанням інтеграла Фермі $\Phi_{1/2}(\xi)$ у формулі (3.15) у ступеневий ряд для великого додатнього параметра ξ . Таке наближення можливе, оскільки для існуючих у металах концентрацій електронів параметр

$\xi = E_{\text{Ф}}/k_{\text{Б}}T$ є досить великим і можна обмежитися двома першими членами ряду, що сходиться:

$$E_{\text{Ф}}(T) = E_{\text{Ф}}(0) \left[1 - \frac{\pi^2}{12} \left(\frac{k_{\text{Б}}T}{E_{\text{Ф}}(0)} \right)^2 \right]. \quad (3.48)$$

Чисельні розрахунки показують, що із зростанням температури енергія Фермі зменшується відносно $E_{\text{Ф}}(0)$ на дуже малу величину. Так, для $T \approx 10^3$ К рівень Фермі зменшується всього на $10^{-3} \cdot 10^{-4}$ еВ, і тільки за температур близько $10^4 \cdot 10^5$ К рівень Фермі зміщується в область від'ємних значень, а електронний газ стає невиродженим.

На рис. 3.6 показані функції розподілу густини станів (а), ймовірність заповнення станів (б) і густини електронів (в) у зоні провідності металу для $T = 0$ і $T > 0$.

Треба відзначити, що для температур, за яких метали знаходяться у твердому стані, електронний газ у металі є виродженим і розподіл електронів по енергіях визначається не тепловим рухом, а принципом Паулі. Підтвердженням цього положення може служити значення середньої енергії одного електрона, яке можна отримати з формули (3.6). Для цього скористаємося розкладанням підінтегральних функцій у ряди і обмежимося першими членами.

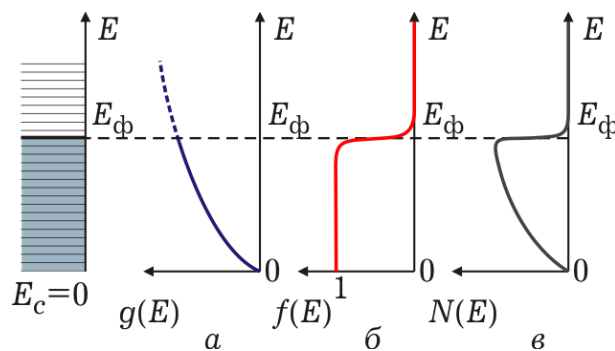


Рис. 3.6. Функції розподілу густини станів (а), ймовірності їх заповнення (б) і енергетичної густини електронів (в) у зоні провідності металів

У результаті отримаємо

$$\bar{E} = E_{\text{Ф}}(0) \cdot \frac{3}{5} \left[1 + \frac{5\pi^2}{12} \left(\frac{k_{\text{Б}}T}{E_{\text{Ф}}(0)} \right)^2 \right]. \quad (3.49)$$

Перший доданок у цьому виразі характеризує середню енергію електрона за нульової температури і відповідно називається «нульовою» енергією, яка має чисто квантову природу, :

$$\bar{E}_0 = 3E_{\text{ф}}(0)/5. \quad (3.50)$$

Маючи на увазі випадковий характер руху електронів, цю енергію розглядають як кінетичну енергію хаотичного руху, який існує і для $T = 0$, з середньою швидкістю

$$v_0 = \sqrt{2\bar{E}_0 / m^*} = \sqrt{6E_{\text{ф}}(0)/5m^*}. \quad (3.51)$$

Другий доданок, який залежить від температури, описує середню теплову енергію електрона

$$\bar{E}_T = \pi^2(k_B T)^2/4E_{\text{ф}}(0). \quad (3.52)$$

Порівняння \bar{E}_0 і \bar{E}_T показує, що навіть за температур, близьких до температури плавлення металів, теплова енергія електрона складає $10^{-3}..10^{-4}$ від «нульової» енергії, а відповідна теплова швидкість - одиниці відсотків від «нульової» швидкості. Така слабка сприйнятливості електронів у металах до нагрівання виходить з самого характеру функції розподілу електронів. Ця обставина пояснює, між іншим, малий внесок теплоємності електронного газу у загальну теплоємність металів.

За класичною теорією питома теплоємність металів має бути вищою за теплоємність діелектриків, оскільки вона, окрім граткової теплоємності, включає ще і теплоємність електронного газу. Але експериментальні дослідження не виявили істотної різниці у значеннях питомої теплоємності металів і діелектриків. Пояснення цьому було дано тільки з використанням квантової статистики. Поведінка функції розподілу Фермі-Дірака (рис. 3.1) зі зміною температури показує, що теплову енергію порядку $k_B T$ отримує дуже мала частина усіх електронів, яка становить величину порядку $k_B T/E_{\text{ф}}$ від загальної їх кількості, і в перерахунку на один електрон має значення близько $(k_B T)^2/E_{\text{ф}}$, що значно нижче за їх повну енергію (3.49). Саме тому електронний газ, який не є теплоємним, дуже слабо впливає на загальну теплоємність металів.

Неспроможність *класичних* представлень проявилася передусім у питанні про температурну залежність електропровідності металів, для якої експеримент показав, що у широкому температурному діапазоні $\sigma \sim T^{-1}$. За класичною теорією причиною розсіювання електронів у металах є їх зіткнення з атомами кристалічної ґратки, тому і довжина вільного пробігу електронів \bar{l} має бути близькою до величини міжатомної відстані і складати декілька ангстремів, тоді як чисельні оцінки дають значення близько сотень періодів ґратки. Застосування для електронного газу класичної статистики Максвелла-Больцмана, відповідно до якої хаотичний рух визначається тільки тепловим, а його швидкість $v_{\text{хаот}} \sim T^{1/2}$, дає функціональну залежність $\sigma \sim \bar{\tau} = \bar{l} / v_{\text{хаот}} \sim T^{-1/2}$.

Квантова теорія, застосовуючи для електронного колективу статистику Фермі-Дірака, припускає, що за звичайних умов хаотичний рух фактично не залежить від температури, а його швидкість $v_{\text{хаот}} \approx v_0$ визначається за (3.51) практично тільки нульовою енергією. Друге принципове твердження квантової теорії пов'язане з тим, що зіткнення електронів відбуваються не безпосередньо з атомами кристалічної решітки, а з її *дефектами*, які виникають у тому числі і в результаті теплових коливань. В чистих металах зазвичай переважає розсіювання на акустичних фононах, кількість яких, особливо довгохвильових, за формулою Планка пропорційна температурі $N_{\text{ф}} \approx k_{\text{Б}}T/\hbar\omega$. У цьому випадку середня довжина вільного пробігу за формулою (2.6) обернено пропорційна $N_{\text{ф}}$ і відповідно $\bar{l} \sim T^{-1}$. Тому $\sigma \sim \bar{\tau} = \bar{l} / v_0 \sim T^{-1}$, що відповідає експериментальним фактам. Крім того, кількісна оцінка довжини пробігу для багатьох металів дає величину близько 10^{-7} м, що на два порядки перевищує міжатомну відстань. Це є додатковим аргументом того, що причиною розсіювання не є зіткнення з атомами ґратки.

Корисно звернути увагу на те, що енергетичний розподіл електронів у зоні провідності металів є подібним до розподілу густини молекул рідини по висоті. Процес переходу електронів на вищі енергетичні рівні у разі нагрівання аналогічний процесу випарування молекул з рідини. З цієї точки зору більш доречною для електронів у металі була б назва «електронна рідина» (рівень Фермі електронів відповідає висоті рівня поверхні рідини). Така аналогія характерна для електронного колективу в цілому, а та частина електронів, енергія яких перевищує рівень Фермі на

декілька одиниць $k_B T$, аналогічна парі над поверхнею рідини і може бути названа «електронним газом», до якого може бути застосовна статистика Максвелла-Больцмана.

3.4. ВЛАСНА ПРОВІДНІСТЬ НАПІВПРОВІДНИКІВ

У напівпровідниках усупільнені електрони беруть участь у створенні ковалентних зв'язків між атомами і, будучи пов'язаними обмінними силами з кристалічною ґраткою, не можуть брати участь в електропровідності. Щоб електрони стали рухливими в електричному полі, вони повинні здолати сили зв'язків, отримавши певну енергію від зовнішніх джерел, і стати вільними. Тут ми розглядаємо утворення (*генерацію*) рухливих носіїв, яка виникає під дією тільки теплового руху.

Процеси теплової генерації у напівпровідниках на мові енергетичних діаграм можна описати таким чином. За нульової температури усі стани валентної зони повністю заповнені електронами, а з підвищенням температури частина електронів, енергія яких стає більшою за ширину забороненої зони, переходить у зону провідності. У валентній зоні створюється така ж кількість дірок. Процеси генерації супроводжуються зворотними переходами електронів у валентну зону і наступним їх зникненням як рухливих носіїв, тобто *рекомбінацією* з дірками. За умови термодинамічної рівноваги швидкість теплової генерації електронно-діркових пар дорівнює швидкості їх рекомбінації, а концентрації електронів і дірок у відповідних зонах встановлюються статистично постійними або *рівноважними*.

Оскільки у бездомішкових (власних) напівпровідниках утворення і зникнення носіїв завжди відбувається парами, то концентрації електронів і дірок рівні між собою, і $n_i = p_i$ (індекс i означає приналежність до власного напівпровідника). Ця рівність і є рівнянням електронейтральності, суть якого в тому, що процеси генерації рухливих носіїв не порушують електричної нейтральності речовини. Тоді, використовуючи (3.26) і (3.27), для невідродженого напівпровідника можна записати

$$N_c \exp [(E_F - E_c)/k_B T] = N_v \exp [(E_v - E_F)/k_B T]. \quad (3.53)$$

Логарифмуючи обидві частини цієї рівності і обчислюючи значення енергії Фермі, отримуємо

$$E_{\Phi} = \frac{E_c + E_v}{2} + \frac{k_B T}{2} \ln \frac{N_v}{N_c} = E_i + \frac{3}{4} k_B T \ln \frac{m_{pd}^*}{m_d^*}, \quad (3.54)$$

де $E_i = (E_c + E_v)/2$ - середина забороненої зони.

Аналіз формули (3.54) показує, що із зростанням температури рівень Фермі лінійно зміщується від середини забороненої зони для $T = 0$ у бік зони з більш легкими носіями (рис. 3.7). Оскільки ефективні маси густини станів m_d^* і m_{pd}^* не відрізняються більш ніж на порядок, то це зміщення не перевищує $(1..2) k_B T$. Наприклад, для $T = 300$ К рівень Фермі розташований в германії на ~ 8 меВ, а в кремнії на ~ 12 меВ нижче за середину забороненої зони.

Зі зміною температури за лінійним законом (1.34) також змінюється і ширина забороненої зони. У таких поширених напівпровідниках, як Ge, Si і GaAs, температурний коефіцієнт α є від'ємним і дорівнює приблизно 10^{-4} еВ/К. Тому температурні зміщення рівня Фермі і країв забороненої зони співмірні. У вузькозонних напівпровідниках з дуже різними ефективними масами носіїв із зростанням температури рівень Фермі може наблизитися до дна зони провідності або навіть опинитися в ній. Так, наприклад, в InSb $\Delta E(T) = 0,26 - 2,7 \cdot 10^{-4} T$ (еВ), $m_{pd}^* \approx 10 m_d^*$, і вже для $T \geq 440$ К рівень Фермі лежить у зоні провідності. Це призводить до виродження газу більш легких носіїв, і для розрахунку концентрацій n і p необхідно користуватися загальними виразами (3.14) і (3.24).

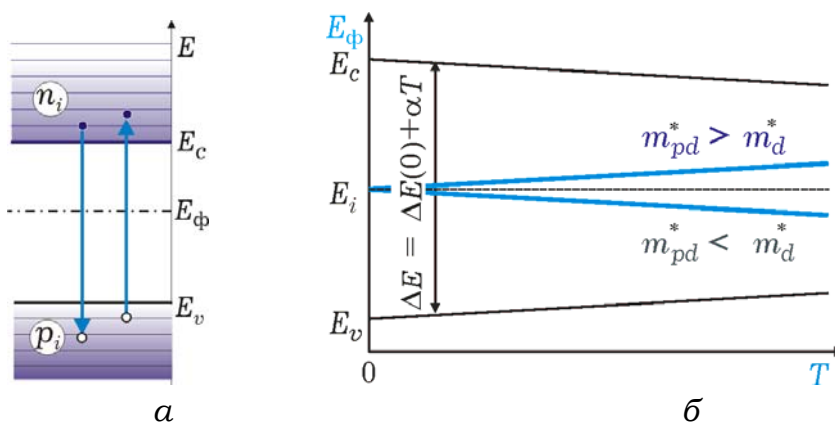


Рис. 3.7. Зміна положення рівня Фермі у власному напівпровіднику в залежності від температури. Зображено також температурну зміну ширини забороненої зони

Якщо напівпровідник залишається невідродженим, то залишаються у силі і вирази (3.26) і (3.27), які для знайденого положення рівня Фермі (3.53) спрощуються і набувають вигляду

$$n_i = p_i = \sqrt{N_c N_v} \exp(-\Delta E / 2k_B T). \quad (3.55)$$

На рис. 3.8 представлена залежність концентрацій електронів і дірок у германії від температури, зображена у різних координатах: а) лінійний масштаб найбільш наочний з точки зору міри впливу температури на концентрацію; б) логарифмічний дозволяє на одному графіку зображати величини, які відрізняються на декілька порядків; в) залежність логарифма концентрації від зворотної температури $1/T$ описується функцією:

$$\ln n_i = \ln(N_c N_v) / 2 - \Delta E / 2k_B T \approx \text{const} - (\Delta E / 2k_B) \cdot (1/T). \quad (3.56)$$

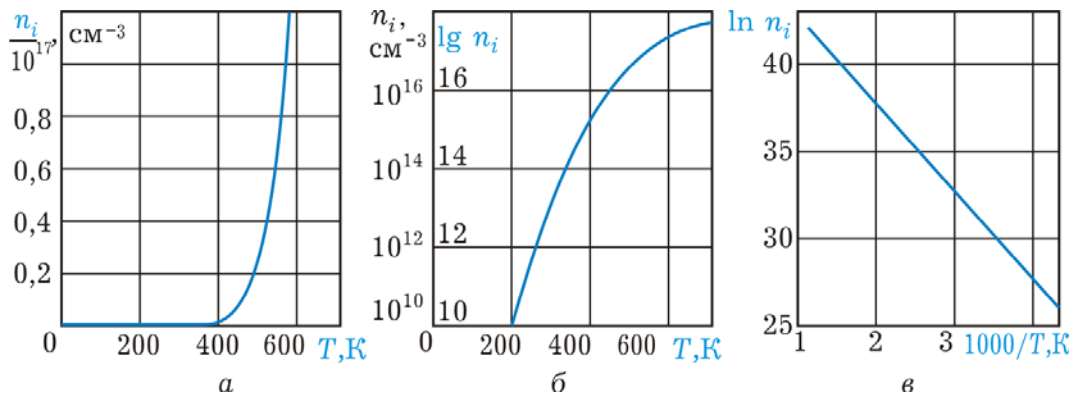


Рис. 3.8. Залежність концентрації електронів і дірок у чистому германії від температури, зображена у різних координатах

Якщо враховувати тільки вказану температурну залежність, то рівність (3.56) описує пряму лінію з кутовим коефіцієнтом $\Delta E / 2k_B$. Перевага останнього способу представлення залежності (3.55) полягає в тому, що, по-перше, лінійність побудованого за результатами експерименту графіку в координатах $\ln n_i = f(1/T)$ підтверджує правильність розрахункової формули (3.55) і, по-друге, дозволяє за нахилом прямої визначити ширину забороненої зони напівпровідника.

У разі більш строгого аналізу температурної залежності $n_i(T)$ треба враховувати, що N_c і N_v залежать від температури згідно з

законом $\sim T^{3/2}$, і що від температури залежить ширина забороненої зони $\Delta E = f(T)$.

Залежність питомої електропровідності від температури визначається в основному концентраційною залежністю і має вигляд

$$\sigma_i = en_i\mu_n + ep_i\mu_p = \sigma_0 \exp(-\Delta E/2k_B T), \quad (3.57)$$

де коефіцієнт $\sigma_0 = e\sqrt{N_c N_v}(\mu_n + \mu_p) = \text{const}(T)$, оскільки залежність рухливості від температури у разі розсіювання електронів і дірок на теплових коливаннях ґрат пропорційна $T^{-3/2}$ і компенсує температурну залежність N_c і N_v .

3.5. ДОМІШКОВА ПРОВІДНІСТЬ

У домішкових напівпровідниках рухливі електрони і дірки можуть з'являтися не лише у результаті порушення валентних зв'язків між атомами основного напівпровідника, але і за рахунок іонізації домішкових атомів. Іонізація донорних атомів призводить до появи у зоні провідності вільних електронів, а іонізація акцепторних до створення дірок у валентній зоні.

На енергетичній діаграмі (рис. 3.9) ці процеси виглядають таким чином. За нульової температури у донорному напівпровіднику електронами зайняті усі стани валентної зони і домішкового рівня, а в акцепторному - усі стани у валентній зоні. З підвищенням температури під дією теплового збудження відбувається перехід електронів з донорних локальних рівнів у зону провідності (1), перехід електронів з валентної зони на акцепторний рівень (2), а також міжзонні переходи (3), за яких генеруються як електрони, так і дірки.

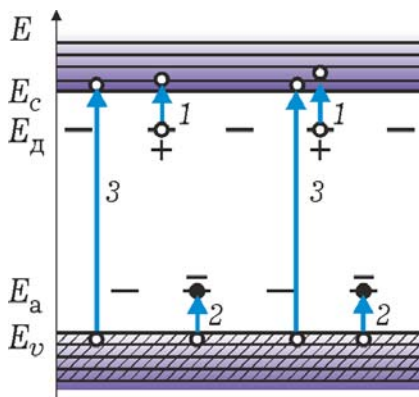


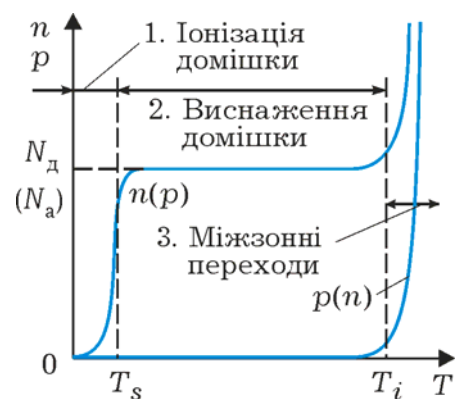
Рис. 3.9. Механізми теплової генерації рухливих носіїв у домішкових напівпровідниках. Світлими кружками умовно позначено енергетичний стан дірок, темними - електронів. Знаки + і - означають, що в результаті генерації носіїв домішкові атоми перетворюються відповідно в іонізовані донори і акцептори

Як і для власного напівпровідника, ці переходи супроводжуються зворотними переходами. Так, у донорному напівпровіднику електрони із зони провідності повертаються на донорні рівні і у валентну зону, а в акцепторному із зони провідності і з акцепторного рівня - у валентну зону, де відбувається їх рекомбінація з дірками. За постійної температури процеси генерації і рекомбінації рухливих носіїв урівноважуються і їх концентрації встановлюються статистично постійними.

Інтенсивність різних переходів залежить від температури і необхідної зміни енергії. Оскільки енергія іонізації домішкових атомів ($\Delta E_d \sim 0,01$ еВ) значно менша за енергію, необхідну для розриву валентних зв'язків власного напівпровідника (ΔE - десятки долі, одиниці електронвольт), то для температур в одиниці і десятки кельвінів помітною буде тільки іонізація домішкових атомів, а міжзонні переходи, інтенсивність яких залежить від ширини забороненої зони, особливо істотні за сотень кельвінів. Такі міркування дозволяють відносно просто описати якісну картину зміни концентрацій із зростанням температури для напівпровідника з одним видом домішки (рис. 3.10).

Розглянемо, наприклад, донорний напівпровідник або напівпровідник типу n (надалі усі величини для донорного напівпровідника позначатимемо індексом n). За низьких температур відбувається іонізація донорних атомів, внаслідок чого з'являються рухливі носії тільки одного виду - електрони у зоні провідності. Їх концентрація дорівнюватиме концентрації нерухомих додатньо заряджених донорів: $n_n = N_d^+$. Іноді ділянку 1 на рис. 3.10 називають областю виморожування домішок, що виправдано для опису процесів, які відбуваються у разі охолодження напівпровідника.

Рис. 3.10. Типові температурні залежності концентрації електронів і дірок у напівпровіднику, що містить донорні (акцепторні) домішки. Позначено три характерних ділянки цих залежностей



Зі зростанням температури зростає ймовірність іонізації донорів і за температур у декілька десятків кельвінів практично усі донори будуть іонізовані. Якщо за цих температур кількість домішок N_d помітно перевищує кількість рухливих носіїв у власному напівпровіднику, то концентрація електронів у зоні провідності буде рівною концентрації домішки $n_n = N_d$. Така ситуація називається *виснаженням домішки*.

Лише за температур у кілька сотень кельвінів стають помітними *міжзонні переходи* (3 на рис. 3.9). Такі переходи призводять до утворення пари рухливих носіїв: електрона і дірки. Тому концентрація електронів у зоні провідності визначається не лише кількістю домішок, але і кількістю створених електронно-діркових пар: $n_n = N_d + p_n$.

Подальше підвищення температури вище за деяке значення T_i (рис. 3.10) призводить до інтенсивної генерації електронно-діркових пар, і їх кількість може помітно перевищити величину N_d . У цьому випадку $n_n \approx p_n \approx n_i$ і процеси зміни концентрації носіїв будуть такими ж, як і у власному напівпровіднику.

Така якісна оцінка процесів у донорному напівпровіднику дозволяє на ділянці низьких і середніх температур визначити електрони як *основні носії*, а дірки як *неосновні*.

В акцепторному напівпровіднику типу p основними носіями будуть дірки, які створюються у валентній зоні за рахунок переходу електронів у домішкові акцепторні стани (іонізація домішок) і в зону провідності (міжзонні переходи), а неосновними - електрони у зоні провідності. Якісна температурна залежність концентрації основних і неосновних носіїв в акцепторному напівпровіднику буде такою ж, як і у донорному. На ділянці високих температур розподіл носіїв на основні і неосновні не має сенсу. До того ж у широкозонних напівпровідниках, наприклад GaN, цей режим може лежати за межами температури плавлення.

3.5.1. Донорні напівпровідники

Скористатися виразами (3.14) і (3.24), або (3.26) і (3.27) для розрахунку концентрації носіїв за даних умов можливо тільки за відомої енергії Фермі. Її величина може бути визначена у разі підстановки у рівняння електронейтральності (3.36), яке за наявності тільки донорної домішки набуває вигляду

$$n_n = p_n + N_d^+, \quad (3.58)$$

відповідних виразів (для невиродженого напівпровідника (3.26), (3.27) і (3.44)) і розв'язку його відносно E_f . Мінімальні спрощення, які допускають аналітичний розв'язок трансцендентного рівняння, яке за цих умов виходить, можливі для двох крайніх, на перший погляд, режимів:

а) на ділянці низьких температур можна знехтувати неосновними носіями ($p_n \ll N_d^+$) і вважати, що електрони у зону провідності збуджуються тільки за рахунок іонізації домішок, тобто

$$n_n = N_d^+, \quad (3.59)$$

б) на ділянці високих температур можна припустити, що всі домішкові атоми іонізовані ($N_d^+ = N_d$), а додаткові носії утворюються парами у результаті міжзонних переходів:

$$n_n = p_n + N_d. \quad (3.60)$$

Насправді, як буде показано подальшими кількісними розрахунками, якщо енергія іонізації і ширина забороненої зони істотно відрізняються, то в області виснаження домішки виконуються обидві ці умови: температури вже досить високі для повної іонізації усіх донорних атомів і ще не такі високі, щоб враховувати міжзонні переходи.

Скористаємося цими припущеннями для невиродженого напівпровідника:

а) *Ділянка низьких температур.* Для спрощення вигляду отримуваних результатів зручно ввести величину

$$n_1 = N_c \exp [(E_d - E_c)/k_B T] = N_c \exp (-\Delta E_d/k_B T), \quad (3.61)$$

яку називають ефективною густиною станів, зведеною до домішкового рівня, приписуючи їй «фізичний» сенс, який витікає з порівняння (3.26) і (3.61): ця величина дорівнюватиме концентрації електронів у зоні провідності, якщо рівень Фермі співпадатиме з донорним рівнем. За цієї умови рівняння електронейтральності (3.59), записане у припущенні несуттєвості процесів генерації електронно-діркових пар, набуває вигляду

$$n_n = N_d / (n_n/n_1 + 1). \quad (3.62)$$

Враховуючи, що n_n пов'язана з E_Φ співвідношенням (3.26), з (3.62) можна знайти значення енергії Фермі

$$E_\Phi = E_d + k_B T \ln \left(\sqrt{N_d/n_1 + 1/4} - 1/2 \right) \quad (3.63)$$

і концентрації електронів

$$n_n = N_d / \left(\sqrt{N_d/n_1 + 1/4} + 1/2 \right). \quad (3.64)$$

У разі необхідності, концентрація неосновних носіїв, дірок у валентній зоні, може бути знайдена із закону діючих мас:

$$p_n = n_i^2/n_n, \quad (3.65)$$

де величина n_i заздалегідь визначається з (3.55).

б) *Ділянка високих температур.* У цій області усі домішкові атоми іонізовані ($N_d^+ = N_d$), а додаткові носії утворюються за рахунок міжзонних переходів і концентрація електронів завжди перевищує концентрацію дірок на величину N_d . Замінюючи у рівнянні електронейтральності (3.60) концентрацію дірок на n_i^2/n_n відповідно до закону діючих мас, отримаємо рівняння $n_n^2 - N_d n_n - n_i^2 = 0$, розв'язок якого для $n_n > 0$ матиме вигляд

$$n_n = \frac{N_d}{2} \left[1 + \sqrt{1 + \left(\frac{2n_i}{N_d} \right)^2} \right]. \quad (3.66)$$

З рівності (3.26) і (3.55) знаходиться положення рівня Фермі:

$$E_\Phi = E_c + k_B T \ln \left\{ \frac{N_d}{2N_c} \left[1 + \sqrt{1 + \left(\frac{2n_i}{N_d} \right)^2} \right] \right\} = E_c + k_B T \ln \left(\frac{n_n}{N_c} \right). \quad (3.67)$$

У разі використання останніх співвідношень для розрахунків необхідно застосовувати вираз (3.55) для попереднього знаходження n_i . Розрахунок концентрації неосновних носіїв може бути виконаний за відомих n_i , n_n і E_Φ різними способами: з рівняння

електронейтральності $p_n = n_n - N_d$; із закону діючих мас $p_n = n_i^2/n_n$; із загального виразу (3.27); з виразу:

$$p_n = N_d \left[\sqrt{1 + (2n_i/N_d)^2} - 1 \right] / 2, \quad (3.68)$$

яке може бути отримане так само, як і (3.66).

Формули, отримані для низьких і високих температур, придатні для розрахунків в усьому температурному діапазоні, оскільки в області виснаження домішки виконуються обидва прийняті допущення, і результати розрахунків перекриваються (рис. 3.11, а): функція (3.64) прагне до N_d з боку низьких температур, а (3.66) - з боку високих.

Іноді може виявитися зручною менш точна, але простіша апроксимація досліджуваних залежностей на трьох характерних ділянках:

а) на ділянці найнижчих температур, де $k_B T \ll \Delta E_d$, або $n_1 \ll 4N_d$, з (3.63) і (3.64) отримуємо співвідношення

$$n_n = (N_c N_d)^{1/2} \exp(-\Delta E_d/2k_B T), \quad (3.69)$$

$$E_\Phi = (E_c + E_d)/2 + k_B T \ln(N_d/N_c)^{1/2}, \quad (3.70)$$

подібні за формою до (3.55) і (3.53) для власного напівпровідника, і функція (3.69) в координатах $\ln n_n = f(1/T)$ приймає вигляд прямої лінії (ділянка 1 на рис. 3.11, б) з нахилом, пропорційним енергії іонізації домішки ΔE_d ;

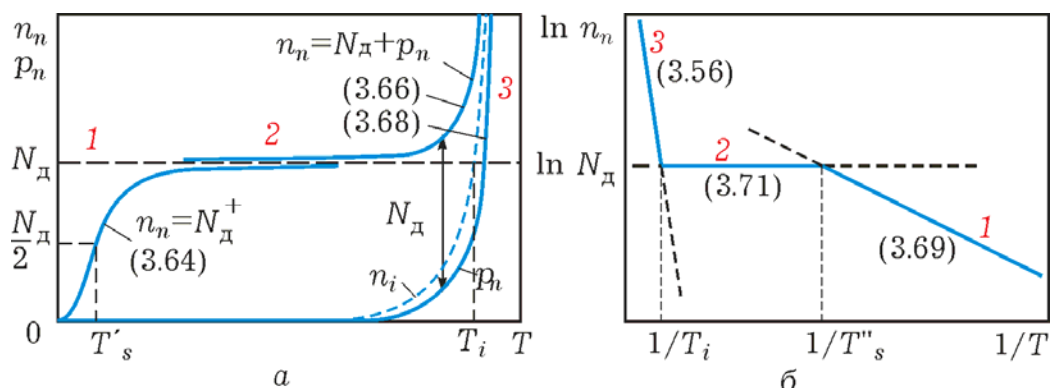


Рис. 3.11. Використання двох наближень (3.59) і (3.60) приводить до формул, які дають результати, що перекриваються у режимі насичення домішки (а). У спрямних координатах (б) зручно використовувати апроксимацію у вигляді трьох лінійних ділянок

б) в області виснаження домішки (ділянка 2) з (3.63) і (3.64) за умови $n_1 \ll N_d$, і з (3.66), (3.67) при $2n_i \ll N_d$ слідує, що

$$n_n = N_d, \quad (3.71)$$

$$E_{\text{ф}} = E_c + k_B T \ln (N_d/N_c), \quad (3.72)$$

а із закону діючих мас отримуємо

$$p_n = n_i^2/N_d; \quad (3.73)$$

в) на ділянці високих температур переважають міжзонні переходи ($n_i \gg N_d$), роль домішки стає несуттєвою, напівпровідник поводить ся як власний, а ділянка 3 на рис. 3.11, б, так звана область власної провідності, відповідає рис. 3.8, в.

Перехідну область між ділянками 2 і 3 за необхідності можна описати більш простими співвідношеннями, ніж (3.66) і (3.68), якщо в них обмежитися двома членами розкладання $\sqrt{1 + (2n_i/N_d)^2}$ у степеневий ряд:

$$n_n = N_d + n_i^2/N_d, \quad p_n = n_i^2/N_d, \quad (3.74)$$

На рис. 3.12 представлені у різних координатах результати розрахунку температурних залежностей концентрації основних і неосновних носіїв і енергії Фермі для германію, легованого донорною домішкою з енергією іонізації 0,01 еВ, за формулами, що відповідають двом припущенням (3.59) і (3.60). В розрахунках не бралася до уваги температурна зміна ширини забороненої зони, а також зміна енергетичних донорних станів зі збільшенням концентрації домішки.

Вид концентраційних залежностей відповідає загальній картині фізичних процесів, описаних на початку розділу. Потрібно також звернути увагу на одну особливість поведінки неосновних носіїв, яка наочно проявляється на початку зростання концентрацій за рахунок власної провідності. Концентрація дірок у валентній зоні завжди менша, ніж її значення у власному напівпровіднику. Це тим помітніше, чим вища концентрація донорів. Формально цей факт описується законом діючих мас, а фізично пояснюється тим, що із збільшенням концентрації домішки за інших рівних умов

збільшується концентрація електронів, тобто заповнених станів у зоні провідності, що зменшує ймовірність міжзонних переходів. Корисно звернути увагу на порядок значень концентрації неосновних носіїв за кімнатної температури: для концентрації донорів (тобто і електронів у зоні провідності) $N_d = 10^{15} \text{ см}^{-3}$ концентрація дірок складає $2 \cdot 10^{11} \text{ см}^{-3}$ у германії і $1 \cdot 10^4 \text{ см}^{-3}$ у кремнії, а для концентрації $N_d = 10^{16} \text{ см}^{-3}$ ці значення на порядок нижчі: $2 \cdot 10^{10} \text{ см}^{-3}$ і $1 \cdot 10^3 \text{ см}^{-3}$ відповідно.

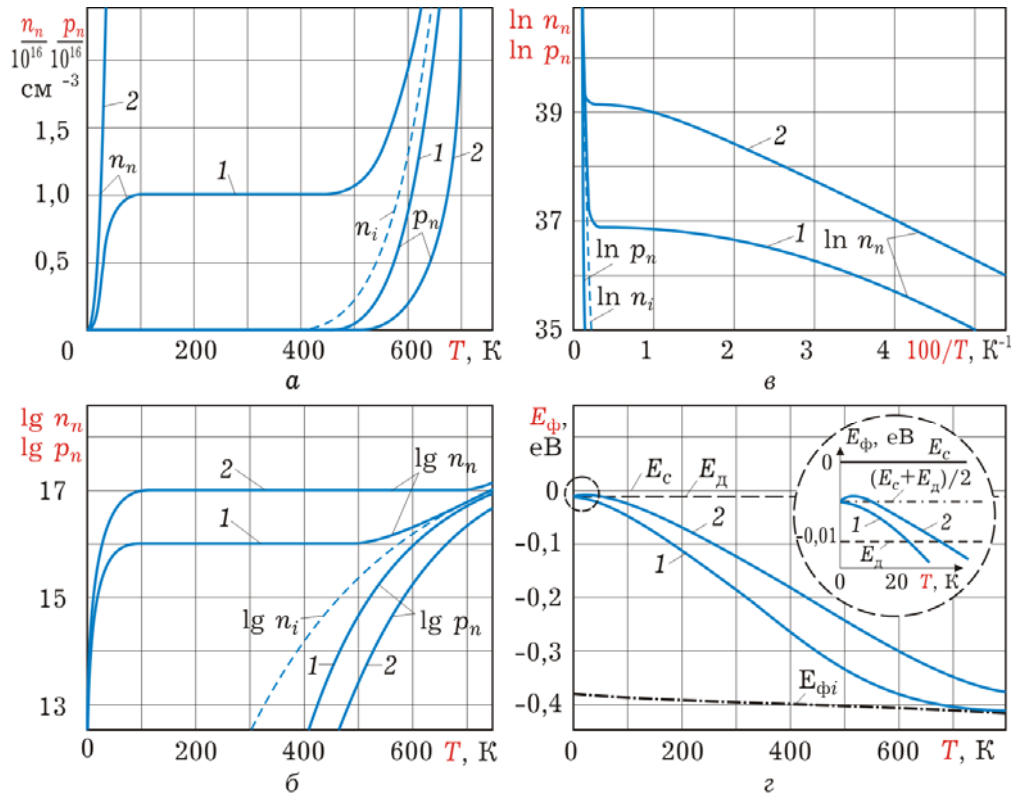


Рис. 3.12

Рис. 3.12. Залежності концентрацій електронів і дірок у різних координатах (а, б, в), а також рівня Фермі (з) від температури в германії ($\Delta E = 0,785 \text{ eV}$), легованої донорними домішками ($\Delta E_d = 0,01 \text{ eV}$) з концентрацією 10^{16} см^{-3} (1) і 10^{17} см^{-3} (2). Температурна зміна ширини забороненої зони не враховувалась. У колі зображено фрагмент залежності $E_\phi(T)$ для найнижчих температур у збільшеному масштабі

Для температурної залежності енергії Фермі (рис. 3.12, з) треба відмітити такі особливості. За нульової температури відповідно до (3.70) рівень Фермі знаходиться посередині між дном зони провідності і донорним рівнем. З підвищенням температури через збільшення $N_c \sim T^{3/2}$ рівень Фермі наближається до зони провідності,

потім $\ln (N_d/N_c)$ змінює знак, рівень Фермі опускається і, як випливає з (3.63) у разі $E_f = E_d$, за температури

$$T'_s = \Delta E_d / k_B \ln (2N_c / N_d) \quad (3.75)$$

перетинає донорний рівень. Це значення часто приймають як температуру виснаження домішки. Проте потрібно пам'ятати про умовність цієї межі, оскільки за температури T'_s тільки половина домішкових атомів іонізована, що видно з (3.64) для $n_1(T'_s) = N_d/2$.

З використанням трикускової апроксимації для такої температури обирають значення

$$T''_s = \Delta E_d / k_B \ln (N_c / N_d), \quad (3.76)$$

що відповідає точці «зламу» на рис. 3.11, б і витікає з (3.69) за умови $n_n = N_d$.

З подальшим підвищенням температури рівень Фермі віддаляється від дна зони провідності. Формально це підтверджується загальним виразом (3.28), в якому співвідношення p/n , залишаючись дуже малим, росте за рахунок збільшення кількості дірок зі збільшенням інтенсивності міжзонних переходів і практично постійного значення концентрації електронів у режимі виснаження домішки. За кімнатної температури рівень Фермі розташований нижче за дно зони провідності на 0,23 еВ у Ge і на 0,26 еВ у Si для значення $N_d = 10^{15} \text{ см}^{-3}$ і відповідно на 0,17 еВ і на 0,20 еВ для концентрації домішки 10^{16} см^{-3} .

На ділянці високих температур, де переважають міжзонні переходи, рівень Фермі практично співпадає з E_{fi} у власному напівпровіднику.

Верхньою межею області виснаження домішки вважають таку температуру T_i , за якої значення n_i стають співмірними з N_d , тобто $n_i = \alpha N_d$, де $\alpha \sim 1$. З урахуванням виразу (3.55) отримаємо

$$T_i = \Delta E / k_B \ln [N_c N_v / (\alpha N_d)^2]. \quad (3.77)$$

Ця температура для $\alpha = 1$ відповідає точці «зламу» між ділянками 2 і 3 на рис. 3.11, б, а значення концентрацій носіїв, як випливає з (3.66) і (3.68), дорівнюють $n_n \approx 1,6 N_d$ і $p_n \approx 0,6 N_d$. Ширина температурного інтервалу від T_s до T_i , в якому

концентрація електронів залишається постійною, має значення у разі вибору матеріалу для напівпровідникових приладів для зменшення залежності їх параметрів від температури. Це стосується у першу чергу таких приладів, де вирішальну роль грають основні носії, бо концентрація неосновних носіїв у цьому інтервалі хоча і є невеликою, але сильно залежить від температури. Як впливає з виразів для T_s і T_i , розширення цього температурного інтервалу можливе за рахунок: а) збільшення ширини забороненої зони, б) зменшення енергії іонізації домішки і в) збільшення її концентрації.

Наприклад, за концентрації домішки 10^{15} см^{-3} у германії цей діапазон має межі від $T_s = 16 \text{ К}$ до $T_i = 360 \text{ К}$ для $\alpha = 0,1$ або до $T_i = 430 \text{ К}$ для $\alpha = 1$, а в кремнії від $T_s = 34 \text{ К}$ до $T_i = 540 \text{ К}$ ($\alpha = 0,1$) або до $T_i = 640 \text{ К}$ ($\alpha = 1$).

Провідність домішкових напівпровідників, як і власних, у загальному випадку забезпечується носіями двох видів, тому

$$\sigma = en_n\mu_n + ep_n\mu_p. \quad (3.78)$$

Наявність домішки призводить до того, що на відміну від власних напівпровідників до розсіювання носіїв, викликаного тепловими коливаннями кристалічної ґратки, додається розсіювання на іонізованих і нейтральних (за найнижчих температур) атомах домішки, а температурна залежність рухливості електронів і дірок описується функціями типу (2.18).

Степеновий характер цих функцій набагато слабший, ніж експоненціальний для залежностей $n_n(T)$ і $p_n(T)$ в областях іонізації домішки і власної провідності. Тому залежність $\ln\sigma$ від $1/T$, яка представлена на рис. 3.13, на цих ділянках практично прямолінійна. І тільки в області виснаження домішки, коли концентрація основних носіїв постійна і дорівнює концентрації донорів, а неосновних знехтувально мала, $\sigma = eN_d\mu_n$, і рухливість є головним чинником, який визначає вплив температури на електропровідність. На цій ділянці температур домішкове розсіювання грає другорядну роль, а основним є механізм розсіювання електронів на теплових коливаннях кристалічної ґратки, за якого зі зростанням температури рухливість, а значить і електропровідність спадають через збільшення інтенсивності теплових коливань.

3.5.2. Акцепторні напівпровідники

У напівпровіднику, який містить тільки акцепторні домішки, як вже відзначалося, процеси генерації у значній мірі подібні до процесів в донорному напівпровіднику з тією різницею, що основними рухливими носіями тут будуть дірки у валентній зоні, які утворюються у першу чергу в результаті переходу електронів з валентної зони у вільні акцепторні стани (переходи 2 на рис. 3.9). Акцепторні атоми у цьому випадку стають від'ємно зарядженими іонами. За високих температур дірки утворюються також за рахунок міжзонних переходів. Одночасно з цим генеруються рухливі електрони у зоні провідності, які для акцепторного напівпровідника будуть неосновними носіями.

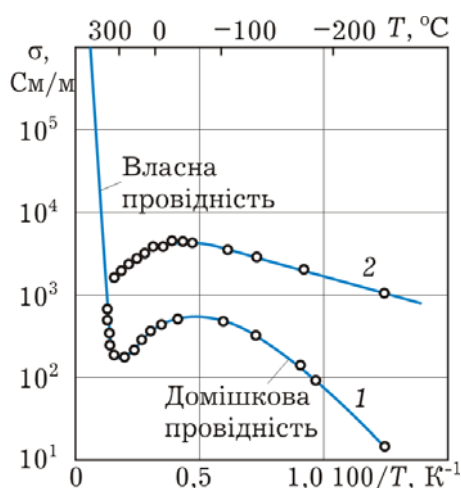


Рис. 3.13. Залежність питомої електропровідності кремнію від зворотної температури для значень концентрації донорної домішки: 1 - $5 \cdot 10^{23} \text{ м}^{-3}$, 2 - $3 \cdot 10^{24} \text{ м}^{-3}$

Концентрацію рухливих носіїв у зонах можна розрахувати із загальних виразів (3.14) і (3.24), а для невідроджених напівпровідників з (3.26) і (3.27). Положення рівня Фермі може бути визначене з рівняння електронейтральності, яке в даному випадку має вигляд

$$p_p = N_a^- + n_p. \quad (3.79)$$

Тут N_a^- - концентрація від'ємно іонізованих акцепторних атомів, яка визначається з (3.45), а індекс p біля концентрацій вказує на приналежність до акцепторного напівпровідника.

Повторювати розв'язок електронейтральності немає необхідності. Зручніше скористатися відміченою раніше симетрією між зоною провідності і валентною зоною, а також між донорними і акцепторними станами. Дійсно, зміна в усіх функціях розподілу

напрямку відліку енергії і заміна відповідних позначень і індексів (n на p і навпаки, E_c на E_v , N_d на N_a і тому подібне) не змінює їх. Тому вирази для рівня Фермі і концентрації рухливих носіїв у зонах для невідродженого акцепторного напівпровідника можуть бути отримані вказаною заміною з відповідних виразів для донорного напівпровідника.

Ділянка низьких температур :

$$E_{\Phi} = E_a - k_{\text{B}}T \ln\left(\sqrt{N_a/p_1 + 1/4} - 1/2\right), \quad (3.80, a)$$

$$p_p = N_a / \left(\sqrt{N_a/p_1 + 1/4} + 1/2\right), \quad (3.80, б)$$

$$p_1 = N_v \exp [(E_v - E_a)/k_{\text{B}}T] = N_v \exp (-\Delta E_a/k_{\text{B}}T). \quad (3.80, в)$$

Ділянка високих температур :

$$E_{\Phi} = E_v - k_{\text{B}}T \ln\left\{\frac{N_a}{2N_v} \left[1 + \sqrt{1 + \left(\frac{2n_i}{N_a}\right)^2}\right]\right\} = E_v - k_{\text{B}}T \ln\left(\frac{p_p}{N_v}\right), \quad (3.81, a)$$

$$p_p = N_a \left(1 + \sqrt{1 + (2n_i/N_a)^2}\right) / 2, \quad (3.81, б)$$

$$n_p = p_p - N_a = n_i^2 / p_p. \quad (3.81, в)$$

Ці вирази перекриваються, як і для донорного напівпровідника, в області виснаження домішки і достатні для розрахунку концентрацій в усьому температурному діапазоні (рис. 3.14). Більш прості вирази, придатні для трьох характерних областей, подібних до поданих на рис. 3.11, можуть бути отримані з (3.80) і (3.81) як окремі випадки.

За рівності концентрацій донорів і акцепторів відбувається повна компенсація домішок і поведінка такого напівпровідника з точки зору концентрації носіїв і положення рівня Фермі багато в чому подібна до поведінки власного напівпровідника. За різних концентрацій домішок, наприклад, для $N_d > N_a$, відбувається часткова компенсація, а температурні зміни концентрації рухливих носіїв і рівня Фермі у першому наближенні можна вважати такими ж, як і в донорному напівпровіднику, що містить домішки з ефективною концентрацією, рівною $N_d - N_a$.

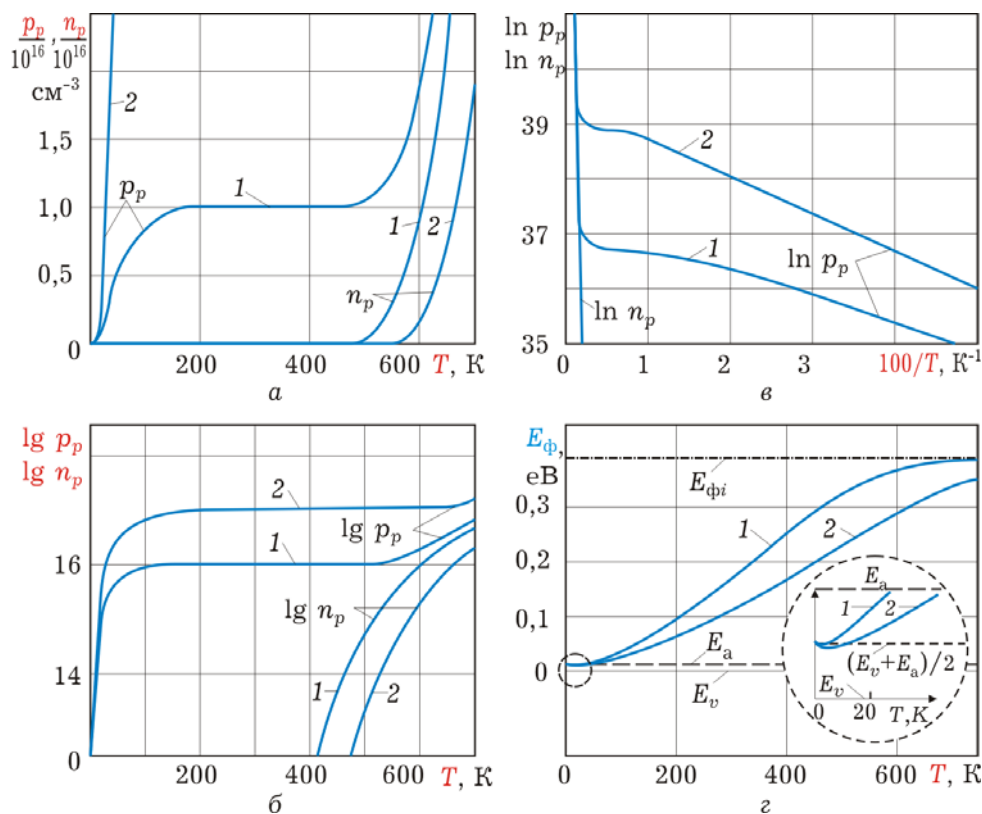


Рис. 3.14. Такі ж залежності, що і на рис. 3.12, для германію, легованого акцепторними домішками

На закінчення треба відзначити, що напівпровідники, які містять тільки один вид домішки, практично не зустрічаються. Разом з домішками одного виду, що спеціально вводяться, у напівпровіднику через недосконале очищення обов'язково є присутніми і домішки іншого виду. Іноді ті і інші домішки вводять спеціально для надання напівпровіднику потрібних властивостей. Наявність домішкових атомів різного виду призводить до того, що навіть за відсутності зовнішнього збудження відбувається іонізація домішок через те, що електрони переходять з донорних станів в енергетично нижчі акцепторні. У такому разі відбувається так звана *компенсація домішки* з меншою концентрацією.

3.6. РІВНОВАЖНІ НОСІЇ В ЕЛЕКТРИЧНОМУ ПОЛІ

3.6.1. «Викривлення» енергетичних зон

Якщо у кристалі яким-небудь зовнішнім джерелом створено електричне поле, що характеризується деяким потенціальним

рельєфом $V(\vec{r})$, то строго кажучи зонною теорією користуватися не можна, оскільки наявність зовнішнього поля порушує періодичність поля у кристалі. Проте ситуація істотно спрощується, якщо зовнішні поля набагато слабші за поле кристалічної ґратки, що має напруженість близько 10^{10} В/м. У такому разі зовнішнє поле хоч і порушує періодичність поля ґрат в цілому, але локально в межах деякої області кристала, що включає досить багато періодів, це збурення можна розглядати як слабке, а потенціальну енергію електрона в полі $U(\vec{r}) = -eV(\vec{r})$ - постійною. Тому природно вважати, що розв'язок рівняння Шредінґера у сумарному полі призводитиме практично до такого ж енергетичного спектру, що і в ідеальних ґратах, але зміщеного локально на величину $U(\vec{r})$.

Це означає, що в слабких зовнішніх полях можна зберегти уявлення про енергетичні зони, але усі рівні цих зон викривлятимуться відповідно до створеного в кристалі потенціального рельєфу $V(\vec{r})$ (рис. 3.15). Це викривлення може бути описане просторовими залежностями для країв зон:

$$E_c(\vec{r}) = E_{c0} + U(\vec{r}) = E_{c0} - eV(\vec{r}), \quad (3.82)$$

$$E_v(\vec{r}) = E_{v0} + U(\vec{r}) = E_{v0} - eV(\vec{r}),$$

де E_{c0} - дно зони провідності, а E_{v0} - стеля валентної зони за відсутності зовнішнього електричного поля, або, як то кажуть, для випадку *плоских зон*.

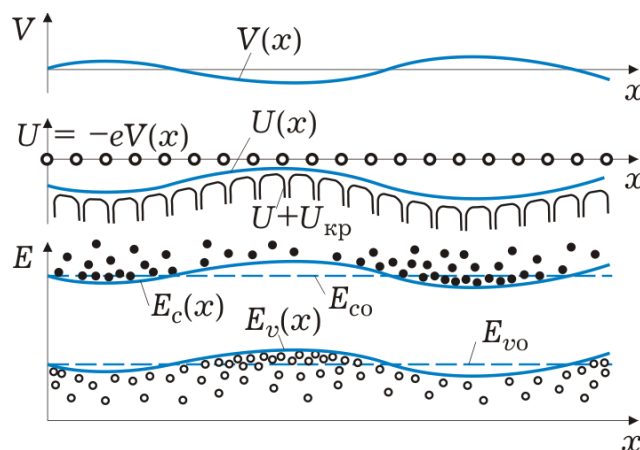


Рис. 3.15. Викривлення потенціального рельєфу і енергетичних зон і зміна концентрацій рухливих електронів і дірок за наявності зовнішнього електричного поля у напівпровіднику

Усе сказане вище відноситься також до внутрішніх макроскопічних полів, які можуть бути створені у результаті перерозподілу просторового заряду. Характерними є контактні поля, які виникають у місці контакту двох різних речовин. У напівпровідникових матеріалах вони мають досить велику протяжність, тому виконуватимуться умови, за яких можна використовувати модель викривлених зон. У металах вони локалізовані на відстані порядку періоду ґратки, їх напруженість співмірна з періодичним полем кристала, і тому ця модель непридатна.

3.6.2. Зміна концентрацій рухливих носіїв

Якщо в таких полях носії не створюють електричного струму, що можливо за відсутності електричного контакту з джерелом зовнішньої напруги або за наявності макроскопічного внутрішнього поля, то стан електронно-діркового колективу залишається рівноважним і значення енергії Фермі, як параметра функції розподілу, також залишається постійним для усього об'єму. Але локальна відстань між рівнем Фермі і межами зон змінюється (рис. 3.15) через викривлення енергетичних зон згідно з (3.82) законом. Відповідно відбувається і перерозподіл рухливих носіїв у зонах, їх концентрація буде просторово неоднорідною і може бути описана виразами (3.14) і (3.24), зміненими відповідно до (3.82) :

$$n(\vec{r}) = N_c \Phi_{1/2}[\xi_0 + eV(\vec{r})/k_B T], \quad (3.83)$$

$$p(\vec{r}) = N_v \Phi_{1/2}[\xi_{p0} - eV(\vec{r})/k_B T].$$

Тут ξ_0 і ξ_{p0} - приведені рівні Фермі ξ і ξ_p у точці, де $V(\vec{r}) = 0$.

Для невироджених напівпровідників у разі використання співвідношень (3.26) і (3.27), модифікованих відповідно до (3.82), неважко отримати

$$n(\vec{r}) = n_0 \exp[eV(\vec{r})/k_B T], \quad (3.84)$$

$$p(\vec{r}) = p_0 \exp[-eV(\vec{r})/k_B T],$$

де n_0 і p_0 - значення концентрацій електронів і дірок у точці, де потенціал приймається рівним 0, або, як іноді говорять, у режимі

«плоских зон». Отримані вирази показують, що рухливі електрони накопичуються у «западинах» потенціального рельєфу, тобто в областях з підвищеним потенціалом (рис. 3.15), а в «опуклостях» їх концентрація знижується. Поведінка рухливих дірок, природно, буде зворотною, причому, як витікає безпосередньо з (3.84), добуток концентрацій електронів і дірок в усіх точках залишається постійним:

$$n(\vec{r}) \cdot p(\vec{r}) = n_0 \cdot p_0. \quad (3.85)$$

Це відповідає закону діючих мас і є однією з формальних ознак рівноважного стану.

Співвідношення (3.84) для електронів було фактично отримане у розділі 2.3.2 з умови рівноваги дифузійних і дрейфових струмів у разі виконання співвідношення Ейнштейна. Дійсно, в кожній точці простору через наявність електричних полів і градієнта концентрацій носіїв виникають їх локальні дрейфові і дифузійні потоки, але, як неважко показати це на підставі виразів (3.84), вони компенсують один одного, що і призводить до рівноважного стану носіїв. Цікаво відзначити, що ця рівновага виконується окремо як для електронів так і для дірок.

Характерною і практично важливою ілюстрацією подібної ситуації є розподіл носіїв поблизу контакту електронного і діркового напівпровідників (рис. 3.16). Контактна різниця потенціалів V_k порядку десятих долей вольт, що виникає у приконтактній області завтовшки $10^{-1} \dots 10$ мкм, створює електричне поле з напруженістю близько $10^4 \dots 10^6$ В/м. Це електричне поле створюється у першу чергу іонізованими домішковими атомами: додатно зарядженими донорами в n - і від'ємно зарядженими акцепторами у p -напівпровіднику.

Розподіл рухливих носіїв у приконтактній області описується виразами (3.84), де n_0 і p_0 відповідають рівноважним концентраціям n_p і p_p в акцепторному напівпровіднику, якщо обрати тут початок відліку потенціалу. У n -області, де концентрації носіїв n_n і p_n , потенціал дорівнює V_k , тому

$$n_n = n_p \exp(eV_k/k_B T), \quad p_n = p_p \exp(eV_k/k_B T). \quad (3.86)$$

Звідси можуть бути отримані два рівноправні вирази для контактної різниці потенціалів

$$eV_{\kappa} = k_{\text{B}}T \ln(n_n/n_p) = k_{\text{B}}T \ln(p_p/p_n). \quad (3.87)$$

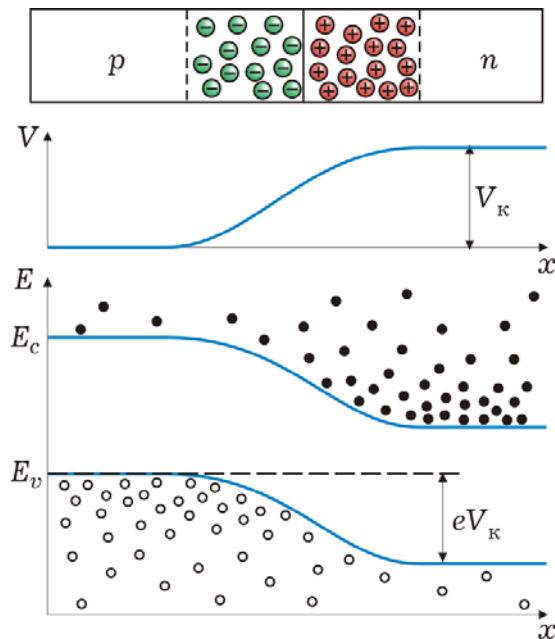


Рис. 3.16. Контактна різниця потенціалів, потенціальний бар'єр і розподіл рухливих носіїв поблизу контакту електронного та діркового напівпровідників

У приконтактній області, як це витікає з (3.84), рухливих носіїв обох знаків буде менше, ніж іонізованих домішок, тому ця область, яку називають *електронно-дірковим переходом*, вважається збідненою рухливими носіями і має нелінійні властивості, характерні для діодних структур.

3.7. ТЕРМОЕЛЕКТРОННА ЕМІСІЯ

Термоелектронною емісією називають вихід електронів з речовини (твердої або рідкої) під дією нагрівання. Суть цього явища полягає в тому, що з підвищенням температури збільшується енергія теплових коливань атомів речовини і, відповідно, теплова енергія електронів. Не зважаючи на те, що середнє значення теплової енергії невелике і складає десяті долі електронвольт за температур у тисячі кельвінів, все ж дуже мала частина електронів набуває енергії, яка у багато разів перевищує середню і може виявитися достатньою для подолання сил, що перешкоджають виходу електронів з речовини. Хоча таких електронів дуже мало, але саме вони обумовлюють процес емісії. За своєю фізичною суттю

термоелектронна емісія багато у чому подібна до процесу випарування молекул з рідини.

3.7.1. Закон термоелектронної емісії

Природа сил, що існують біля поверхні твердого тіла і перешкоджають емісії електронів, розглядалася у першій главі. Ці сили створюють біля межі твердого тіла з вакуумом потенціальний бар'єр, зображений на енергетичній діаграмі (рис. 1.34) і форма якого може бути розрахована методом дзеркального відображення. На рис. 3.24 з енергетичною діаграмою поєднана функція розподілу електронів по енергіях, з якої виходить, що за температури абсолютного нуля усі електрони заповнюють стани, що лежать нижче за рівень Фермі, тобто у напівпровідниках - у валентній зоні. У стани вище рівня вакууму E_0 , з яких можливий їх вихід, електрони переходять з підвищенням температури, і їх кількість, пропорційна заштрихованій області, зростає з температурою. Проте не усі електрони, що мають енергію $E > E_0$, можуть брати участь в емісії, оскільки швидкості їх спрямовані в різні боки. Необхідно, щоб енергія їх руху в напрямку до поверхні була достатньою для подолання сил дзеркального відображення ($v_x > v_{x\min}$). Значення паралельних до поверхні складових швидкості v_y і v_z не грають ролі і можуть бути будь-якими.

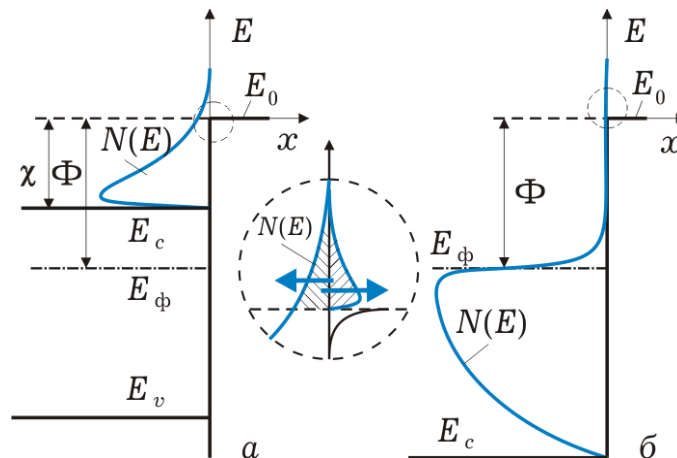


Рис. 3.17. Енергетичні діаграми і функції розподілу електронів у зоні провідності напівпровідника (а) і металу (б). На вставці наведено ці функції для внутрішніх електронів і тих, що вийшли, у станах, вищих за рівень вакууму. Площі заштрихованих областей рівні за умови термодинамічної рівноваги

Оскільки умови виходу формулюються для значень швидкостей v_x , то кількість електронів, що беруть участь в емісії, не може бути визначена через енергетичну функцію розподілу. Для цього потрібно скористатися розподілом за складовими швидкості. Такий «класичний» розгляд наведено у додатку ДЗ.

Для виведення основного закону емісії скористаємося наступним прийомом. Розглянемо електронний газ емітованих електронів над поверхнею катода. Якщо зовнішнє електричне поле не відводить електрони, то електронний газ у вакуумі знаходитиметься у термодинамічній рівновазі з газом в об'ємі напівпровідника і рівні Фермі для них співпадатимуть (рис. 3.17). За цих умов потоки електронів, емітованих катодом і тих, що повертаються назад, будуть рівні між собою, і густину струму емісії легко знайти, знаючи густину потоку часток, що повертаються. Число часток, що потрапляють на одиницю поверхні за одиницю часу, визначається відповідно до відомих результатів кінетичної теорії газів і дорівнює $nv_T/4$, де n – концентрація електронів у вакуумі, а $v_T = \sqrt{8kT/\pi m}$ їх середня теплова швидкість. Тому

$$j_s = env_T/4 = en\sqrt{kT/2\pi m}. \quad (3.88)$$

Використовувати результати класичної кінетичної теорії газів у даному випадку можна, оскільки рівень Фермі знаходиться нижче за рівень вакууму E_0 на величину $\Phi = E_0 - E_\phi$, яка зазвичай у багато разів перевищує $k_B T$, і електронний газ над поверхнею є невинородженим. З цієї ж причини його концентрацію легко розрахувати, як і концентрацію рухливих електронів у зоні провідності невинородженого напівпровідника, замінивши у виразі (3.26) рівень E_c на мінімальну енергію електронів у вакуумі E_0 , а в (3.13) – ефективну масу на масу вільного електрона:

$$n = 2(2\pi mkT/h^2)^{3/2} \exp[-(E_0 - E_\phi)/kT]. \quad (3.89)$$

Підставляючи $\Phi = E_0 - E_\phi$ і (3.84) в (3.83), приходимо до виразу

$$j_s = AT^2 \exp(-\Phi/kT), \quad (3.90)$$

де $A = 4\pi m e k^2 / h^3 = 1,2 \cdot 10^6 \text{ А/м}^2 \cdot \text{К}^2$. Вираз (3.90) називається формулою Річардсона-Дешмана і є основним законом термоелектронної емісії. Він характеризує залежність струму емісії від температури і термодинамічної роботи виходу (рис. 3.18). Графік $j_s(T)$ показує, що із зростанням температури густина емісійного струму зростає дуже різко, оскільки ця залежність визначається в основному експоненціальним множником. Її зручно представити в логарифмічному масштабі від зворотної температури. Дійсно, оскільки $\ln(j_s/T^2) = \ln A - \Phi/k_B T$, то функція (3.90) у координатах рис. 3.18, б має вигляд прямої лінії, кутовий коефіцієнт якої дорівнює $-\Phi/k_B$, а точка її перетину з віссю ординат - $\ln A$. Ця особливість використовується для експериментального визначення термодинамічної роботи виходу Φ і константи Річардсона A у рівнянні (3.90).

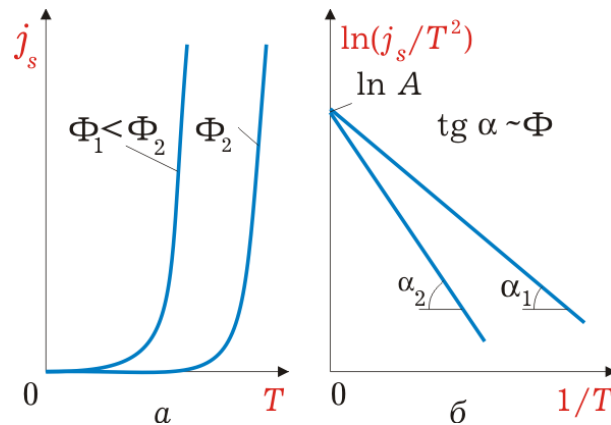


Рис. 3.18. Залежність густини струму термоелектронної емісії від температури у природних (а) і спрямлених (б) координатах. Нахил прямої тим більший, чим більша робота виходу

Наведене отримання закону термоелектронної емісії у принципі придатне як для металів, так і напівпровідників. Проте, незважаючи на формальний збіг цих законів, температурна залежність струму емісії для напівпровідників зазвичай набагато складніша, ніж для металів. І викликано це тим, що термодинамічна робота виходу напівпровідників може сильно залежати від температури. Для порівняно простих випадків, наприклад, для донорних і акцепторних напівпровідників, розглянутих в розділі 3.5, зміна рівня Фермі, а значить, і роботи виходу може бути розрахована. Проте для багатьох реальних напівпровідникових катодів, що мають складний склад домішок і структуру, такий розрахунок представляє

труднощі, і першочергового значення набувають експериментальні методи дослідження термоелектронних властивостей напівпровідникових катодів.

Цікаво звернути увагу на той факт, що у напівпровідниках, як і в металах, струм емісії визначається відстанню між рівнем Фермі і рівнем вакууму. Це може на перший погляд здатися дивним, враховуючи, що рівень Фермі напівпровідника зазвичай лежить у забороненій зоні і тому жоден з електронів у процесі виходу не виконує роботи, рівної Φ . Але якщо згадати, що енергія Фермі відповідає середній енергії електронів, що беруть участь у теплових переходах, і це середнє значення може знаходитися у забороненій зоні, то стане зрозумілим, чому в напівпровідниках термодинамічна робота виходу відлічується від рівня Фермі.

Розглянемо, від яких чинників залежить робота виходу напівпровідникових матеріалів. Істотний вплив на роботу виходу чинять вид і концентрація домішки. Якщо розглянути енергетичну діаграму домішкового напівпровідника (рис. 3.19), то стане зрозумілим, що за інших рівних умов меншу роботу виходу мають донорні напівпровідники, оскільки рівень Фермі у них розташований вище, ніж у акцепторних. З цієї ж причини збільшення концентрації донорної домішки, а також зменшення енергії іонізації домішкових атомів знижує роботу виходу напівпровідників.

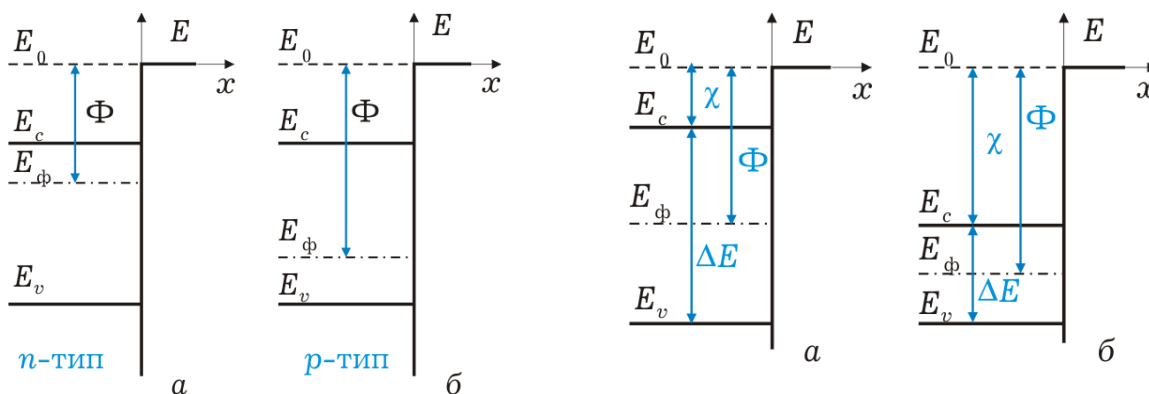


Рис. 3.19. Вплив виду домішки на роботу виходу. Донорні напівпровідники (а) є кращими емітерами, ніж акцепторні (б)

Рис. 3.20. Вплив електронної спорідненості на роботу виходу. За рівної відстані між рівнем вакууму і валентною зоною меншу роботу виходу має напівпровідник з більшою забороненою зоною

Важливим параметром напівпровідникових емітерів є енергія електронної спорідненості χ , тобто відстань між рівнем вакууму і

дном зони провідності. Напівпровідник з меншою енергією електронної спорідненості має і меншу роботу виходу, якщо, звичайно, рівень Фермі однаково розташований відносно дна зони провідності. Цікаво також відзначити, що з двох власних напівпровідників з однаковою відстанню між рівнем вакууму і стелею валентної зони кращим термоелектронним емітером буде той (рис. 3.20, а), який має меншу електронну спорідненість. Другий (рис. 3.20, б) матиме кращу електропровідність, оскільки за меншої ширини забороненої зони концентрація носіїв у ньому буде вищою.

На роботу виходу напівпровідників впливає присутність адсорбованих атомів. Як і у металах, плівка електропозитивних атомів знижує роботу виходу, створюючи біля поверхні електричне поле, яке сприяє виходу електронів, а наявність електронегативних

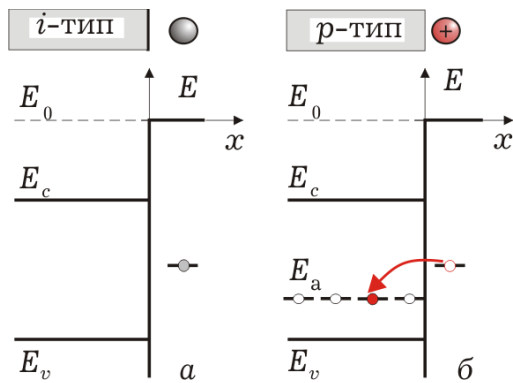


Рис. 3.21. Наявність акцепторної домішки підвищує ймовірність іонізації адсорбованих атомів

атомів «отрує» емітер, підвищуючи роботу виходу. Потрібно відзначити, що на властивості адсорбованих атомів впливає характер домішки у напівпровіднику. На рис. 3.21, а зображена ситуація, коли енергетичний стан валентного електрона атома потрапляє у заборонену зону і атом залишається нейтральним. За таких умов можлива тільки його поляризація за рахунок взаємодії з приповерхневими (наприклад,

таммівськими) станами напівпровідника.

Введення акцепторної домішки у напівпровідник (рис. 3.21, б) призводить до появи вільних станів, що лежать нижче за рівень валентного електрона, і відповідно до іонізації адсорбованого атома. У цьому разі збільшується електричне поле біля поверхні і зниження роботи виходу буде більш значним, ніж у разі бездомішкового напівпровідника.

Вплив адсорбованих атомів не вичерпується зміною форми і висоти потенціального бар'єру «зовні» напівпровідника. Не меншу роль грає також викривлення енергетичних рівнів «усередині» під дією електричного поля, що утворюється. Причина цього впливу буде розглянута нижче.

3.7.2. Вплив електричного поля на роботу виходу

Сили дзеркального відображення, які існують біля поверхні твердого тіла і перешкоджають виходу з нього електронів, можуть бути послаблені зовнішнім електричним полем, якщо поле є прискорюючим, тобто якщо його силові лінії спрямовані до поверхні катода. Це призводить до зниження потенціального бар'єра і відповідно до зменшення роботи виходу.

Для кількісного аналізу розглянемо плоский металевий катод (рис. 3.22 а), біля поверхні якого діє сила дзеркального відображення, рівна згідно (1.40), $-e^2/16\pi\epsilon_0 x^2$, і сила електричного поля $e\mathcal{E}$. Ці сили спрямовані назустріч одна одній і в точці

$$x_0 = \sqrt{e/16\pi\epsilon_0\mathcal{E}}. \quad (3.91)$$

взаємно компенсуються, що безпосередньо виходить з рівності цих сил. З рис. 3.22, б видно, що сили кристалічної ґратки

перешкоджають виходу електрона тільки до точки x_0 . На відстанях, більших за x_0 , поле буде прискорюючим. Потенціальна енергія електрона у такому полі визначатиметься сумою енергій в полі ґрат і в зовнішньому полі $-e\mathcal{E}x$ (рис. 3.22, в):

$$U(x) = -e^2/16\pi\epsilon_0 x - e\mathcal{E}x. \quad (3.92)$$

Точка x_0 відповідає максимуму потенціальної енергії і визначає найбільшу висоту потенціального бар'єру. Оскільки значення $U(x_0)$ відраховується від рівня вакууму за відсутності зовнішнього поля, то воно дорівнює зменшенню роботи виходу:

$$\Delta\Phi = U(x_0) = -e\sqrt{e\mathcal{E}/4\pi\epsilon_0}. \quad (3.93)$$

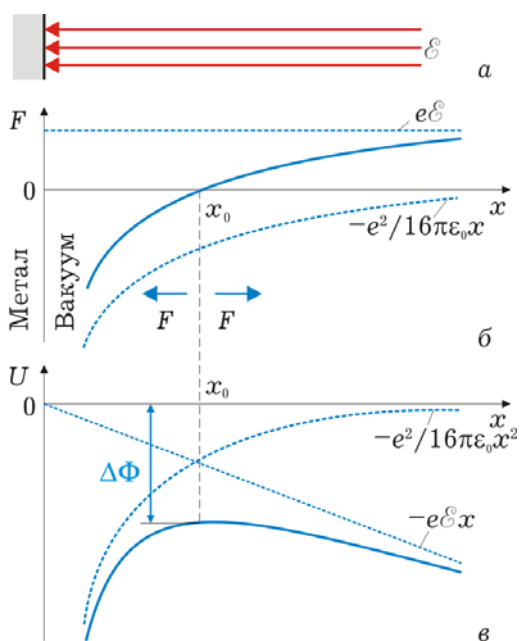


Рис. 3.22. Зміна сили (б) і потенціальної енергії (в) біля поверхні металу для прискорювального зовнішнього поля (а). У точці x_0 сили поля і дзеркального зображення врівноважуються, що відповідає максимуму потенціального бар'єру і його зниження на величину $\Delta\Phi$

Таке зниження бар'єру приводить до того, що його зможе подолати більша кількість електронів, тобто до збільшення струму термоелектронної емісії:

$$j_s = j_{s0} \exp[(e/kT) \sqrt{e\mathcal{E}/4\pi\epsilon_0}]. \quad (3.94)$$

Тут j_{s0} - струм емісії, визначуваний з (3.90) за відсутності зовнішнього поля. Таке явище називається *ефектом Шоттки*.

Підтвердженням правильності отриманих результатів служать представлені на рис. 3.23 експериментальні результати в координатах $\ln j_s$ і $\mathcal{E}^{1/2}$, що випрямляють залежність (3.94). Відхилення експериментальних точок від прямої у слабких полях може бути пояснено тим, що просторовий заряд електронів, що емітуються, екранує катод від зовнішнього поля, а також більш складними ефектами, пов'язаними з емісійною неоднорідністю поверхні катода.

Присутність біля катода гальмівного електричного поля (рис. 3.24) не призводить до зростання потенціального бар'єру безпосередньо поблизу поверхні, як це мало місце для ефекту Шоттки. Кількість електронів, які досягають анода, буде визначатися у цьому випадку не силою дзеркального відображення, яка локалізована біля поверхні катода, а силою зовнішнього електричного поля, або потенціальним бар'єром заввишки $e|V_a|$, який створюється у проміжку анод-катод від'ємним анодним потенціалом (рис. 3.24). Оскільки за енергії $E < e|V_a|$ емітовані електрони повертаються полем назад на катод, анодний струм створюватиметься тільки найбільш швидкими електронами, кількість яких дорівнює інтегралу від функції їх розподілу по енергіях в межах $e|V_a| < E < \infty$. Тому

$$j_a \sim \int_{e|V_a|}^{\infty} N(E) dE = f(V_a), \quad (3.95)$$

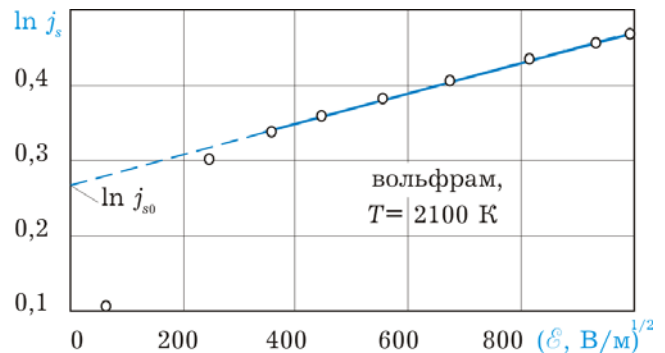


Рис. 3.23. Вплив зовнішнього прискорювального поля на струм термоелектронної емісії. Одиниці струму довільні

і вигляд вольт-амперної характеристики проміжку катод-анод, що дістала назву *кривої затримки*, за затримуючих анодних потенціалів визначається видом функції розподілу $N(E)$. Оскільки $dj_a/dV_a \sim N(E)$, то, диференціюючи експериментально виміряну криву затримки, неважко визначити вид функції розподілу емітованих електронів по енергіях. Цей метод, широко використовуваний при дослідженні емісійних явищ, називається *методом затримуючого потенціалу*.

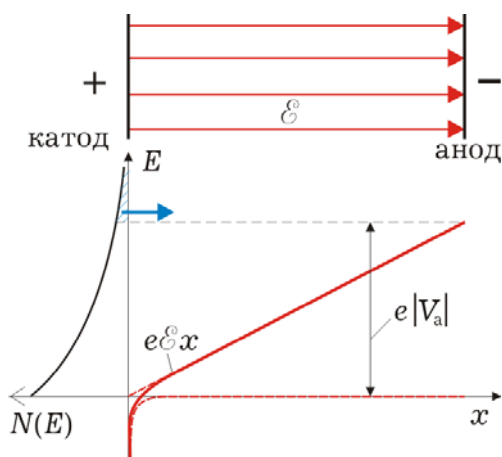


Рис. 3.24. Потенціальний бар'єр, створений гальмуючим полем анода, долають електрони з енергією, що перевищує висоту бар'єра

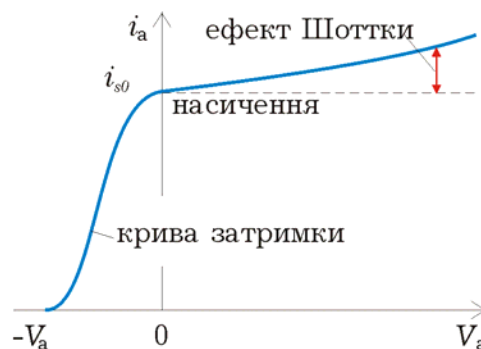


Рис. 3.25. Вольтамперна характеристика діода з металевим катодом. За затримуючої анодної напруги струм обумовлений початковими швидкостями вильоту електронів. За прискорюючої спостерігається збільшення струму через ефект Шоттки

Необхідно уточнити, що умова $E \geq e|V_a|$ придатна тільки для сферичної конструкції, коли радіус катода набагато менший за радіус анода. У цьому випадку, незалежно від напрямку вильоту з катода, електрони летять вздовж радіальних траєкторій. У системі плоских електродів електрон зможе здолати гальмівне поле, якщо буде достатньою складова його швидкості v_x у напрямі силових ліній поля, або $mv_x^2/2 \geq e|V_a|$. Очевидно, що у такому разі диференціювання кривої затримки дасть інформацію тільки про розподіл електронів по складовій швидкості v_x . Так само можна показати, що у циліндричній системі електродів вигляд кривої затримки залежить від розподілу електронів тільки по радіальних складових швидкості.

У результаті вольт-амперна характеристика вакуумного діода з металевим катодом (рис. 3.25) за від'ємних анодних потенціалів

повинна мати вигляд зростаючої кривої, оскільки зі збільшенням V_a усе більша кількість електронів долає гальмівне поле. За додатньої анодної напруги анодний струм зростає за рахунок ефекту Шоттки. Потрібно пам'ятати, що такий вигляд характеристика матиме, якщо рух окремих електронів не залежить один від одного. За великої густини струму на вигляд характеристики сильно впливає величина просторового заряду, що створюється самими електронами.

Ефект Шоттки проявляється і для напівпровідникових катодів. Проте картина фізичних явищ та їх кількісний аналіз набагато складніші, оскільки у цьому випадку, окрім зниження потенціального бар'єру, спостерігається проникнення електричного поля у приповерхневі шари напівпровідника. Через порівняно малу концентрацію рухливих носіїв у напівпровідниках заряди, які компенсують зовнішнє електричне поле (рис. 3.26, а), розташовуються не на самій поверхні, як у металах, а в приповерхневому шарі завтовшки в десятки і сотні періодів ґрат. Відповідно і напруженість зовнішнього поля дорівнює нулю не на поверхні напівпровідника, а поступово спадає усередині цього шару. Таке положення міняє не лише форму потенціального бар'єру, але і, як відзначалося у розділі 3.6, викликає викривлення енергетичних зон (рис. 3.26, в). Тому результуюче зменшення роботи виходу $\Delta\Phi$ визначається двома складовими: зниженням бар'єру по Шоттки $\Delta\Phi_{ш}$ і величиною вигину зон - $\Delta\Phi_B = eV_s$.

Подібне явище може бути викликане не лише зовнішнім прискорюючим полем, але і наявністю заряджених поверхневих станів. Як вже відзначалося, для металів присутність на поверхні електропозитивних адсорбованих атомів змінює форму потенціального бар'єру, знижуючи його. У напівпровідниках до цього потрібно додати викривлення енергетичних зон біля поверхні, викликане електричним полем іонізованого шару адсорбованих атомів. Зміну роботи виходу у такому разі враховують зазвичай через зміну енергії ефективної електронної спорідненості $\chi_{\text{еф}}$, визначуваної як відстань від дна зони провідності в об'ємі до рівня вакууму (рис. 3.27). Викривлення енергетичних зон може бути викликане не лише адсорбованими атомами, але і таммівськими поверхневими станами, які можуть заповнюватися електронами з об'єму або віддавати їх, утворюючи шар поверхневого заряду.

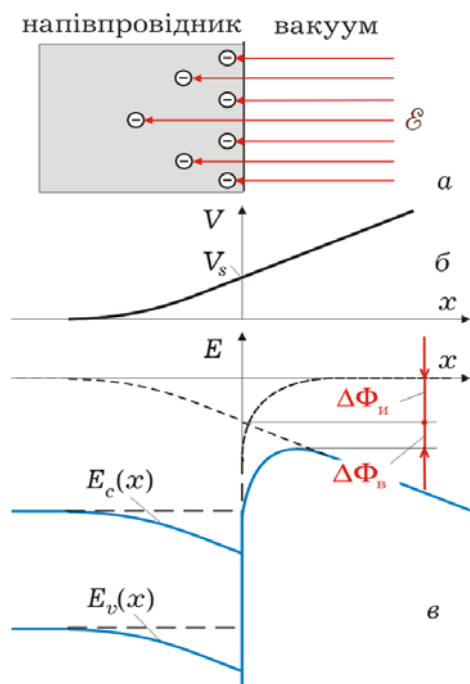


Рис. 3.26. Робота виходу з напівпровідника у зовнішньому полі зменшується за рахунок вигину зон біля поверхні і зниження потенціального бар'єру через ефект Шоттки

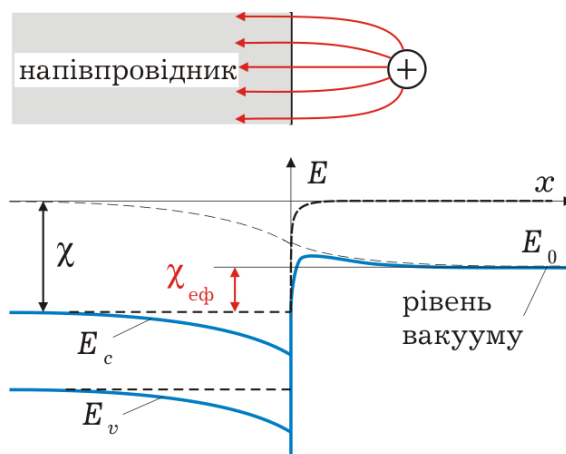


Рис. 3.27. Адсорбовані електропозитивні атоми деформують потенціальний бар'єр і викривляють енергетичні зони біля поверхні. У результаті ефективна енергія електронної спорідненості зменшується

Величина ефективної електронної спорідненості залежить не лише від характеру поверхневих станів, але і від виду домішки в напівпровіднику, а також від її концентрації. У напівпровідниках *n*-типу більш ймовірним є такий обмін, коли електрони з донорних рівнів переходять у вільні поверхневі стани, що лежать нижче, і утворюють подвійний електричний шар, поле якого викривляє енергетичні зони вгору (рис. 3.28, а), підвищуючи ефективну електронну спорідненість.

В акцепторних напівпровідниках (рис. 3.28, б), як правило, спостерігаються зворотні переходи електронів із заповнених поверхневих станів на вільні акцепторні рівні, що викликає вигин зон вниз і знижує ефективну електронну спорідненість. Важливо мати на увазі, що міра вигину зон і глибина проникнення поля залежать від густини просторового заряду. Збільшення концентрації домішки призводить до більшої зміни електронної спорідненості і одночасно зменшує товщину шару поверхневого просторового заряду. У цьому разі ефективна електронна спорідненість може стати навіть від'ємною (рис. 3.28, в). Варто відзначити також, що за

великої густини поверхневих станів шар просторового заряду дуже ефективно екранує об'єм напівпровідника від зовнішнього електричного поля. У таких випадках за прискорюючого поля ефект Шоттки проявляється так само, як і у металах.

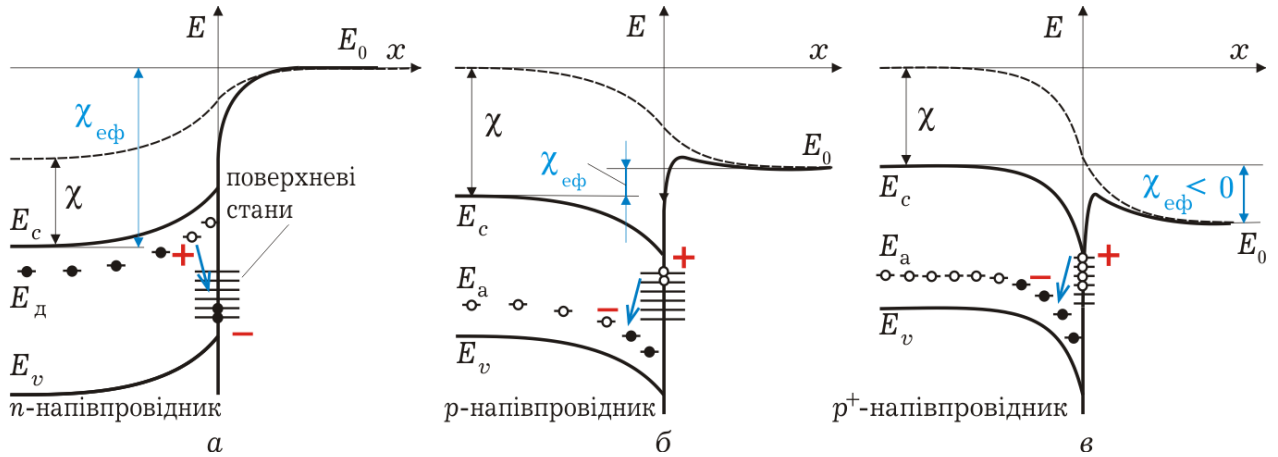


Рис. 3.28. Вплив поверхневих станів на енергію ефективної електронної спорідненості в залежності від виду та концентрації домішки

3.7.3. Термоелектронні емітери

Явище термоелектронної емісії використовується для створення катодів у різних електронних приладах. Для виготовлення катодів найбільшого поширення набув вольфрам, рідше використовуються молібден, тантал, реній, цирконій і деякі інші. Завдяки великій тугоплавкості робоча температура вольфрамового катода може досягати 2500 К, за якої, незважаючи на велику роботу виходу (4,5 еВ), густина струму досягає 0,3..0,5 А/см². Металеві катоди мають малу ефективність, оскільки вимагають великих витрат для нагрівання до високої температури. Тому область їх використання обмежується потужними лампами, що працюють за високої напруги, і приладами, що вимагають тривалої стабільності емісійних властивостей, яка не може бути забезпечена іншими типами катодів.

Значно більшу ефективність мають так звані *плівкові катоди*. Вони є металевою поверхнею, покритою шаром адсорбованих атомів. Фізичні процеси, що відбуваються у такій структурі, були розглянуті у гл. 1 і зводяться до наступного. У результаті обміну електронами між плівкою адсорбованих атомів і металом виникає подвійний електричний шар, поле якого деформує потенціальний бар'єр. Якщо адсорбовані атоми електропозитивні по відношенню до

металу, то сили електричного поля сприяють виходу електронів з металу, знижуючи потенціальний бар'єр і зменшуючи роботу виходу. Хороші емісійні властивості мають плівки торію, барію, цезію і деяких інших речовин на поверхні вольфраму (таблиця. 3.1).

Таблиця 3.1

Матеріал	Робота виходу, еВ	Стала $A \cdot 10^{-6}$, $A/m^2 \cdot K^2$	Робоча температура, К	Густина струму емісії, A/cm^2
Вольфрам	4,5	0,5...0,6	2500	0,4
Молібден	4,3	0,5...0,6	2300	0,2
Тантал	4,1	0,4...0,6	2100	0,15
Торій на вольфрамі	2,6	0,03	1800	0,4
Барій на вольфрамі	1,6	0,15	1000	0,02
Оксидний катод	1,0	$10^{-5} \dots 10^{-4}$	1100	0,1...10

З даних таблиці видно, що у плівкових катодів істотно зменшується і стала A у рівнянні (3.90). Це пов'язано, по-перше, з сильною «плямистістю» плівкових катодів, оскільки адсорбовані атоми утворюють на поверхні катода плівки у вигляді окремих груп атомів або плям. По-друге, наявність атомів сильно змінює форму, а значить, і прозорість потенціального бар'єру і сприяє тим самим зменшенню струму емісії.

Серйозним недоліком плівкових катодів, що обмежує їх використання, є невисока стабільність емісії. У разі перегрівання або бомбардування іонами залишкових газів поверхневий шар легко руйнується і в результаті робота виходу збільшується.

Найбільш поширеним видом термоелектронних емітерів є оксидний катод. Його виготовляють зазвичай з суміші карбонатів барію, кальцію і стронцію, які у разі нагрівання в процесі виготовлення перетворюються на відповідні оксиди. Готовий оксидний катод є (рис. 3.29, а) пористим шаром дрібнодисперсійної суміші оксидів, нанесеним на металеву підкладку (кern). Між підкладкою і оксидом у результаті хімічної взаємодії оксидів і речовин, що додаються в kern для надання катода певних властивостей, утворюється тонкий шар, який має дуже високий опір і тому називається запірним шаром. Основні процеси в оксидному катоді визначаються окислом барію, енергетична діаграма якого представлена на рис. 3.29, б.

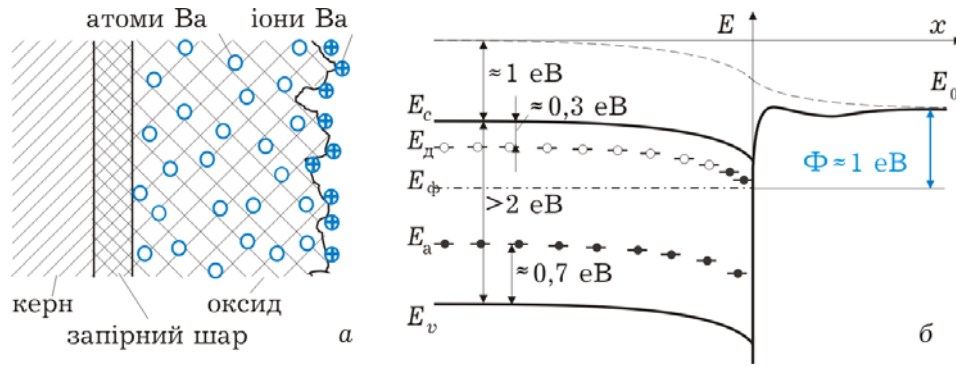


Рис. 3.29. Структура (а) і енергетична діаграма (б) оксидного катода

Окисел барію є напівпровідником з шириною забороненої зони більшою за 2 еВ і енергією електронної спорідненості близькою до 1 еВ. Наявність акцепторних станів, природа яких зазвичай зв'язується з неминучими у дрібнозернистій структурі дефектами ґрат, приводить до того, що робота виходу такого катода не може бути малою і складає величину, близькою до 3 еВ. Для поліпшення емісійних властивостей ще на етапі виготовлення приладу такий катод активують. Активування полягає в нагріванні катода до температури приблизно у 1300 К. Під час розкладання оксиду утворюються вільні атоми барію, які по відношенню до оксиду барію є донорними. Поява донорних атомів на початковому етапі активування, по-перше, нейтралізує акцепторні стани і, по-друге, зміщує рівень Фермі ближче до зони провідності. Крім того, атоми барію на поверхні катода утворюють плівку додатних іонів, що також знижує роботу виходу. У готового активованого катода робота виходу дорівнює приблизно 1 еВ, що за робочої температури оксидного катода близько 1100 К забезпечує густину струму емісії, більшою за 1 А/см².

Розглянемо деякі елементарні процеси, що відбуваються в оксидному катоді. У процесі термоелектронної емісії катод покидають електрони, енергія яких перевищує середнє значення. За цих умов рівноважний максвеллівський розподіл термоелектронів у зоні провідності оксиду порушується, і електронний газ, а разом з ним і катод, охолоджується. Одночасно відбувається і зворотний процес розігрівання анодним струмом, що проходить. Це розігрівання неоднорідне і найбільше там, де є ділянки з великим омичним опором (запірний шар і місця контактів зерен оксиду) або ділянки підвищеної густини струму (виступи на шорсткій поверхні катода).

У процесі роботи оксидного катода кількість атомів барію зменшується через їх випаровування, що призводить передусім до утворення збідненого поверхневого шару, що має підвищений у порівнянні з об'ємом питомий опір. Відновлення барію відбувається за рахунок термічного розкладання окислу барію, а також у результаті електролітичних процесів під дією електричного поля, що виникає у катоді під час протікання струму. Іони барію рухаються до підкладки, там нейтралізуються і дифундують по всьому об'єму. Іони і атоми кисню, що утворюються під час розкладання оксиду, а також залишкові гази (особливо галогени, пари води) можуть утворювати на поверхні катода шар від'ємно заряджених іонів, що підвищують роботу виходу і отруюють катод. Отруєння катода призводить до зменшення струму емісії у процесі роботи. Цьому ж сприяє і втягування іонів барію електричним полем всередину катода.

За великих емісійних струмів і зниженої температури катода на поверхні можна спостерігати іскріння, що супроводжується відривом частинок оксиду. Це явище може бути пов'язане з локальним перегріванням окремих ділянок катода, що виникає через нерівномірність емісії з поверхні. Дійсно, співвідношення $dj_s/j_s = -d\Phi/k_B T$, що легко витікає з (3.90), означає, що за робочої температури локальне зменшення роботи виходу навіть на 0,1 еВ викликає зміну струму з цієї ділянки на величину $\Delta j_s \approx j_{scp}$, тобто. удвічі. Зменшення температури посилює цю нерівномірність, тому робота катода за умови недорозжарювання може призвести до руйнування оксидного шару.

Цікаві особливості оксидного катода проявляються в імпульсному режимі. Якщо на анод приладу подавати високу напругу у вигляді коротких (близько мікросекунд) і порівняно не частих імпульсів, то струм емісії може зростати в сотні і навіть тисячі разів у порівнянні з неперервним режимом. Збільшення густини струму викликається сильним електричним полем, що виникає у шарі оксиду під час протікання великого струму і призводить до такої зміни максвеллівського розподілу електронів, яка відповідає «розігріванню» електронного газу. Важливу роль грає, ймовірно, і звільнення просторового заряду, що накопичується у порах оксиду.

Наведений розгляд деяких процесів, що відбуваються у оксидному катоді, не описує усіх його властивостей, а лише підкреслює складність явищ і частково пояснює,

чому до теперішнього часу не існує точної теорії катода, незважаючи на численність досліджень, що проводилися.

Контрольні запитання та вправи

3.1. Опишіть суть основного методу статистичного аналізу великих колективів частинок і роль функцій розподілу. Які функції розподілу типові для електронних колективів?

3.2. Опишіть метод і основний результат розрахунку концентрації електронів в зоні провідності. Який вид він має для невідроджених напівпровідників?

3.3. Дайте відповідь на попереднє питання для дірок у валентній зоні, звернувши увагу на симетрію понять.

3.4. Що таке закон діючих мас для електронно-діркового колективу і за яких умов він застосовний?

3.5. Який метод розрахунку і його результати для концентрації іонізованих домішкових (донорних і акцепторних) станів?

3.6. Поясніть залежність провідності металів від температури, використовуючи властивість виродженості електронного газу і особливості його розсіювання.

3.7. Поясніть хід температурної залежності питомої електропровідності, а також рівня Фермі, для власних напівпровідників, пов'язавши її з процесами генерації носіїв і механізмами розсіювання. Зверніть увагу на форми представлення цієї залежності у різних координатах і на її зв'язок з параметрами зонної структури.

3.8. Виконати попереднє завдання для домішкових напівпровідників (донорних і акцепторних).

3.9. Наведіть і поясніть зв'язок концентрації рівноважних носіїв і потенціального рельєфу, що створюється електричним полем (тут зручно скористатися моделлю викривлених зон).

3.10. Опишіть механізм термоелектронної емісії і приведіть її основний закон. Які параметри і яким чином можуть бути визначені із експериментальної залежності струму емісії від температури?

3.11. Які фактори впливають на роботу виходу термоелектронних емітерів? Як впливає прискорююче електричне поле на струм термоемісії (ефект Шоттки)? Яку інформацію можна отримати, вимірюючи анодний струм за гальмуючого електричного поля (метод затримуючого потенціалу)?

4. ГАРЯЧІ НОСІЇ

У СИЛЬНОМУ ЕЛЕКТРИЧНОМУ ПОЛІ

Для більшості мікроелектронних приладів з характерними розмірами порядку мікрометрів для напруг у кілька вольт значення напруженості електричного поля має порядок 10^6 В/м. Такі поля зазвичай вважають сильними. Загальне визначення *сильних електричних полів* дати важко через велику різноманітність ефектів, що в них відбуваються. Найчастіше сильними вважають такі поля, в яких спостерігаються відхилення від омічної залежності між струмом і напругою, або між дрейфовою швидкістю та напруженістю поля. Різні прояви цих явищ можна поділити на дві групи: ті, що пов'язані зі зміною кінетичних властивостей (рухливості) носіїв у сильних полях, і ті, що спричинені зміною їх концентрації.

4.1. ДРЕЙФ НОСІЇВ

Аналіз процесів дрейфового руху часто передбачає незалежність від електричного поля часу вільного пробігу в «одночастковій» моделі або часу релаксації імпульсу в колективній моделі (розд. 2), що можливо тільки для слабких електричних полів. За цих умов швидкість дрейфового руху залишається пропорційною напруженості електричного поля, що відповідає омічній залежності між густиною струму і напруженістю. У сильних електричних полях наближення часу релаксації стає, строго кажучи, неточним, бо швидкість зміни функції розподілу перестає бути пропорційною її відхиленню від рівноважної. Але в багатьох випадках цю нелінійність враховують, зберігаючи таке наближення лише формально і вважаючи час релаксації (а відповідно й рухливість) таким, що залежить від середньої енергії електрона, а отже, і від напруженості електричного поля. Це викликає нелінійність вольт-амперної характеристики однорідних зразків матеріалу.

4.1.1. Розігрівання електронного газу

Незважаючи на різні прояви ефектів сильного електричного поля, вони мають у своїй основі, як правило, одну першопричину – значне розігрівання електронного газу. Перевищення температури електронного колективу над температурою кристалічної ґратки спостерігається і в слабких полях, але воно не викликає істотної зміни кінетичних параметрів руху електронів.

Електронна температура визначає основні релаксаційні параметри електронного колективу напівпровідників, а значить і їх кінетичні та динамічні параметри, такі як рухливість, реакцію на імпульсні і високочастотні поля і тому подібне. Залежність рухливості від напруженості електричного поля можна визначити, лише розглядаючи *розігрівання електронного колективу* (газу), тобто аналізуючи ступінь впливу електричного поля на середню енергію хаотичного руху, мірою якого є електронна температура, яку визначають із співвідношення

$$E_T = 3k_B T_e / 2. \quad (4.1)$$

Основна фізична причина, яка пояснює підвищення електронної температури у сильних полях, – *непружність співударянь*, що формально проявляється через відмінність часів релаксації імпульсу та енергії. Електрон, прискорюючись електричним полем під час вільного пробігу, набуває імпульсу $p = m^* v$ і, відповідно, енергії $E_d = m^* v^2 / 2$ спрямованого дрейфового руху. Але ця спрямованість практично повністю втрачається вже після першого співударяння, яке ефективно хаотизує рух, бо час релаксації імпульсу має той самий порядок, що й середній час вільного пробігу. Проте повного розсіювання одержаної за цей час енергії не відбувається, оскільки співударяння всіх видів мають певний ступінь neprужності, і для повного розсіювання необхідним був би час порядку τ_E , який зазвичай більший, ніж τ_p . Залишок енергії поповнює запас не спрямованого, а хаотичного теплового руху, що означає збільшення електронної температури, яку визначають із рівності (4.1).

Розглянемо спочатку механізм розігрівання в *однодолінній моделі*, придатний для кремнію або германію, за умови

збереження в ній кількості електронів. Для такого аналізу необхідно використовувати як рівняння збереження імпульсу (2.46) так і рівняння збереження енергії (2.47) за умови незмінної кількості часток у зоні ($dn/dt = 0$) у такій формі:

$$d\vec{p}/dt - e\vec{\mathcal{E}} = -\vec{p}/\tau_p, \quad (4.2)$$

$$dE/dt - e\vec{v}\vec{\mathcal{E}} = -(E - E_0)/\tau_E. \quad (4.3)$$

Для кількісного аналізу розігрівання електронів необхідно з рівнянь балансу повної енергії та енергії дрейфового руху отримати рівняння балансу енергії тільки для хаотичного руху. У рівнянні балансу *повної енергії* (4.3) доданок $e\vec{v}\vec{\mathcal{E}}$ характеризує енергію, набуту від поля за одиницю часу, а доданок $(E - E_0)/\tau_E$ – втрачену в результаті зіткнень за той самий час. Щоб отримати рівняння балансу для *енергії дрейфового руху* $E_d = m^* v^2/2$ помножимо всі доданки рівняння (4.2) скалярно на \vec{v} і отримаємо

$$dE_d/dt - e\vec{v}\vec{\mathcal{E}} = -2E_d/\tau_p. \quad (4.4)$$

Тут права частина – це енергія дрейфу, яка втрачається за 1 с, а $\tau_p/2$ – час її релаксації. Віднімаючи останнє рівняння від (4.3), отримуємо рівняння балансу тільки *хаотичної (теплової) енергії*, яке й визначає динаміку зміни електронної температури:

$$\frac{dE_T}{dt} = -\frac{E_T - E_0}{\tau_E} - \frac{E_d}{\tau_E} + \frac{2E_d}{\tau_p} = -\frac{E_T - E_0}{\tau_E} + \frac{E_d}{\tau_E} \left(\frac{2\tau_E}{\tau_p} - 1 \right). \quad (4.5)$$

Доданок $2E_d/\tau_p$ дорівнює енергії дрейфу, яка за 1 с перетворюється безпосередньо в теплову, але її частина, рівна E_d/τ_E , за той самий час релаксує і переходить до кристалічної ґратки. У підсумку збільшення теплової енергії за рахунок дрейфової за 1 с становить $E_d(2\tau_E/\tau_p - 1)/\tau_E$. Відношення $2\tau_E/\tau_p \equiv \gamma$ є параметром, який характеризує ефективність переходу енергії дрейфу в теплову і визначає ступінь непружності співударянь щодо напрямленої енергії. Тому розігрівання електронного газу можливе

тільки за умови, що швидкість розсіювання імпульсу перевищує швидкість розсіювання енергії, тобто коли $\tau_E > \tau_p/2$. І тільки за рахунок перевищення електронної температури над рівноважною ($E_0 = 3k_B T/2$), що дорівнює температурі кристалічної ґратки, можливий *стаціонарний стан*, за якого віддана та отримана електронами енергії врівноважують одна одну. Це відбувається, коли електронна температура встановлюється на рівні, який відповідає рівності

$$E_T = E_0 + E_d(\gamma - 1). \quad (4.6)$$

Беручи до уваги, що в стаціонарному стані з огляду на (4.2) швидкість дрейфу $v = e\tau_p \mathcal{E}/m^* = \mu \mathcal{E}$, де $\mu = e\tau_p/m^*$ – рухливість, отримаємо для електронної температури вираз

$$T_e = T + \mu^2 \mathcal{E}^2 m^* (\gamma - 1) / 3k_B. \quad (4.7)$$

У слабкому полі, коли часи релаксації не залежать від електронної температури, отримана залежність $T_e(\mathcal{E})$ є параболічною. Але у процесі розігрівання питома вага різних механізмів суттєво змінюється: у сильних електричних полях переважають механізми за участю оптичних фононів. Тому в загальному випадку отримання залежності $T_e(\mathcal{E})$ в аналітичному вигляді практично неможливе і її зручніше знаходити як обернену до $\mathcal{E}(T_e)$:

$$\mathcal{E} = \sqrt{\frac{3k_B(T_e - T)}{m^* \mu^2 (\gamma - 1)}} = \frac{v_T}{\mu} \sqrt{\frac{T_e / T - 1}{\gamma - 1}}, \quad (4.8)$$

де $v_T = \sqrt{3k_B T / m^*}$ – рівноважна теплова швидкість «холодних» електронів, яка відповідає температурі ґратки.

Кремній. На рис. 4.1 наведено залежності електронної температури, часів релаксації імпульсу та енергії від напруженості електричного поля, розраховані для температури $T = 300$ К для механізмів розсіювання, розглянутих у підрозд. 2.3,

за тих самих числових значень параметрів, які було використано для розрахунку слабопольової рухливості (підрозд. 2.4). Тут потрібно звернути увагу на те, що оптичне розсіювання є переважаючим, оскільки акустичне розсіювання через велику пружність практично не впливає на час релаксації енергії, а домішкове помітне тільки за дуже слабкого розігрівання ($T \approx T_e$).

Треба звернути увагу також на те, що електронна температура і час релаксації енергії практично не залежать від концентрації домішки, а час релаксації імпульсу, тобто й рухливість, цю залежність виявляють у відносно слабких полях (до 10...20 кВ/см). Причому τ_p спадає зі збільшенням напруженості електричного поля (у сильному полі $1/\tau_p \sim \mathcal{E}$), а τ_E – зростає. У результаті посилюється нерівність $\tau_E > \tau_p$, що означає збільшення ступеня непружності розсіювання.

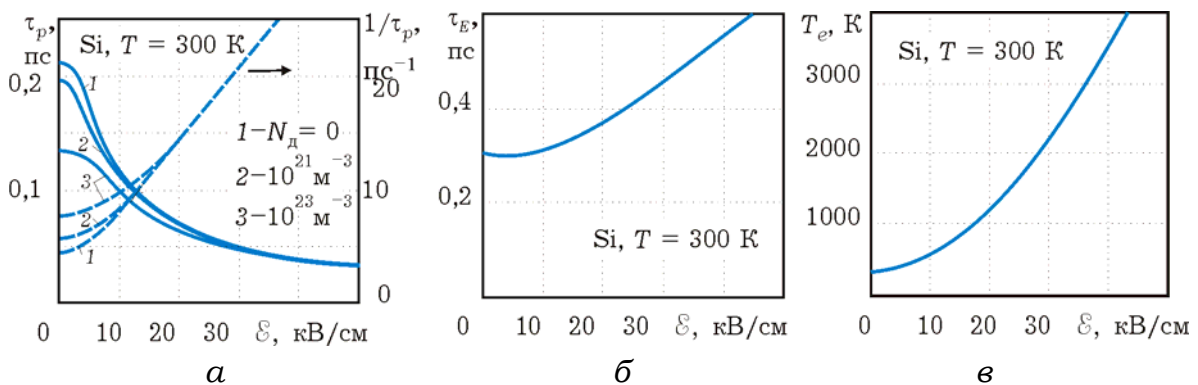


Рис. 4.1. Залежності часів релаксації імпульсу (а), енергії (б) та електронної температури (в) від напруженості електричного поля для кремнію. Штрихові лінії (а) характеризують вплив концентрації домішки на обернений час релаксації імпульсу

Германій. Подібні результати розрахунків для германію подано на рис. 4.2. Основні відмінності від кремнію пов'язані перш за все з переважанням у германії акустичного виду розсіювання, що впливає на значення часу релаксації імпульсу (рис. 2.20, а) і особливо помітно збільшує час релаксації енергії (рис. 4.2, б, в).

Передусім це пояснюється більшою пружністю акустичного розсіювання, що потребує збільшення кількості співударянь для розсіювання енергії, набутої в електричному полі. Крім того, більша пружність акустичних співударянь, а також вища

рухливість викликають більший ступінь розігрівання: у таких саме полях електронна температура в германії на порядок вища.

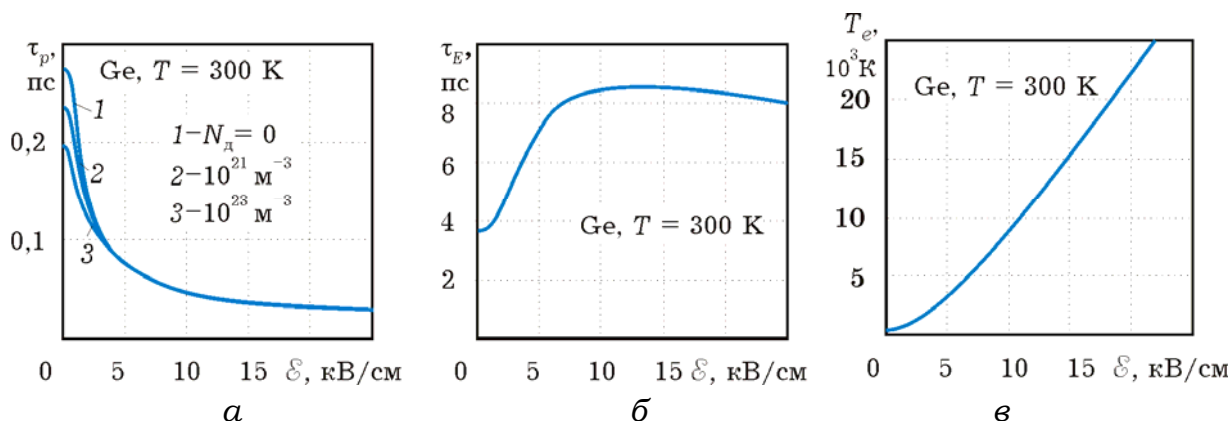


Рис. 4.2. Такі самі залежності, що й на рис. 4.1, для германію. Слід звернути увагу на значно більшу вагу акустичного розсіювання

У багатодолинних напівпровідниках крім Γ -долини біля дна зони провідності, є й більш високі L- та X-долини. У кожній з них механізм розігрівання, у принципі, такий самий, але ситуація ускладнюється тим, що, по-перше, у різних долинах електрони мають різні параметри, тому і ступені їх розігрівання будуть різнитися. По-друге, зі зростанням електронної температури змінюється і концентрація електронів у долинах. Вважаючи, що повна концентрація електронів $n(T)$ у зоні провідності залишається постійною і визначається температурою ґратки, а їх перерозподіл по долинах у результаті міждолинного розсіювання залежить від електронної температури, відносну заселеність долин можна охарактеризувати коефіцієнтами ($i, j = \Gamma, L, X$):

$$b_i(T_e) = n_i/n = n_i/(n_i + n_j) = 1/(1 + n_j/n_i), \quad (4.9)$$

які визначають відповідно до (3.20)

$$n_i/n_j = (m_{di}^*/m_{dj}^*)^{3/2} \exp(-\Delta_{ij}/k_B T_e).$$

Вираз (4.9) можна узагальнити і на випадок трьох долин.

Вихідними рівняннями для багатодолинної моделі необхідно обирати рівняння балансу імпульсу та енергії електронного колективу у формі (4.2) і (4.3) для кожної долини. Відповідно рівняння балансу тільки для теплової енергії можна отримати підсумовуванням рівнянь типу (4.5) з ваговими коефіцієнтами, які

дорівнюють відносній заселеності долин b_i . Уводячи поняття середньої теплової енергії електронів $E_T = b_\Gamma E_{T\Gamma} + b_L E_{TL}$ (для стислості запису розглянемо тільки дводолину Γ - L -модель), отримаємо для обчислення її значення у стаціонарному стані

$$E_T = E_0 + b_\Gamma E_{d\Gamma} (\gamma_\Gamma - 1) + b_L E_{dL} (\gamma_L - 1). \quad (4.10)$$

Звідси й можна установити зв'язок між середньою для долин електронною температурою і напруженістю електричного поля:

$$T_e = T + \mathcal{E}^2 [b_\Gamma \mu_\Gamma^2 m_\Gamma^* (\gamma_\Gamma - 1) + b_L \mu_L^2 m_L^* (\gamma_L - 1)] / 3k_B. \quad (4.11)$$

Так само, як для однодолинної моделі, залежність електронної температури від напруженості електричного поля зручніше знаходити як обернену до $\mathcal{E}(T_e)$:

$$\mathcal{E} = \sqrt{\frac{3k_B(T_e - T)}{b_\Gamma \mu_\Gamma^2 m_\Gamma^* (\gamma_\Gamma - 1) + b_L \mu_L^2 m_L^* (\gamma_L - 1)}}, \quad (4.12)$$

оскільки параметри b_i , μ_i , γ_i є функціями електронної температури.

Тут потрібно мати на увазі, що в останній вираз входить деяка усереднена по долинах температура. Але, як вже відзначалося, через відмінності параметрів електронів у різних долинах міра розігрівання в них може істотно різнитись.

Для **арсеніду галію** скласти уявлення про часи релаксації імпульсу та енергії в Γ - та L -долинах можна, аналізуючи відповідні швидкості розсіювання на рис. 2.14 та рис. 2.21. Тут слід звернути увагу на те, що домішкове розсіювання, як пружне, не бере участі в розсіюванні енергії і є суттєвим тільки для Γ -долини, причому його роль слабшає із підвищенням електронної температури, а акустичне впливає, практично, тільки на час релаксації імпульсу. Міждолинне розсіювання стає помітним у Γ -долині лише для температур, вищих за ~ 1000 К, а в L -долині можна знехтувати розсіюванням у Γ -долину через низьку густину станів у ній.

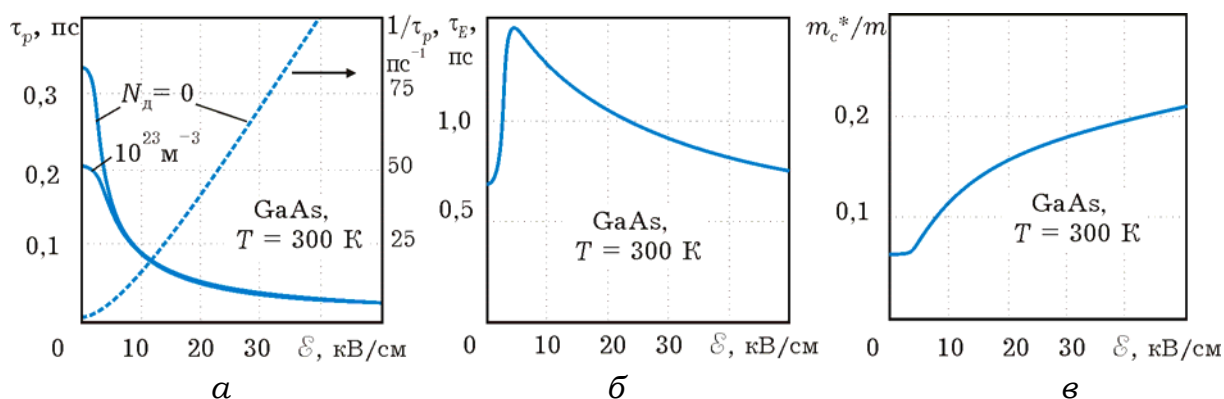


Рис. 4.3 Залежності усереднених по долинах часів релаксації імпульсу (а), енергії (б) та зведеної ефективної маси (в) від напруженості електричного поля для арсеніду галію. Штрихова лінія (а) – обернений час релаксації імпульсу

Знання часів релаксації в окремих долинах і розрахунок заселеностей долин (4.9) з використанням (3.20) дозволяє знайти їх усереднені по долинах значення (рис. 4.3), а також залежність між електронною температурою та напруженістю електричного поля, а за нею визначити поле-температурну характеристику $T_e(\mathcal{E})$. На цій залежності (рис. 4.4, а) можна виділити три характерні ділянки: область слабких полів з напруженістю менше 4 кВ/см, де розігрівання визначається, практично, тільки параметрами Γ -долини, перехідну область (до 20 – 40 кВ/см), у якій суттєво змінюється заселеність долин, та область дуже сильних полів, для якої характерне заповнення електронами переважно верхніх долин (рис. 4.4, б) та повільне, майже лінійне зростання електронної температури.

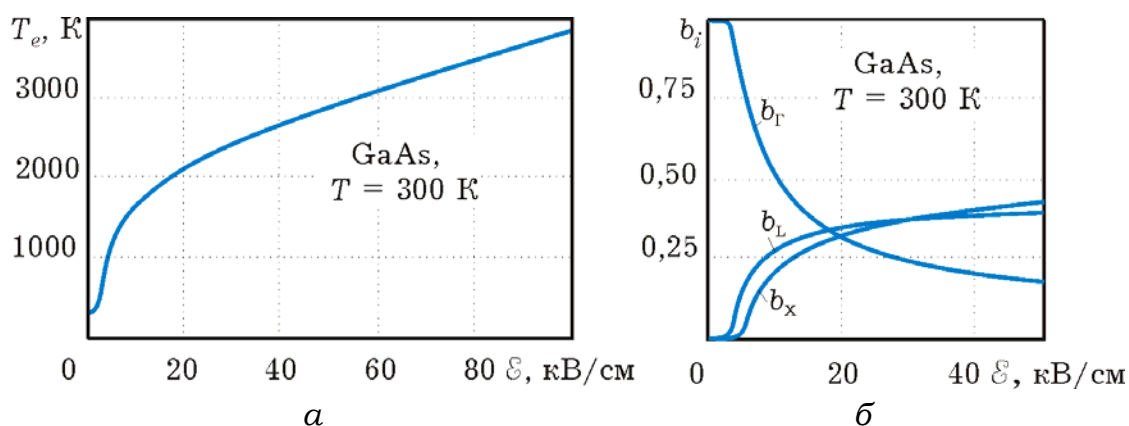


Рис. 4.4. Залежності електронної температури (а) й заселеності долин (б) від напруженості електричного поля в арсеніді галію для кімнатної температури

Нітрид галію. Аналогічна ситуація спостерігається і для кубічної структури нітриду галію: переважаючими є механізми оптичного розсіювання; акустичне, через пружність, практично не впливає на часи релаксації і є значимим лише для X-долин. Міждолинне розсіювання досить вагомо проявляється у Γ -долині у разі підвищення температури до ~ 2500 К.

Незважаючи на суттєво інші міждолинні відстані у законі дисперсії (рис. 1.26) і розташування долин, подібна, але кількісно дещо інакша картина спостерігається для поле-температурної характеристики нітриду галію, в якому помітне зростання електронної температури проявляється лише з позначки 120 – 125 кВ/см. А в області сильних полів (починаючи з 200 кВ/см) відбувається досить швидко зростання електронної температури і досягаються її значення порядку $\sim 10^4$ К (рис. 4.5, а).

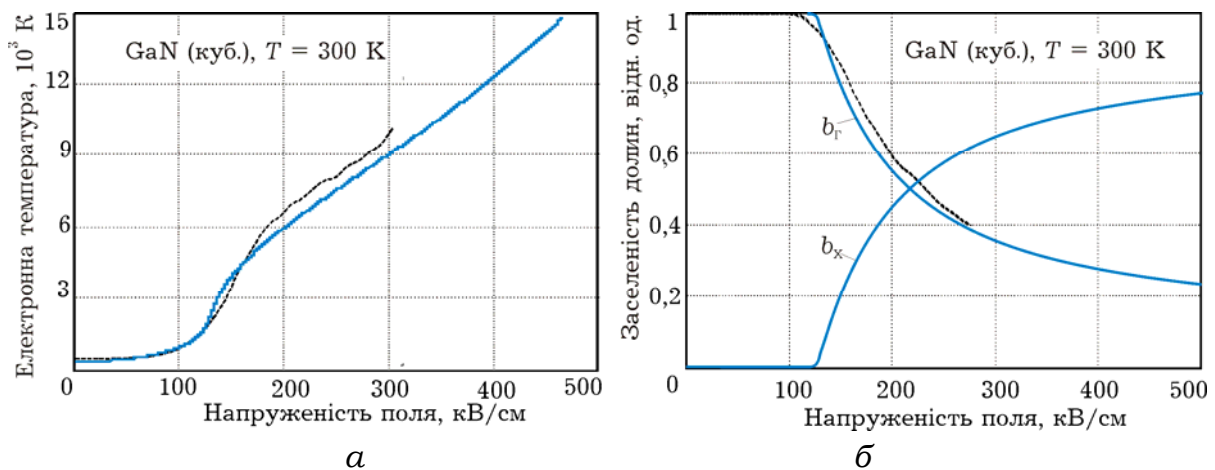


Рис. 4.5. Поле-температурна залежність (а) і залежність заселеності долин від напруженості електричного поля (б) в GaN для кімнатної температури. Штрихові лінії - довідникові

Отримані таким чином усереднені по долинах параметри у вигляді функцій від \mathcal{E} дозволяють використовувати рівняння однодолинної моделі для багатодолинних напівпровідників у наближенні заданого поля.

4.1.2. Насичення дрейфової швидкості

Розігрівання електронного газу в сильних електричних полях супроводжується так званим *насиченням дрейфової швидкості*, тобто виникненням ситуації, коли спрямована швидкість стає

практично незалежною від напруженості поля, а на вольтамперній характеристиці з'являється горизонтальна ділянка.

Найпростіше уявлення про причини такого явища можна скласти на прикладі одночасткової моделі дрейфу (розд. 2), де середній час вільного пробігу обернено пропорційний середній швидкості, яка має як хаотичну (теплову), так і спрямовану (дрейфову) складову. У слабких електричних полях $v_{др} \ll v_{хаот}$ і середня швидкість визначається тільки хаотичною складовою. Тому рухливість залишається постійною, а вольтамперна характеристика – лінійною. Але в міру того, як із зростанням \mathcal{E} дрейфова швидкість стає співмірною з хаотичною, середня швидкість на вільному пробігу зростає як за рахунок дрейфової складової, так і за рахунок хаотичної, яка збільшується із розігріванням електронного газу. У результаті зменшується час вільного пробігу і рухливість, а вольтамперна характеристика стає похилішою з тенденцією до насичення. Звідси випливає і можливий критерій для визначення терміну «*сильне електричне поле*»: у ньому дрейфова та хаотична складові швидкості електронів співмірні.

Повніше описати це явище можна з використанням колективної моделі дрейфу. Як видно на прикладі кремнію (рис. 4.1, а), обернений час релаксації імпульсу зростає зі збільшенням напруженості поля практично лінійно. Це спостерігається для $\mathcal{E} > 10...20$ кВ/см, коли переважаючим стає розсіювання за участю оптичних фононів, причому здебільшого з їх збудженням. За цих умов дрейфова швидкість електронів не може перевищити деяке значення v_n , оскільки чим більша напруженість поля, тим пропорційно вища швидкість розсіювання. Тому тільки-но прискорюваний полем електрон досягне певної (в середньому!) швидкості, що дорівнює швидкості насичення v_n , з великою ймовірністю відбудеться розсіювання із збудженням оптичного фонуна і відповідним зменшенням енергії електрона на величину $\hbar\omega_0$. Отже, зростання напруженості електричного поля приводить до збільшення не дрейфової, а хаотичної швидкості, тобто електронної температури.

Деяким аргументом на користь такого уявлення є той факт, що в германії (див. далі), у якому переважає акустичне розсіювання, перехід до режиму насичення помітно «затягнуто» в область високих електронних температур, а в режимі насичення спостерігається помітне підвищення швидкості.

Щоб зрозуміти чинники, які визначають швидкість насичення, іноді використовують таку просту модель. Відповідно до рівняння

балансу імпульсу (4.2), розсіяний за час τ_p імпульс $m^* v_H$ дорівнює набутому від поля $-e\mathcal{E}\tau_p$, а відповідно до рівняння балансу енергії (4.3) для дрейфового руху набута від поля за той самий час енергія дрейфу $e v_H \mathcal{E} \tau_p$ у *стаціонарному стані* дорівнює енергії випромінюваного фонона $\hbar\omega_0$. Вилучивши з цих рівностей напруженість електричного поля, отримаємо

$$v_H = (\hbar\omega_0/m^*)^{1/2}. \quad (4.13)$$

Це співвідношення, використовуване зазвичай у навчальній літературі, слід сприймати лише як оціночне. Щоб отримати розрахункові співвідношення, потрібно враховувати особливості механізмів розсіювання у конкретних матеріалах.

Кремній. Помноживши співвідношення (4.8) на рухливість, одержимо вираз для дрейфової швидкості:

$$v = v_T \sqrt{\frac{T_e/T - 1}{\gamma - 1}}. \quad (4.14)$$

Цей вираз дозволяє, розрахувавши часи релаксації з урахуванням отриманого раніше зв'язку між електронною температурою і напруженістю електричного поля, отримати *поле-швидкісну характеристику* $v(\mathcal{E})$. Особливістю цього співвідношення є те, що коефіцієнт непружності також пропорційний температурі $\gamma = \beta T_e / T$.

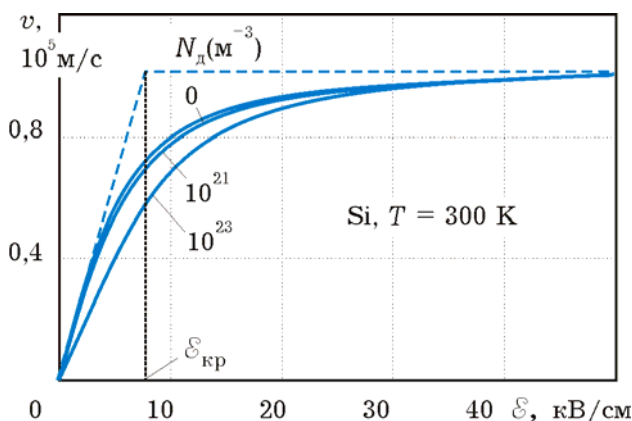


Рис. 4.6. Залежність дрейфової швидкості від напруженості електричного поля у кремнії за різних концентрацій домішки та її кусково-лінійна апроксимація (штрихова лінія)

Для кремнію її подано на рис. 4.6: у слабких полях швидкість пропорційна напруженості, а в сильних ($\mathcal{E} > 20$ кВ/см) маємо дуже похилу ділянку, яка і відповідає насиченню. У цьому режимі

практично не спостерігається залежності швидкості від концентрації домішки, що говорить про незначну роль домішкового розсіювання і, відповідно, про переважання фононних механізмів у сильних електричних полях.

За відомих значень v_n і слабкопольової рухливості μ_0 іноді використовують кусково-лінійну апроксимацію (див. рис. 4.6) поле-швидкісної характеристики $v(\mathcal{E})$:

$$v = \mu_0 \mathcal{E} \text{ для } \mathcal{E} < \mathcal{E}_{\text{кр}}, \quad v = v_n \text{ для } \mathcal{E} > \mathcal{E}_{\text{кр}}, \quad (4.15)$$

де $\mathcal{E}_{\text{кр}} = v_n / \mu_0$ – критична напруженість поля, що відповідає умовній межі між сильними та слабкими полями. Така апроксимація зручна, оскільки використовувані параметри v_n та μ_0 можна відносно просто ідентифікувати.

Ще зручнішою і досить точною може виявитися явна функція $v(\mathcal{E}) = \mu(\mathcal{E})\mathcal{E}$, у якій залежність рухливості від поля задано наближеним виразом

$$\mu(\mathcal{E}) = \mu_0 / [1 + (\mathcal{E} / \mathcal{E}_{\text{кр}})^2]^{1/2}. \quad (4.16)$$

Цю залежність можна використовувати тільки для фононних механізмів розсіювання. Вплив домішки тут помітний у слабких полях і його можна наближено врахувати, якщо в чисельнику замість μ_0 використати рухливість $\mu(N_d)$ для певної концентрації домішки, але $\mathcal{E}_{\text{кр}}$ слід розраховувати через μ_0 .

Германій. Переважання акустичного розсіювання викликає, як зазначалось у попередньому розділі, нелінійність залежності $\gamma(T_e)$. Але не зважаючи на це, загальний коефіцієнт непружності зручно подати, як і для кремнію, тільки в коефіцієнті непружності параметр β залежатиме від напруженості.

У підсумку можна вважати, що відповідно до (4.14) при $T_e / T \gg 1$ настає насичення, але в режимі насичення спостерігається відносно повільне зростання швидкості дрейфу. Це уявлення підтверджують результати розрахунків, наведені на рис. 4.7, де зображено поле-швидкісну характеристику (а) і залежність рухливості від напруженості електричного поля (б), отримані з

урахуванням поле-температурної характеристики $T_e(\mathcal{E})$ (рис. 4.2, в).

На рисунках також наведено результати експериментів, отримані з різних першоджерел.

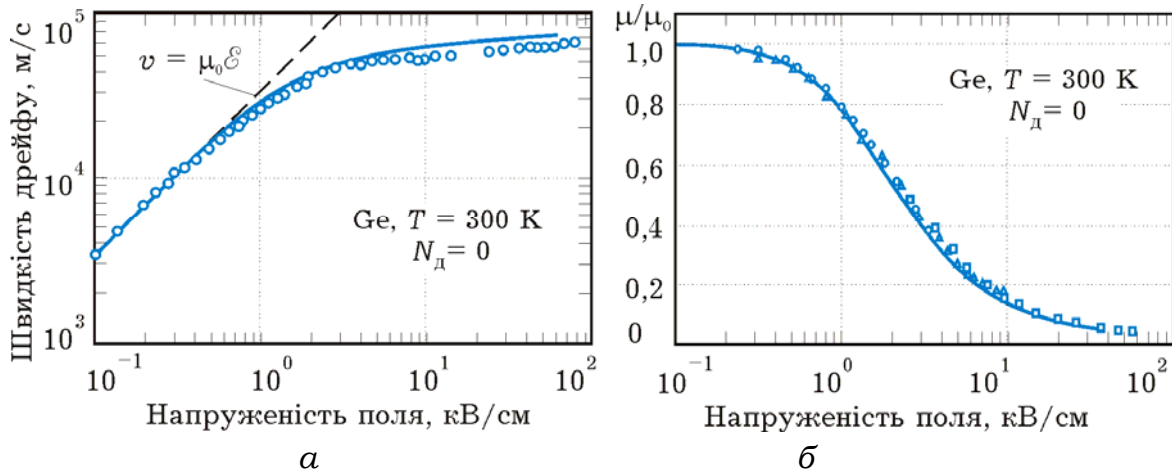


Рис. 4.7. Залежності дрейфової швидкості (а) і нормованої на слабо-польове значення рухливості (б) від напруженості електричного поля в германії за кімнатної температури

Арсенід галію. Поле-швидкісна характеристика GaAs має принципово інший вигляд, що пов'язано передусім з багатодолинною структурою закону дисперсії, а також з особливостями механізмів розсіювання у різних долинах (рис. 2.14, а). На рис. 4.8, а наведено поле-швидкісні характеристики для різних долин. У кожній з них спостерігається насичення, але суттєво різне через відмінність перш за все ефективних мас. Для Γ -долини швидкість $2 \cdot 10^5$ м/с встановлюється за напруженості біля 5 кВ/см. У верхніх долинах швидкість насичення більш ніж вдвічі менша для напруженості майже на порядок більшої.

Для знаходження середньої швидкості треба прийняти до уваги кількість електронів у долинах (рис. 4.4, б), яка змінюється за рахунок міждолинного розсіювання. Видно, що зі збільшенням напруженості внесок електронів Γ -долини зменшується, а верхніх долин збільшується.

Середню швидкість електронів знаходять як суму швидкостей в окремих долинах з ваговими коефіцієнтами, що дорівнюють відповідним заселеностям, які також можна подати функціями напруженості поля, використовуючи попередньо знайдену

залежність (4.12) між T_e та \mathcal{E} . На рис. 4.8, б подано результуючу поле-швидкісну характеристику і внески до неї електронів кожної долини, а також результати (штрихові лінії) від 0 до 10 кВ/см, взяті з літературних джерел.

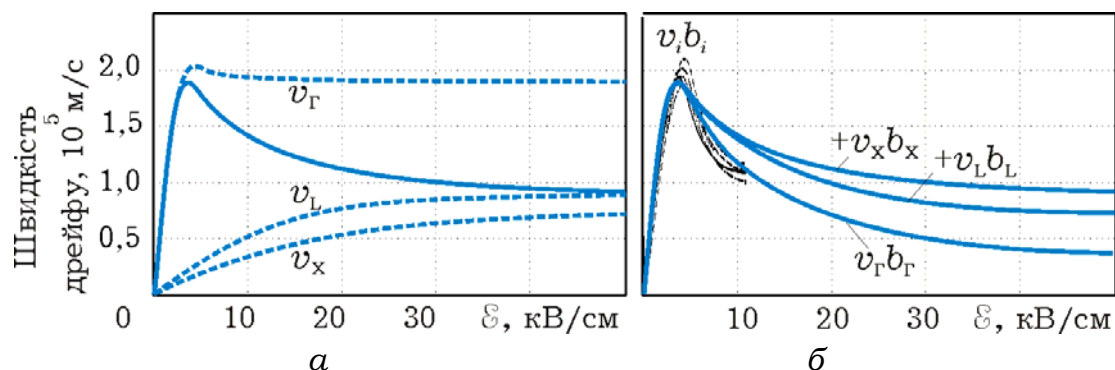


Рис. 4.8. Залежність від напруженості електричного поля швидкості дрейфу електронів для кожної долини та їх усереднена залежність (суцільна лінія) (а). Та сама залежність (б), побудована підсумовуванням швидкостей в окремих долинах з ваговими коефіцієнтами, рівними їх заселеностям (див. рис. 4.4, б). Штрихові лінії (б) від 0 до 10 кВ/см - довідникові

Нітрид галію з кубічною кристалічною структурою має значно меншу рухливість у слабкому полі перш за все через більшу ефективну масу. Найістотнішою відміною його поле-швидкісної характеристики є те, що перехід до режиму насичення відбувається за значно вищої напруженості електричного поля біля 100 кВ/см. Крім того, різке спадання дрейфової швидкості, викликане інтенсифікацією переходів у верхні долини, починається для напруженості близько 150 кВ/см, що відповідає міждолинній ГХ-відстані в 1,3 еВ.

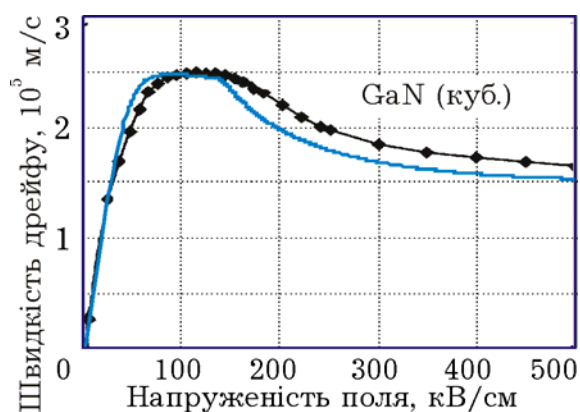


Рис. 4.9. Поле-швидкісні характеристики нітриду галію з кубічною структурою. Крива з позначками взята з літературних джерел

Наведені на рис. 4.8 та 4.9 параметри характеристик вказують на особливості цих матеріалів та на перевагу нітриду галію для використання у приладах підвищеної потужності через більшу швидкість і напруженість електричного поля.

4.1.3. Динамічна провідність

Диференціальна провідність. Статична питома електропровідність визначається як коефіцієнт пропорційності між густиною струму провідності й напруженістю електричного поля: $j = \sigma \mathcal{E}$. Відповідно визначимо динамічну (диференціальну) провідність як коефіцієнт пропорційності між змінами dj і $d\mathcal{E}$: $\sigma_d = dj/d\mathcal{E}$. Оскільки густина струму $j = env$, то за незмінної концентрації $dj = endv$. У слабких електричних полях швидкість дрейфу пропорційна напруженості поля $v = \mu \mathcal{E} = (e\tau_p/m^*)\mathcal{E}$, а оскільки рухливість постійна, то $dv = \mu d\mathcal{E}$ і диференціальна провідність збігається зі статичною. У сильних полях через особливості оптичного розсіювання відбувається насичення швидкості, зменшується час релаксації імпульсу і, відповідно, рухливість. За цих умов із стаціонарного рівняння балансу імпульсу випливає, що $(e/m^*)d\mathcal{E} = (\tau_p^{-1})dv + vd(\tau_p^{-1})$. У режимі насичення $dv = 0$ і $d(\tau_p^{-1}) = (e/m^*v_H)d\mathcal{E}$. Якщо цей зворотний зв'язок між рухливістю і напруженістю поля екстраполювати на перехідну область, то $dv = \mu(1 - v/v_H)d\mathcal{E}$ і

$$\sigma_d = \sigma(1 - v/v_H). \quad (4.17)$$

У слабких полях $v \ll v_H$ і диференціальна провідність збігається зі статичною, але в міру наближення до режиму насичення зменшується до 0. Отримане співвідношення (4.17) придатне для однодолинних напівпровідників (Si, Ge), якщо концентрація електронів у зоні залишається незмінною.

Характерною рисою арсеніду галію й інших багатодолинних напівпровідників є наявність спадної ділянки на залежності дрейфової швидкості від напруженості електричного поля.

Пояснюється це тим, що з розігріванням збільшується частка електронів, які під час міждолинного розсіювання переходять у верхні долини, де їхня ефективна маса стає більшою. У результаті зменшується середня дрейфова швидкість i , відповідно, струм провідності. Така ситуація приводить до того, що позитивні зміни поля викликають негативні зміни струму провідності в матеріалі, тобто *від'ємну динамічну провідність* (ВДП). На рис. 4.10 (штрихові лінії) наведено залежності такої провідності від напруженості поля, отримані чисельним диференціюванням функції $v(\mathcal{E})$.

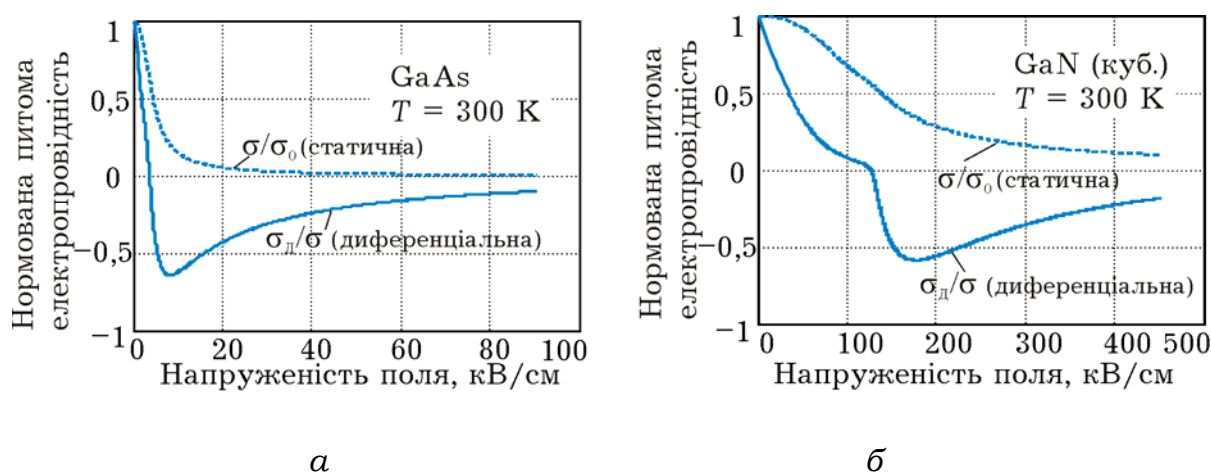


Рис. 4.10. Залежності від напруженості електричного поля питомої електропровідності для арсеніду галію (а) і нітриду галію (б): статичної, нормованої на її слабопольове значення (пунктирна лінія); диференціальної, нормованої на статичну й отриману чисельним диференціюванням функції $v(\mathcal{E})$ (суцільна лінія)

На рис. 4.10 наведено нормовані залежності σ_d/σ від напруженості електричного поля отримані чисельним диференціюванням кривих $v(\mathcal{E})$. Найхарактерніша особливість одержаних результатів – наявність області ВДП, що виявляється практично відразу після початку міждолинних переходів ($\mathcal{E} \approx 4\text{ кВ/см}$ для арсеніду галію та біля 120 кВ/см для нітриду галію) і набуває максимального значення, коли концентрації електронів у верхній і нижній долинах приблизно однакові ($b_{\Gamma}b_L \approx 0,25$).

Розрахована таким чином диференціальна провідність придатна лише для *повільних* змін електричного поля і, відповідно, струму. Із підвищенням швидкості змін поля (частоти) необхідно враховувати інерційність таких процесів: запізнювання зіткнень, що характеризується часом релаксації імпульсу (τ_p), релаксації енергії (τ_E), міждолинних переходів (τ_n), використовуючи для цього нестационарні рівняння збереження.

Високочастотна провідність. Для високочастотних змін електричного поля визначальними є прояви інерційних властивостей різних фізичних ефектів. Обмежувачись розглядом так званої малосигнальної провідності, насамперед необхідно зазначити, що провідність, як і в розд. 2.3.1, має комплексний характер, що проявляється на високих частотах. На рис. 4.11 наведено частотні залежності дійсної (а) і уявної (б) частин електропровідності, нормованої на її статичне значення $\sigma = e^2 n \tau_{p0} / m^*$.

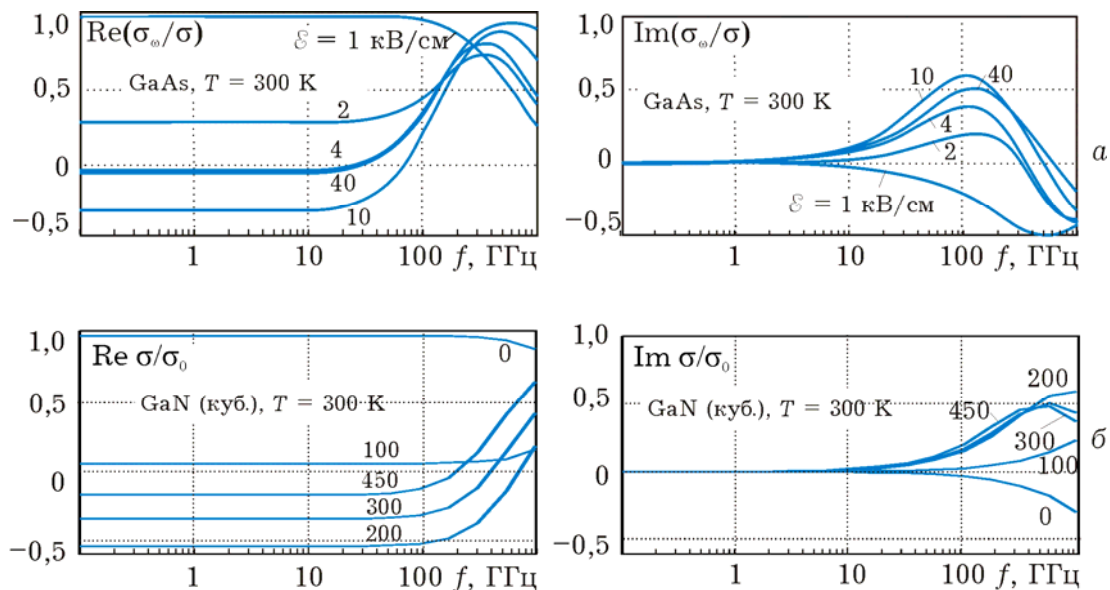


Рис. 4.11. Частотні залежності дійсної та уявної частин динамічної питомої електропровідності, нормованої на її статичне значення для арсеніду галію (а) і нітриду галію (б)

Частотна залежність σ_ω визначається доданками, що містять, як і (2.44), множники типу $(1 + i\omega\tau)^{-1}$, що включають три часові константи τ_p , τ_E і τ_m . Перший множник залежить від частоти

зіткнень, тобто від часу релаксації імпульсу, і пояснюється інерційністю змін струму зі зміною поля. Другий множник пов'язаний з переходом дрейфової швидкості в режим насичення, що вимагає участі оптичних фононів, і можливий тільки у разі підвищення температури електронів. Інерційність процесів розігрівання визначається часом релаксації енергії, тому другий множник містить τ_E . Для процесів, пов'язаних з перерозподілом електронів по долинах, характерна інерційність як процесів розігрівання, так і міждолинних переходів. Тому в третій множник входить час міждолинної релаксації τ_m . Важливо зазначити, що він пропорційний енергетичній відстані Δ_{ij} між долинами і саме він визначає від'ємні значення високочастотної електропровідності.

Із цих трьох часів найбільше значення (біля 1 пс для GaAs і близько до 0,3 пс для GaN) має τ_E , тому верхня частотна межа ВДП залежить насамперед від інерційності процесів розігрівання електронів і становить десятки (для GaAs) та сотні (для GaN) гігагерц. У слабкому полі (верхні криві на рис. 4.11) процеси розігрівання і міждолинних переходів несуттєві, тому спад провідності визначається значенням τ_p і відбувається на частотах порядку 100 ГГц.

Реактивна провідність стає помітною на більш низьких частотах (порядку одиниць або десятків ГГц), що пояснюється особливостями множників $(1 + i\omega\tau)^{-1} = (1 - i\omega\tau)/(1 + \omega^2\tau^2)$: реактивна провідність, на відміну від активної, пропорційна першому степеню $\omega\tau$. У слабкому полі вона має індуктивний характер, але в області ВДП змінює знак разом із активною складовою і стає ємнісною. Далі, на дуже високих частотах за час періоду поля міжзонні переходи не встигають відбуватися і реактивна провідність знову стає індуктивною. А оскільки більшість електронів знаходиться у верхніх долинах, значення провідності визначається параметрами цих долин.

Важливим результатом проведеного аналізу є можливість розрахунку граничної частоти існування ВДП. Його можна провести, знехтувавши інерційністю міждолинних переходів ($\omega\tau_m \ll 1$), запізнюванням зіткнень ($\omega\tau_p \ll 1$) і вважаючи, що в режимі ВДП швидкість електронів досягає насичення. Зручно критерієм

граничної ситуації обрати умову $\text{Re } \sigma_\omega = 0$, яка і дозволяє одержати для граничної частоти

$$f_m = \frac{1}{4\pi\tau_{E0}} \sqrt{\frac{\Delta_{ji}}{k_B T_e}}. \quad (4.18)$$

Як видно з рис. 4.12, гранична частота, яка визначається здебільшого часом релаксації енергії, має значення порядку десятків гігагерц для арсеніду галію, зменшуючись із зростанням напруженості електричного поля, що в цілому відповідає відомим результатам. Але для нітриду галію граничні частоти існування ВДП на порядок вищі.

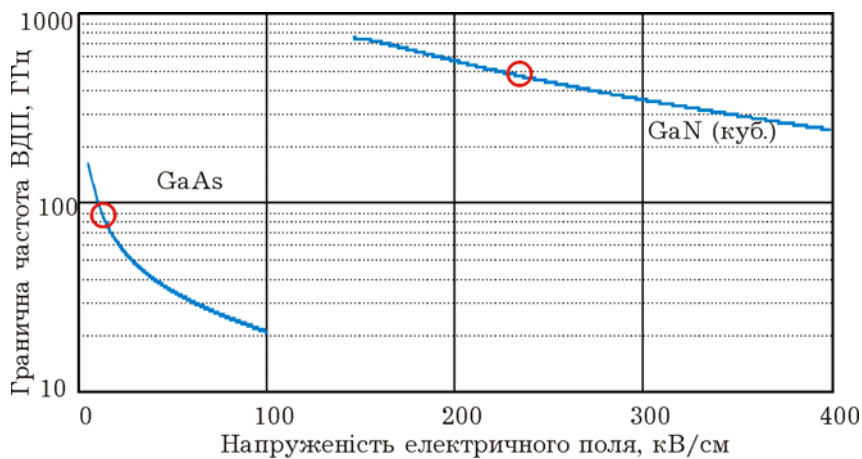


Рис. 4.12. Залежність максимальної частоти існування ВДП від напруженості електричного поля. Позначено напруженості, які відповідають максимальним значенням ВДП

4.1.4. «Сплеск» дрейфової швидкості

Дрейфові процеси виявляють цікаві особливості в імпульсному електричному полі. У разі різкого збільшення напруженості поля значення швидкості дрейфу може в кілька разів перевищувати стаціонарне. Це явище відоме як «сплеск» (*overshoot*) дрейфової швидкості. Воно пов'язане з інерційністю процесів розсіювання та із залежністю часів релаксації від напруженості поля. Для розуміння такої поведінки електронів слід звернути увагу на деякі деталі релаксаційного процесу (рис. 2.8, б), який після вмикання електричного поля описується формулою (2.52):

$$v(t) = \mu \mathcal{E} \{1 - \exp[-(t - t_0^+)/\tau_p]\}.$$

На початковій ділянці похідна цієї функції, тобто прискорення, визначається як для вакууму відношенням $e\mathcal{E}/m^*$, а швидкість зростає за законом $v(t) = (e/m^*)\mathcal{E}(t - t_0^+)$, який можна екстраполювати прямою $1'$ на рис. 4.13, б. У слабкому полі \mathcal{E}_1 за час вільного пробігу, або практично за час релаксації імпульсу τ_{p1} , вона досягла б значення стаціонарної дрейфової швидкості $(e/m^*)\mathcal{E}_1\tau_{p1} = \mu_1\mathcal{E}_1$. Тобто в процесі наростання швидкість прагне до значення, що визначається високою рухливістю у слабкому полі і напруженістю \mathcal{E}_1 .

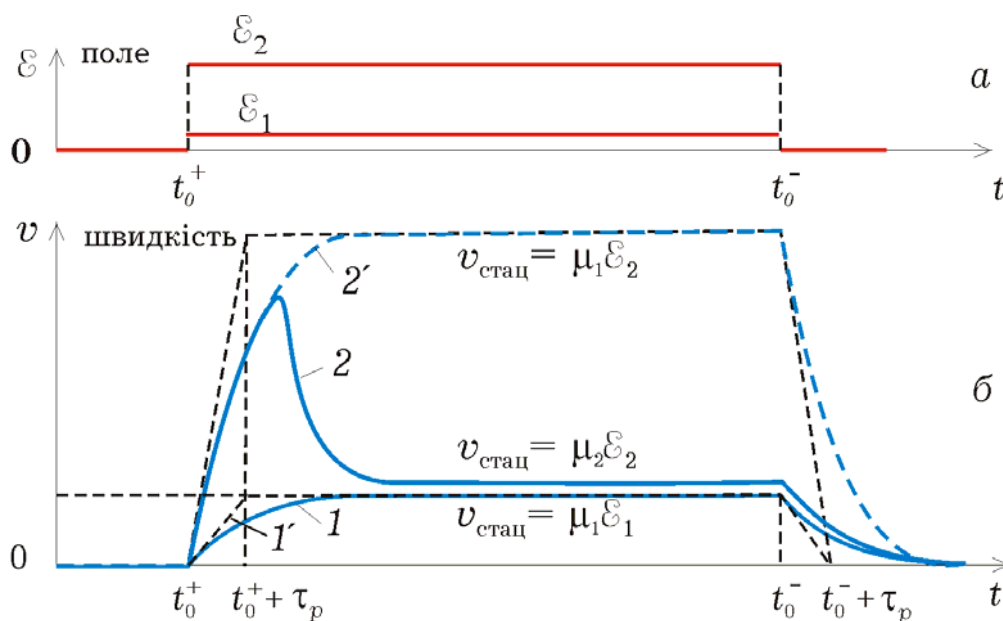


Рис. 4.13. Динамічні (часові) залежності швидкості дрейфу (б) з увімкненням імпульсу слабого (1) і сильного (2) електричного поля (а)

Включення імпульсу сильного електричного поля \mathcal{E}_2 викликає відповідно більше прискорення електронів і в момент часу $t_0^+ + \tau_{p1}$ їх швидкість досягла б значення $(e/m^*)\mathcal{E}_2\tau_{p1} = \mu_1\mathcal{E}_2$, яке відповідає високій рухливості у слабкому полі і великому значенню напруженості (крива $2'$).

Однак цього значення швидкість досягти не встигає, бо одночасно із зростанням швидкості відбувається подальше

розігрівання електронного газу з переважанням майже пружного акустичного й частково полярного оптичного розсіювання, що спричиняє значне зменшення часу вільного пробігу i , відповідно, рухливості (див. розд. 4.1.1).

Така зміна τ_p викликає швидке зменшення швидкості дрейфу, причому після різкого спаду швидкість може падати нижче стаціонарного значення для \mathcal{E}_2 , а після закінчення імпульсу поля вона може бути навіть від'ємною. Це явище, відоме як *від'ємний «сплеск» (undershoot)*, пов'язано з різними часами релаксації електронів, які рухаються вздовж поля або назустріч йому.

Інерційність такого процесу визначається не стільки процесами релаксації енергії, як цього можна було б очікувати з огляду на (2.42), скільки інтенсивною хаотизацією спрямованого руху. Швидкість буде зростати, доки триватиме *балістичний* рух, тобто рух без зіткнень. За час, близький до часу вільного пробігу, або часу релаксації імпульсу, відбудеться зіткнення усіх електронів, які «забувають» про свій попередній стан, і перехід до дрейфового руху.

Наведене пояснення підтверджують також результати чисельного моделювання перехідних процесів (рис. 4.14) для різних амплітуд імпульсу електричного поля. Із збільшенням напруженості поля, природно, зростає максимальне значення дрейфової швидкості (рис. 4.14, *a*), зменшення часу релаксації імпульсу і відповідно скорочення тривалості «сплеску» швидкості.

Аналогічні процеси відбуваються і в статичному, але *просторово неоднорідному* електричному полі. Це легко зрозуміти, враховуючи, що в загальному випадку рівняння збереження (4.2) та (4.3) містять повну похідну за часом від функцій координат і часу: $d/dt = \partial/\partial t + \vec{v}\partial/\partial\vec{r}$. У змінному в часі, але однорідному в просторі полі $\partial/\partial\vec{r} = 0$ й $d/dt = \partial/\partial t$, а в статичному, але просторово змінному полі $d/dt = \vec{v}\partial/\partial\vec{r}$. Тому рівняння i , відповідно, їх розв'язки матимуть подібний вигляд, якщо замінити $x = \int_{t_0}^t v(t') dt'$.

Ілюструє це рис. 4.14, *б*, на якому подано просторові залежності швидкості електронів, що влітають в статичне поле з певною напруженістю. Із графіків видно, що після вльоту

електрона в простір сильного поля спостерігається так званий *балістичний рух*, тобто рух без розсіювання, подібний до руху у вакуумі. Він триває кілька сотих і десятих долей пікосекунди, за цей час електрон встигає пролетіти відстань у декілька десятих долей мікрометра (порядку довжини вільного пробігу) і прискоритися до значень, які в декілька разів перевищують стаціонарні значення з поле-швидкісної характеристики (рис. 4.9). Після цього, як і в динамічному випадку, у результаті зіткнень, що починаються, зменшується час релаксації майже на порядок, а швидкість релаксує до значення, що визначається низьким значенням рухливості в сильному електричному полі.

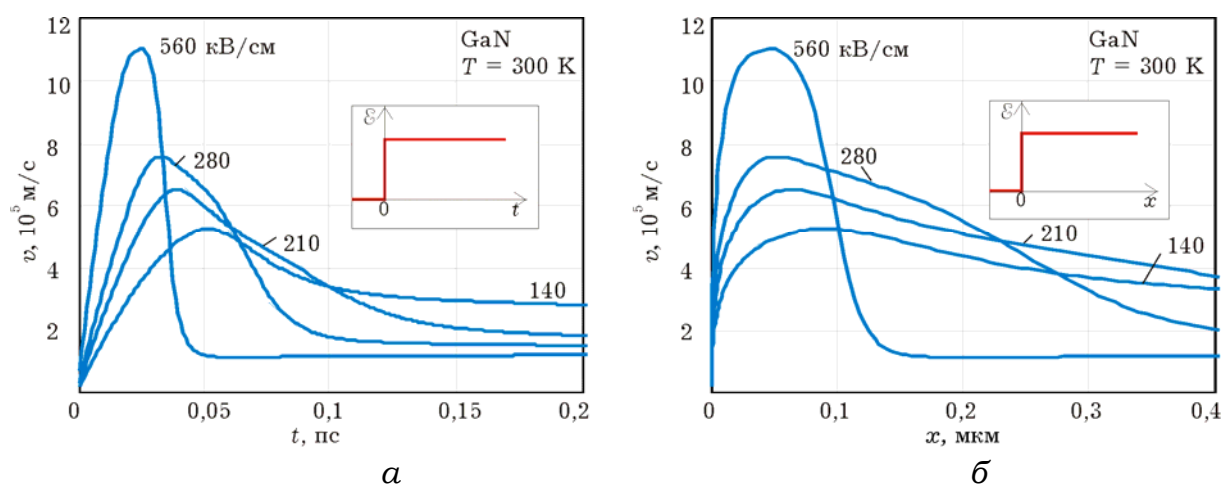


Рис. 4.14. Часова (а) та просторова (б) залежності дрейфової швидкості для різних значень амплітуди напруженості поля у нітриді галію

Ефект «сплеску» може виявлятися і в тому випадку, коли напруженість поля змінюється у просторі не стрибкоподібно, а відносно плавно. Він помітний на відстанях порядку $0,1$ мкм, тому у субмікронних приладах середню швидкість прольоту можна отримати істотно вищою за стаціонарну. Крім того, цей ефект можна використовувати для одержання надкоротких імпульсів.

4.1.5. Ефект Ганна

Ефект Ганна, один з ефектів сильного поля, проявляється у протіканні в однорідному зразку пульсуючого струму за постійної

напруги живлення. Ефект спостерігається у напівпровідниках, які подібно до арсеніду галію мають багатодолинний характер зонної структури. Нагадаємо, що в ньому на дисперсійній кривій (розд. 1.1.6), нижнім електронним станам біля дна зони провідності відповідає Γ -долина, розташована в центрі зони Бріллюена, яка характеризується великою кривизною й, отже, малою ефективною масою електронів: $m^* \approx 0,063m$. Стани найближчої верхньої L-долини, розташованої на 0,3 еВ вище дна зони провідності, відповідають майже на порядок більшій ефективній масі. Обернено пропорційно ефективній масі змінюється рухливість електронів із переходом їх з нижньої долини до верхньої. Характерною є істотно більша густина станів у верхній долині, викликана значною відмінністю ефективних мас. За звичайних температур, коли $k_B T < 0,3$ еВ, і в слабких електричних полях електрони заповнюють практично тільки стани нижньої долини й, маючи високу рухливість, забезпечують високу питому електропровідність і різке зростання вольт-амперної характеристики (початкова ділянка кривої на рис. 4.15, а).

У сильному електричному полі за напруженості вище критичної (для арсеніду галію $\mathcal{E}_{кр} \approx 4$ кВ/см) спостерігається розігрівання електронного газу, що викликає перехід електронів у верхню долину. У результаті на вольт-амперній характеристиці, що за формою повторює поле-швидкісну характеристику (див. рис. 4.8), спостерігається спадна ділянка, що відповідає ВДП напівпровідника. Такий зразок, що має від'ємний динамічний опір до змінного струму (рис. 4.11), будучи включеним у коливальний контур, компенсує активні втрати в ньому і визначає тим самим можливість існування незгасаючих коливань. Генерування коливань струму в однорідному зразку GaAs у разі прикладання до нього постійної напруги вперше спостерігалось Ганном, на ім'я якого й названо цей ефект.

Виникнення коливань струму в напівпровіднику викликано *нестійкістю*, зумовленою наявністю ВДП. Ця нестійкість виникає, коли до зразка прикладають напругу, достатню для створення критичного поля, що відповідає точці 1 на рис. 4.15, а.

Значення струму j_1 нестійке, тому що за рахунок неминуче існуючої неоднорідності зразка в деякій локальній області з підвищеним опором (точка x_d на рис. 4.15, б) напруженість поля

може дещо перевищувати значення $\mathcal{E}_{кр}$. Саме у цій точці починається процес зародження домену, викликаний підвищенням розігріванням електронів, що приводить до міждолинних переходів, зменшення їхньої рухливості й, відповідно, до деякого відставання гарячих електронів від основної їх маси в дрейфовому русі.

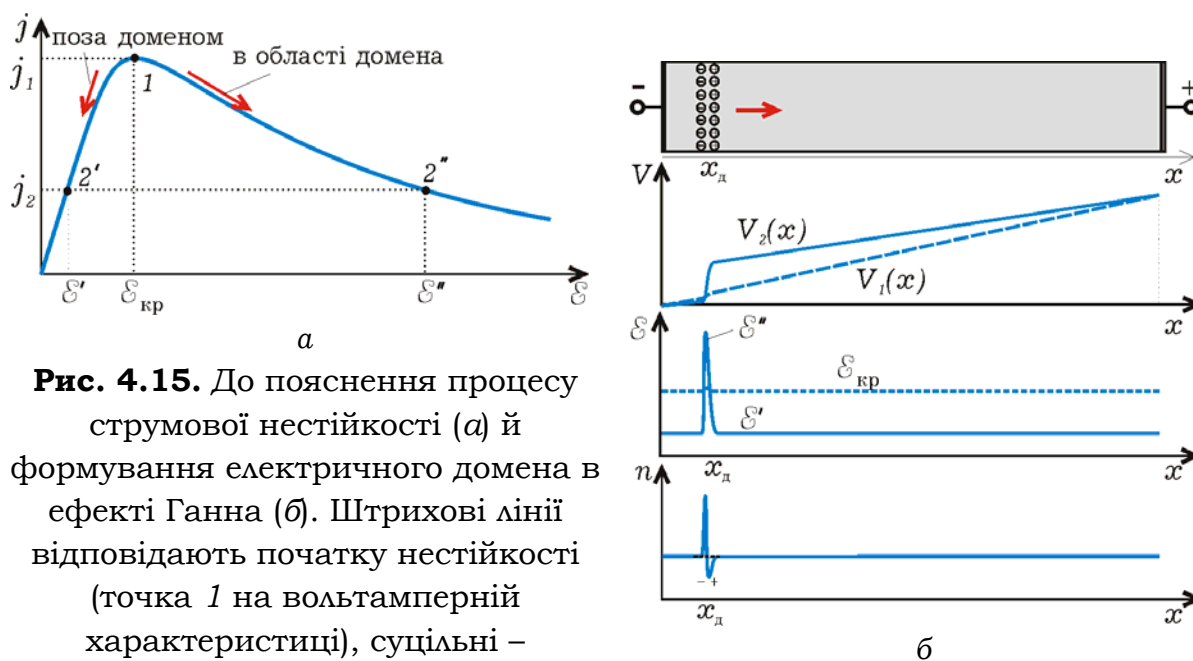


Рис. 4.15. До пояснення процесу струмової нестійкості (а) й формування електричного домена в ефекті Ганна (б). Штрихові лінії відповідають початку нестійкості (точка 1 на вольтамперній характеристиці), суцільні – утворенню домену (точки 2' і 2'')

Відмінність у дрейфовій швидкості електронів приводить до перерозподілу просторового заряду (рис. 4.15, б) і появи внутрішнього локального електричного поля, що збігається за напрямком із зовнішнім полем і ще більше посилює розігрівання електронного газу. Цей процес нестійкий і триватиме лавиноподібно доти, доки напруженість поля не досягне значення \mathcal{E}'' (рис. 4.15, а), за якого більшість електронів перейде у верхню долину. Точка 2'' відповідає стану стійкої рівноваги й визначається перетином вольт-амперної характеристики й навантажувальної прямої. Оскільки різниця потенціалів на зразку залишається постійною (рис. 4.15, б), то збільшення поля у локальній області приводить до його зменшення в іншій частині зразка до значення \mathcal{E}' , що відповідає струму j_2 .

У результаті описаного процесу відбувається формування *електричного домену*, тобто області просторового заряду, утвореного гарячими електронами, з підвищеною напруженістю електричного поля. Зазвичай домен не залишається нерухомим, а переміщується вздовж кристала зі швидкістю, рівною середній швидкості дрейфу електронів. Досягши анода, домен руйнується, напруженість поля відновлюється до критичного значення $\mathcal{E}_{\text{кр}}$ і формування домену починається спочатку. Повторюючись періодично, цей процес викликає коливання густини струму (рис. 4.16) у межах між значеннями j_1 та j_2 .

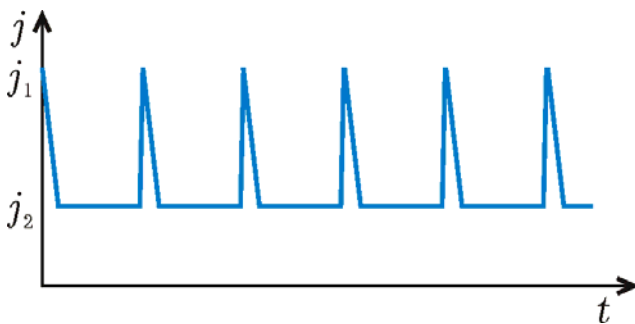


Рис. 4.16. Часова діаграма зміни струму в ефекті Ганна

Інерційність процесу формування домену визначається здебільшого часом релаксації енергії та максвеллівським часом релаксації просторового заряду і становить пікосекунди. Тому у відносно довгих зразках період коливань визначається переважно часом переміщення домену від місця формування до анода. Оскільки область катодного контакту дуже неоднорідна, то домен зазвичай формується на катоді або в спеціально сформованій області біля катода. Якщо, наприклад, довжина зразка – 100 мкм, а середня швидкість дрейфу – 10^5 м/с, то час його руху до анода становить 10^{-9} с, а коливання струму відбуваються із частотою 10^9 Гц. З огляду на простоту конструкції можна зрозуміти, чому діоди з ефектом Ганна набули значного поширення як генератори надвисокочастотних коливань.

4.1.6. Дифузія гарячих електронів

З особливостями механізму розігрівання електронів також пов'язана характерна для багатодолинних напівпровідників

залежність коефіцієнта дифузії D від напруженості електричного поля (рис. 4.17).

Зрозуміти її можна, використовуючи співвідношення Ейнштейна: $D = \mu k_B T_e / e$. Враховуючи залежність рухливості від електронної температури та залежність електронної температури від напруженості поля, можна пояснити хід функції $D(\mathcal{E})$: у слабких полях час релаксації імпульсу, а отже й рухливість, майже не залежать від напруженості, тому зростання коефіцієнта дифузії пояснюється тільки підвищенням температури електронного газу. У сильних полях, де переважають оптичне та міждолинне розсіювання, розігрівання не таке значне і коефіцієнт дифузії спадає через зменшення часу релаксації імпульсу і, відповідно, рухливості

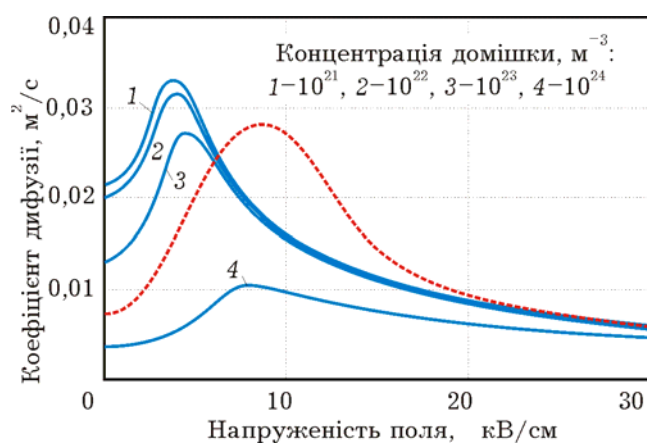


Рис. 4.17. Залежності коефіцієнта дифузії від напруженості електричного поля для різних концентрацій домішки для арсеніду галію. Штрихова лінія взята з наукової літератури

Максимальне значення коефіцієнта дифузії має за напруженості поля, близької до критичної. Вплив концентрації домішки спостерігається практично тільки в слабкому полі через збільшення домішкового розсіювання і відповідне зменшення рухливості за більш високого рівня легування.

Слід зазначити, що наявні в науковій літературі експериментальні та теоретичні залежності $D(\mathcal{E})$, загалом маючи однаковий вигляд, різняться між собою як максимальним значенням (до $0,1 \text{ м}^2/\text{с}$), так і положенням максимуму. Це пов'язано, найімовірніше, з різними методами вимірювань та

деякою невизначеністю умов експерименту (концентрації домішки).

4.2. ПОЛЬОВА ГЕНЕРАЦІЯ

Польовою генерацією називають процес утворення рухливих носіїв (електронів і дірок) у речовині або вільних зарядів (електронів й іонів) у вакуумі або газі під дією сильного електричного поля. Тут розглянемо такі ефекти сильного поля, за яких воно впливає на кількість носіїв.

4.2.1. Ударна іонізація

Ударна іонізація – це один з видів розсіювання, що приводить до утворення нових рухливих носіїв. Електричне поле, викликаючи спрямований потік зарядів, збільшує їх кінетичну енергію. У сильному електричному полі частинки, наприклад, електрони, можуть нагромадити енергію, достатню для того, щоб під час зіткнення з нейтральними атомами (кристалічною ґраткою) акт розсіювання завершився утворенням нових носіїв. Електрони, що утворюються при цьому, також прискорюються електричним полем і можуть, у свою чергу, породжувати нові заряджені частинки. Такий процес розмноження носіїв називають *лавинним*. Процеси лавинного розмноження носіїв заряду відбуваються у плазмі твердого тіла або газоподібного середовища.

Найпростіший аналіз процесу лавинного розмноження можна провести у такий спосіб. Виділимо в середовищі шар завтовшки dx (рис. 4.18). Нехай через одиницю поверхні в цей шар надходить за одиницю часу $n(x)$ електронів, а виходить на dn більше.

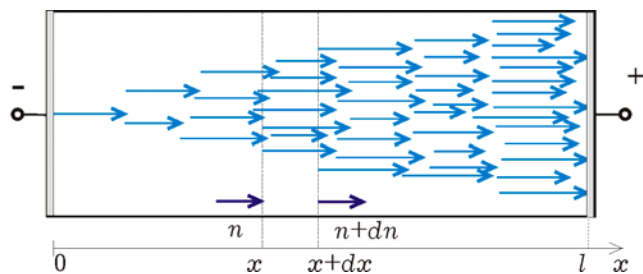
Природно вважати, що це збільшення пропорційне кількості електронів, що надходять, і товщині шару

$$dn = \alpha n(x) dx. \quad (4.19)$$

Фізичний зміст коефіцієнта пропорційності α , названого *коефіцієнтом іонізації*, видно з (4.19):

$$\alpha = dn/ndx, \quad (4.20)$$

тобто він дорівнює відносному (у розрахунку на один електрон) збільшенню кількості електронів на одиниці шляху.



Розв'язком рівняння (4.19) є функція

$$n(x) = n(0) e^{\alpha x}, \quad (4.21)$$

яка показує, що електронний струм

Рис. 4.18. Умовне зображення процесу лавинного розмноження електронів у результаті ударної іонізації

$$j_n(x) = en(x) = j_n(0) e^{\alpha x}, \quad (4.22)$$

експоненціально наростає від катода до анода й кожному електрону, який вийшов з катода, відповідає в середньому $e^{\alpha l}$ електронів, що надходять на анод. Цю величину, рівну відношенню струму в анодному колі до струму, стимульованому зовнішніми джерелами,

$$M = j_n(l) / j_n(0) = e^{\alpha l}, \quad (4.23)$$

називають *коефіцієнтом множення*. Гаданого порушення закону сталості повного струму тут не відбувається, тому що в цій моделі не враховано струм, створюваний позитивними зарядами. Крім того, ці заряди, у свою чергу, рухаючись до катода, також можуть викликати іонізацію й зустрічний лавинний процес.

У напівпровідниках лавинні процеси створюють і електрони, і дірки. Припустимо, що з боку катода в об'єм напівпровідника вводяться (інжектуються) електрони. Потрапляючи в зону провідності, вони прискорюються електричним полем. Поки їхня кінетична енергія (відстань від дна зони провідності до рівня їхньої повної енергії, зображеного стрілками на енергетичній діаграмі рис. 4.19, а) залишається меншою за ширину забороненої зони, зіткнення електронів з ґраткою мають майже пружний характер. Лише незначна частина їхньої енергії витрачається на збудження коливань ґратки (фононів). Але коли енергія електрона досягає значень, більших за ширину забороненої зони, результатом зіткнення може бути порушення

валентного зв'язку, тобто перехід електрона з валентної зони в зону провідності. Триваючи, такий процес приводить до утворення електронної лавини.

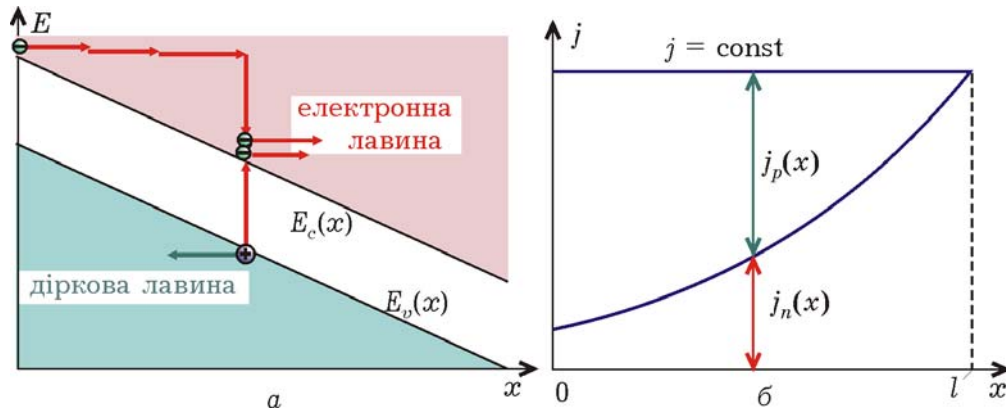


Рис. 4.19. Зміна енергії під час руху електронів у сильному електричному полі й у результаті зіткнень (а). При ударній іонізації утворюється електронно-діркова пара. Збільшення електронного й діркового струмів відбувається в різних напрямках (б), але так, що повний струм залишається постійним

Генеровані у валентній зоні дірки прискорюються електричним полем і збуджують *діркову лавину*, що розвивається у напрямку від анода до катода. Зустрічні потоки електронів і дірок створюють повний струм у колі, який у кожній точці буде однаковим й, природно, більшим за струм інжекції (рис. 4.19, б).

Утворення нової електронно-діркової пари можливе, якщо енергія електрона досягає деякого мінімального значення

$$E_{\min} = [1 + m^*/(m^* + m_p^*)]\Delta E, \quad (4.24)$$

яке називають *порогом іонізації* E_i . Це співвідношення легко впливає із законів збереження енергії й імпульсу первинного електрона і нової електронно-діркової пари. Залежно від співвідношення ефективних мас електронів і дірок $E_i = (1...2)\Delta E$.

Для створення найпростішої моделі лавинного процесу вважатимемо, що іонізувальна здатність дірок така сама, як і в електронів, тобто на одиниці шляху вони також утворять α нових електронно-діркових пар. Припустимо, що в якій-небудь точці утворилася пара рухливих носіїв. Рухаючись під дією електричного поля в різні сторони, електрон і дірка пройдуть

шлях, у сумі рівний міжелектродній відстані, а в результаті зіткнень утворюють у середньому αl нових електронно-діркових пар. Такі «вторинні» частки поведуться подібно «первинним», тобто кожна пара утворює, як і раніше, αl нових або всього $(\alpha l)^2$ «третинних» пар. Ті, у свою чергу, розмножуються в αl разів і т. д. Повну кількість пар, утворених в решті-решт єдиною первісною парою, знаходять як суму ряду

$$1 + \alpha l + (\alpha l)^2 + (\alpha l)^3 + \dots = 1/(1 - \alpha l) = M. \quad (4.25)$$

Отримана величина i є вже згадуваним коефіцієнтом множення. Коефіцієнт іонізації є сильною функцією напруженості електричного поля. Характер цієї залежності можна оцінити з таких простих міркувань. У сильному електричному полі, де електрони дрейфують зі швидкістю насичення v_n , у стаціонарному стані вони мають енергію $E \approx e\mathcal{E}v_n\tau_E$. На довжині релаксації енергії $l = v_n\tau_E \gg l_p$ вона практично повністю переходить у теплову, тому $k_B T_e \approx 2e\mathcal{E}l/4$. Імовірність того, що енергія перевищить поріг іонізації визначається больцманівським множником $\exp(-E_i/k_B T_e) = \exp(-3E_i/2e\mathcal{E}l)$, характерним і для коефіцієнта іонізації. Тому для залежності $\alpha(\mathcal{E})$ часто користуються функцією

$$\alpha(\mathcal{E}) = \alpha_0/\exp(\mathcal{E}_0/\mathcal{E})^m, \quad (4.26)$$

де, як встановлено експериментально, $m = 1$ для германію й кремнію і $m = 2$ для деяких сполук типу $A^{III}B^V$, а константи α_0 й \mathcal{E}_0 залежать від напівпровідника й типу носіїв заряду. За порядком величин $\alpha_0 \approx (10^5 \dots 10^6) \text{ см}^{-1}$, а $\mathcal{E}_0 \approx 10^3 \text{ кВ/см}$.

Останні вирази визначають вид вольт-амперної характеристики лавинного процесу:

$$j(\mathcal{E}) = j_0 M = j_0/[1 - \alpha(\mathcal{E})l]. \quad (4.27)$$

З (4.27) можна зробити висновок, що коефіцієнт множення i , відповідно, повний струм у колі зростають із зростанням напруженості поля. За деякої напруженості поля, називаною критичною, $\alpha l \rightarrow 1$ і можливе нескінченне наростання струму через діодний проміжок (рис. 4.20, а). Таке явище називають *пробоєм*. Насправді в режимі пробою струм переважно обмежується опором зовнішнього кола.

Якщо електричне поле неоднорідне уздовж зразка, що зазвичай спостерігається за більш-менш значної інтенсивності лавинних процесів, то замість αl треба підставляти $\int_0^l \alpha dx$ і

$$M = 1 / (1 - \int_0^l \alpha dx). \quad (4.28)$$

Якщо дірковий коефіцієнт іонізації $\beta = dp/pdx$ відрізняється від електронного, то умова пробою матиме такий вигляд:

$$\int_0^l \alpha \exp[-\int_0^x (\alpha - \beta) dx'] dx = 1. \quad (4.29)$$

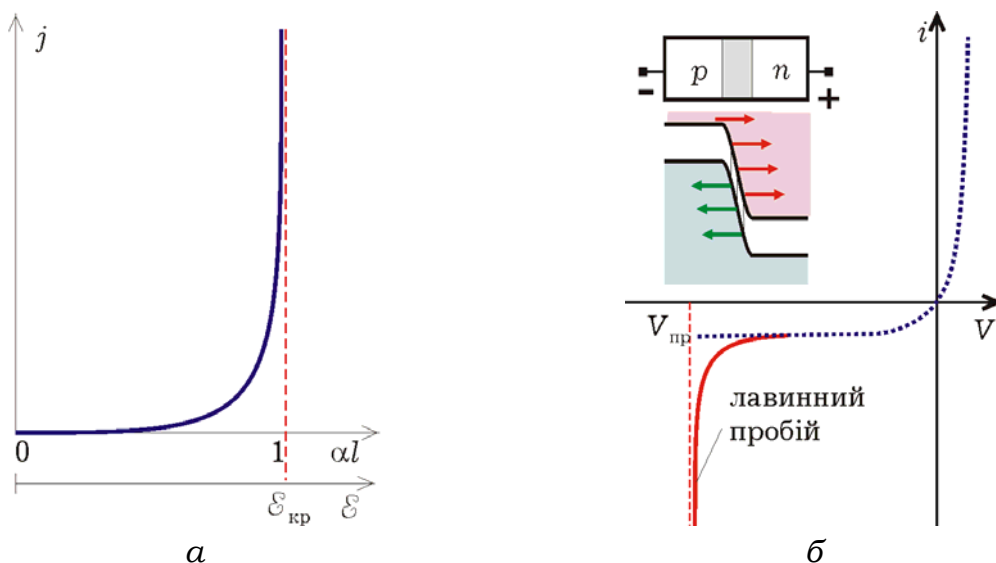


Рис. 4.20. Зростання струму в напівпровіднику зі збільшенням електричного поля викликано ростом інтенсивності лавинних процесів. За певного значення напруженості поля коефіцієнт множення зростає нескінченно й настає лавинний пробій (а). На вольт-амперній характеристиці (б) електронно-діркового переходу (штрихова лінія) лавинні процеси викликають різке зростання зворотного струму (суцільна лінія)

Лавинний пробій можна спостерігати в електронно-дірковому переході (рис. 4.20, б). При зворотному зміщенні переходу навіть при не дуже великих напругах у тонкій збідненій області можна створити сильне електричне поле, що викликає лавинні процеси. Утворення електронно-діркових пар у збідненій області еквівалентне збільшенню струму неосновних носіїв, що приводить до збільшення зворотного струму для напруги, близької до пробивної. Якщо вжити заходів проти необоротних процесів руйнування, то цей режим можна використовувати для стабілізації напруги, оскільки у великому діапазоні зміни струму через перехід напруга на ньому змінюється незначно.

Процеси лавинного розмноження використовують також для створення так званих лавинно-пролітних діодів (ЛПД), які можуть працювати як генератори на частотах порядку сотень гігагерц, а також у різних діодних і транзисторних структурах, де бажане збільшення робочого струму.

Характерною для лавинно-пролітного діода є $p^+-n-i-n^+$ -структура, яку зображено на рис. 4.21. За своєю суттю вона являє собою електронно-дірковий перехід, в якому за рахунок меншої концентрації домішки збіднена область займає n -шар. Практично вся зворотна напруга падає на ній, створюючи сильне електричне поле

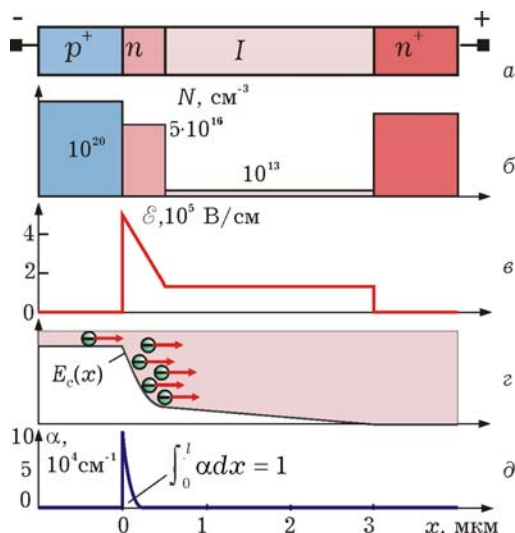


Рис. 4.21. Структура ЛПД (а), профіль легування в ньому (б), розподіл напруженості електричного поля (в) в області іонізації та області дрейфу, потенціальний рельєф у зоні провідності (г), де відбувається лавинне розмноження, та розподіл коефіцієнта іонізації (д) у цій області.

Саме у збідненій області відбувається лавинне розмноження, інтенсивність якого практично відповідає умові пробою $\int_0^l \alpha dx = 1$.

Струм, який створюється за цієї умови, забезпечує так званий баланс амплітуд, властивий генераторним пристроям..

Друга умова, баланс фаз, забезпечується у відносно довгій слабо легованій i -області, де електрони дрейфують з постійною швидкістю, що дорівнює швидкості насичення. Тривалість дрейфу має відповідати часу затримки, необхідної для того, щоб електронні згущення, які утворюються у процесі лавинного розмноження під час прискорювального півперіоду поля у режимі генерації, приходили до анодної n^+ -області під час гальмівного поля. Протифазні коливання змінних напруги і струму відповідають від'ємній динамічній провідності, що і забезпечує можливість генерації надвисокочастотних коливань.

4.2.2. Електростатична іонізація

Ефект *електростатичної іонізації* у напівпровідниках (ефект Зінера) у зовнішній прояві дуже схожий на процес лавинного розмноження: різке зростання струму через напівпровідник зі збільшенням прикладеної напруги. Однак природа цього явища зовсім інша, і пояснити його можна тільки з позицій квантової механіки.

В електричному полі спостерігається викривлення енергетичних зон. В однорідному полі це виглядає (рис. 4.22) як нахил зон під кутом, тангенс якого дорівнює силі $F = e\mathcal{E}$. При цьому певному значенню енергії електрона можуть відповідати стани й у валентній зоні, і в зоні провідності. У полі з напруженістю порядку 10^9 В/м відстань

$$\Delta x = \Delta E / e\mathcal{E} \quad (4.30)$$

між точками x_1 і x_2 , що відповідають межах дозволених зон, складає величини порядку 1 \AA , що дозволяє електронам валентної зони перейти в зону провідності за рахунок *тунельного ефекту*, тобто без енергетичного збудження. Імовірність тунелювання зростає зі зменшенням Δx , тобто зі збільшенням напруженості електричного поля, і визначають її за допомогою

відомого з квантової механіки виразу для коефіцієнта прозорості трикутного потенціального бар'єра заввишки ΔE :

$$D = D_0 \exp[-4\sqrt{2m\Delta E} / \hbar^2 \Delta x / 3]. \quad (4.31)$$

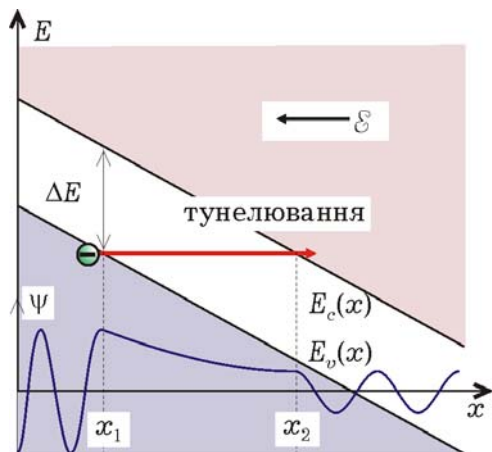


Рис. 4.22. Імовірність тунельного переходу електронів з валентної зони в зону провідності пов'язана з виникненням у зонах під дією поля станів, що відповідають одній і тій самій енергії

Збільшення напруженості електричного поля зменшує товщину бар'єра і підвищує ймовірність тунелювання електронів у зону провідності. У результаті пропорційно ймовірності зростає кількість електронно-діркових пар, що беруть участь у провідності. Це дозволяє вольт-амперну характеристику зразка напівпровідника у разі електростатичної іонізації описати виразом

$$j = j_0 \exp[-4\sqrt{2m^*} / \hbar^2 \Delta E^{3/2} / 3e\mathcal{E}]. \quad (4.32)$$

За зовнішнім виглядом вольт-амперна характеристика нагадує рис. 4.20. Однак вона не має певного асимптотичного значення $E_{кр}$, як у лавинному пробі. Тим не менше швидкість зростання струму, характерна для експоненціальної функції (4.32), настільки велика, що це явище також називають *електростатичним (тунельним) пробієм*.

Його також можна спостерігати у разі зворотного зміщення електронно-діркового переходу (рис. 4.23), коли в тонкій збідненій області виникають поля з напруженістю $10^8 \dots 10^9$ В/м.

Часто електростатичний пробій маскується лавинним, який може відбуватися у слабших полях. Тому електростатичний пробій у чистому вигляді можна спостерігати тільки на дуже тонких переходах, де лавинний процес не встигає розвиватися.

Крім того, ідентифікувати той чи інший вид пробою можна за формою вольт-амперної характеристики через зазначену відмінність.

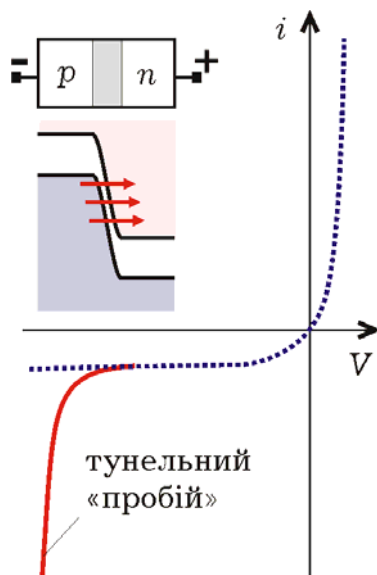


Рис. 4.23. Вольт-амперна характеристика електронно-діркового переходу (штрихова лінія) може мати ділянку різкого зростання зворотного струму (суцільна лінія), викликану тунелюванням електронів.

Електростатичний пробій в електронно-дірковому переході, як і лавинний, використовують для створення стабілізаторів напруги.

4.2.3. Іонізація у домішкових напівпровідниках

Явище електростатичної іонізації можна спостерігати і в домішкових напівпровідниках. Сферично-симетричне поле донорного атома, яке втримує «зайвий» електрон (рис. 4.24, а), взаємодіє із зовнішнім електричним полем так, що праворуч від атома сила, яка утримує електрон, зменшується. На енергетичній діаграмі (рис. 4.24, в) це зображено як зміна форми потенціального бар'єра в околі домішкового атома. У досить сильних полях (на 2...3 порядки нижчих, ніж для власних напівпровідників) потенціальний бар'єр стає настільки вузьким, що з'являється ймовірність тунелювання електрона з донорного рівня у зону провідності.

Однак цей ефект можна спостерігати тільки за дуже низьких температур, коли ще не відбувається іонізації домішки за рахунок теплового збудження. Уже при температурах в 20...30 К для мілких домішкових рівнів (у германії, кремнії) стає помітним

теплове збудження електронів з домішкових рівнів у зону провідності. За цих умов може відбуватися *термоелектронна іонізація* (ефект Френкеля), тобто в області іонізації домішки концентрація електронів у зоні провідності зростає за наявності електричного поля. Це пояснюється тим, що в електричному полі знижується енергія іонізації домішки й при певній температурі підвищується ймовірність її теплового збудження.

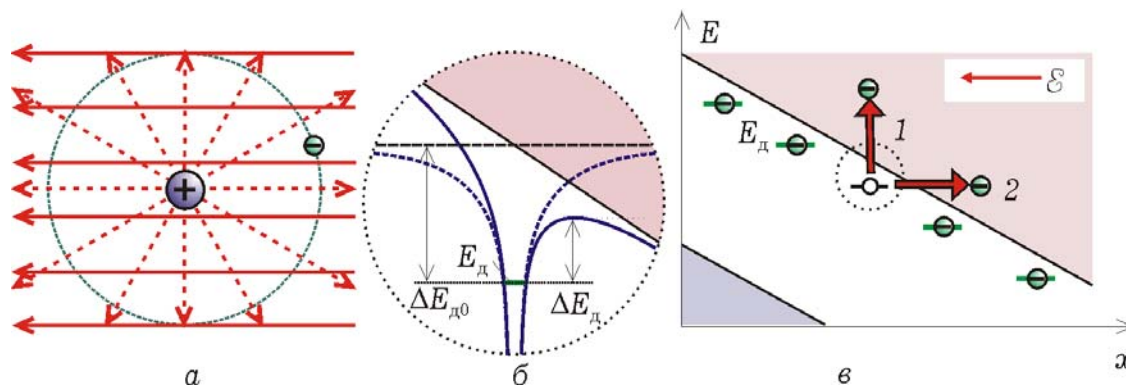


Рис. 4.24. Зовнішнє електричне поле деформує поле домішкового іона (а), зменшує енергію іонізації домішки (б) і полегшує її термічну іонізацію – 1 (в). Крім того, можливе тунелювання «зайвого» електрона домішки через деформований бар'єр – 2 (в)

Оскільки від напруженості електричного поля залежить концентрація носіїв, тобто й електропровідність, то вольт-амперна характеристика зразка відрізнятиметься від лінійної. Тому цей ефект належить до ефектів сильного поля, незважаючи на те, що по суті він є комбінованим. Роль електричного поля зводиться до зниження енергії іонізації домішки, а саме збудження відбувається за рахунок енергії теплових коливань ґратки.

4.2.4. Автоелектронна емісія

Автоелектронною називають емісію електронів з речовини за наявності біля її поверхні сильного, порядку $10^8 \dots 10^9$ В/м, електричного поля, що прискорює електрони від поверхні. Інші назви цього виду емісії – холодна, електростатична і, рідше, тунельна – повніше відображають природу цього явища. Вихід електронів під час автоелектронної емісії на вимагає зовнішнього

збудження і може відбутися з холодних катодів за наявності електростатичного поля. Остання назва підкреслює квантово-механічну природу цього явища й говорить про те, що причиною виходу електронів є *тунельний ефект*.

Зовнішнє електричне поле у вакуумі знижує потенціальну енергію електрона за межами твердого тіла, змінюючи її за законом $U_{\text{вн}} = -e\mathcal{E}x$. На енергетичній діаграмі для металу (рис. 4.25) це зображено трикутним потенціальним бар'єром, трохи скругленим за рахунок сил дзеркального відображення. За напруженості поля порядку $10^8 \dots 10^9$ В/м товщина бар'єра поблизу рівнів, що відповідають заповненим електронним станам, стає настільки малою (одиниці й десятки ангстрем), що з'являється ймовірність виходу електронів із цих станів у вакуум за рахунок тунельного ефекту.

Отже, автоелектронна емісія – суто квантовий ефект, пов'язаний із хвильовими властивостями електронів, і роль електричного поля зводиться тут до формування відповідного потенціального бар'єра, а не до «виривання» електронів, як іноді вважають.

Характерною рисою автоелектронної емісії для металів є вихід електронів з енергіями, близькими до енергії Фермі. Оскільки в холодному катоді немає більш енергійних електронів, а зі зменшенням енергії прозорість бар'єра експоненціально зменшується, розкид швидкостей емітованих електронів дуже невеликий і становить одну-дві десятих електрон-вольта.

Основний закон автоелектронної емісії називається законом Фаулера – Нордгейма і виражає залежність густини емісійного струму від напруженості електричного поля. Для металів, якщо знехтувати силами дзеркального зображення, вигляд цієї залежності за структурою подібний до закону термоелектронної емісії

$$j(\mathcal{E}) = B\mathcal{E}^2 \exp[-4\sqrt{2m / \hbar^2} \Phi^{3/2} / 3e\mathcal{E}], \quad (4.33)$$

де B – константа, що залежить від параметрів енергетичної структури металу, а експоненціальний множник аналогічний (4.32) із заміною ΔE на роботу виходу Φ . Це природно, тому що

механізм електростатичної іонізації й автоелектронної емісії однаковий.

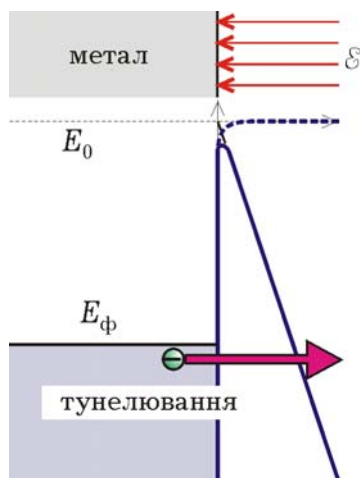


Рис. 4.25. Зміна форми потенціального бар'єра (б) біля поверхні металу під дією зовнішнього прискорювального поля (а) створює передумови для тунелювання електронів у вакуум

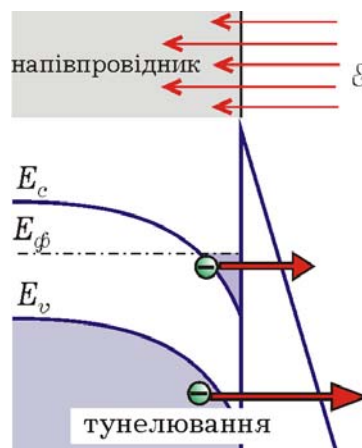


Рис. 4.26. У напівпровіднику електричне поле, проникаючи в поверхневий шар (а), підвищує там концентрацію електронів. У результаті емісія може відбуватися як із зони провідності, так і з валентної (б)

Основні особливості автоелектронної емісії з напівпровідників пов'язані з проникненням зовнішнього електричного поля у приповерхневий шар (рис. 4.26) і відповідним заповненням електронами енергетичних зон. Оскільки концентрація електронів у зоні провідності в багато разів менша, ніж у валентній, то емісія валентних електронів може виявитися істотною, незважаючи на меншу ймовірність їх тунелювання через ширший бар'єр. Поряд із цим електричне поле викривляє енергетичні зони так, що рівень Фермі може опинитися в зоні провідності, і це спричинить підвищення концентрації електронного газу в поверхневому шарі та його виродження.

Струм, що протікає через товщу катода і який під час автоелектронної емісії може досягти дуже великої густини, розігріває катод за рахунок його омичного опору. Тому в сильному електричному полі поверхневого шару можливе розігрівання електронного газу та генерація носіїв за рахунок лавинних процесів або за рахунок електростатичної іонізації. Якщо врахувати, що зовнішні поля й поверхневі стани можуть знижувати ефективну електронну спорідненість до від'ємних

значень, то стає зрозумілим, як складно врахувати спектр усіх цих явищ у теорії автоелектронної емісії з напівпровідників.

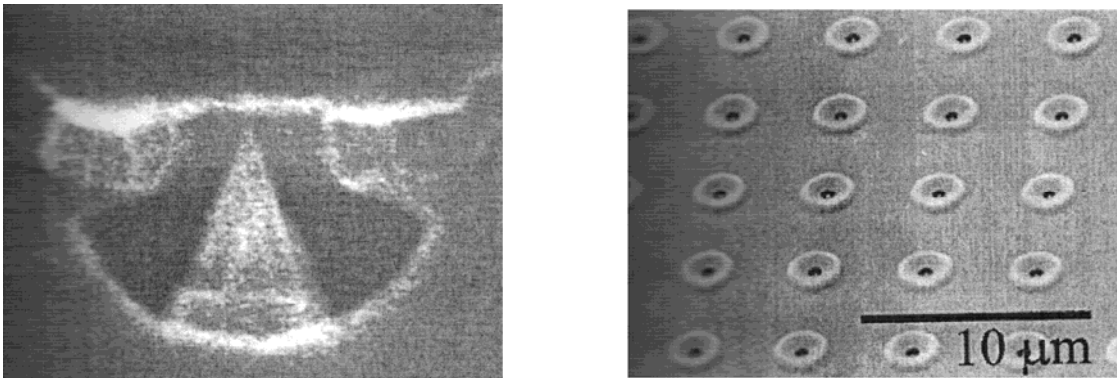


Рис. 4.27. Приклад тріодної структури з вістряним катодом

Автоелектронну емісію використовують для створення холодних катодів (рис. 4.27), які забезпечують малий розкид швидкостей електронів і дають можливість формувати дуже тонкі електронні потоки з малими шумами. Наявні холодні катоди мають ряд недоліків, зокрема: нестабільність струму емісії, пов'язану з впливом адсорбованих на поверхні атомів, сильним бомбардуванням катода іонами залишкових газів і зміною в процесі роботи геометрії поверхні катода, який найчастіше виготовляють у вигляді вістря, й тому він сильно розігрівається за великої густини струму. Тому дослідження щодо усунення зазначених недоліків тривають і досі.

Контрольні запитання та вправи

4.1. Які електричні поля прийнято називати сильними? Перерахуйте ефекти, що спостерігаються у сильних полях, розділивши їх на дві характерні групи (які?).

4.2. З якою особливістю зіткнень і чому пов'язаний ефект розігрівання електронного газу? Які види зіткнень найбільшою мірою сприяють розігріванню? Охарактеризуйте різні види розсіювання з огляду на їх пружність.

4.3. З яким видом розсіювання пов'язане насичення дрейфової швидкості у ковалентних кристалах (Ge, Si)? У полях якої напруженості воно спостерігається? Який порядок швидкості насичення?

4.4. У чому полягає суттєва відмінність механізмів насичення дрейфової швидкості в арсеніді галію від таких у ковалентних напівпровідниках? Наведіть типовий вигляд поле-швидкісної (вольт-амперної) характеристики для GaAs і поясніть зв'язок міждолинних переходів з наявністю ділянки негативної диференціальної рухливості (провідності).

4.5. Поясніть, чому в динамічному режимі швидкість дрейфу може бути отримана вищою, ніж в стаціонарному (ефект «сплеску»).

4.6. Що таке ефект Ганна і яка особливість закону дисперсії необхідна для його прояву? Поясніть, як утворюється домен сильного поля і від чого залежить частота його повторення.

4.7. У чому полягає макроскопічний прояв ефекту Зінера і яке його квантово-механічне пояснення? За яких умов він спостерігається і які властивості електрона відповідальні за його прояв? Яка особливість електростатичного «пробою» зумовлює сферу його застосування?

4.8. Поясніть причини і макроскопічні прояви термоелектронної іонізації (ефекту Френкеля), звернувши увагу на роль електричного поля і термічного збудження.

4.9. Що таке автоелектронна емісія, які її необхідні умови, внутрішній механізм і основний закон? Які її особливості зумовлюють області використання вістряних катодів? (Корисно прослідкувати спільні риси і відмінність автоелектронної емісії і ефекту Зінера).

4.10. Що таке лавинні процеси? Наведіть найпростішу модель лавинного розмноження. У чому її недолік при розгляді носіїв лише одного знаку?

4.11. Опишіть стаціонарну модель (фізичну і математичну) утворення електронно-діркових лавин у напівпровідниках. Сформулюйте умову лавинного пробою, звернувши увагу на його відмінність від електростатичного «пробою».

4.12. Опишіть структуру, принцип дії і призначення простих конструкцій лавинно-пролітних діодів (ЛПД).

5. РЕКОМБІНАЦІЯ НОСІЇВ ЗАРЯДУ

5.1. ЗАГАЛЬНІ ЗАКОНОМІРНОСТІ РЕКОМБІНАЦІЇ

Рекомбінація – це процес, обернений до генерації, за якого відбувається зникнення рухливих носіїв заряду. Питання про конкретні механізми рекомбінації виникає зазвичай, коли в речовині, крім рівноважних, є надлишкові носії заряду, що з'являються за будь-якого зовнішнього впливу. Після припинення цього впливу відбувається повернення у рівноважний стан, швидкість якого визначається швидкістю рекомбінації. Одним з основних для дослідження рекомбінаційних процесів є твердження про те, що *швидкість рекомбінації* R , тобто кількість носіїв, що зникають в одиниці об'єму за одиницю часу, пропорційна добутку концентрацій часток, що беруть участь у рекомбінації. Для електронно-діркової рекомбінації у напівпровіднику

$$R = \alpha_r n p, \quad (5.1)$$

де α_r – коефіцієнт пропорційності, який називають *коефіцієнтом рекомбінації*, $\text{м}^3/\text{с}$.

Очевидно, що в рівноважному стані швидкості генерації G_0 і рекомбінації R_0 рівні між собою:

$$R_0 = \alpha_r n_0 p_0 = G_0,$$

а зміна надлишкової концентрації Δn за одиницю часу дорівнює перевищенню швидкості рекомбінації над швидкістю рівноважної генерації:

$$d(\Delta n)/dt = -(R - G_0) = -\alpha_r (np - n_0 p_0). \quad (5.2)$$

Аналогічний вираз можна отримати і для дірок. Оскільки нерівноважна концентрація дорівнює сумі концентрацій рівноважних і надлишкових носіїв:

$$n = n_0 + \Delta n; \quad p = p_0 + \Delta p,$$

то

$$d(\Delta n)/dt = d(\Delta p)/dt = -\alpha_r(n_0\Delta p + p_0\Delta n + \Delta n\Delta p). \quad (5.3)$$

Це рівняння дуже просто розв'язати, якщо надлишкові концентрації електронів і дірок однакові, для двох крайніх, але характерних випадків.

У випадку *низького рівня збудження*, коли надлишкова концентрація набагато менша за рівноважну концентрацію основних носіїв, але може бути більшою за концентрацію неосновних, $\Delta n = \Delta p \ll (n_0 + p_0)$ і третій доданок в (5.3) знехтувально малий. Тому

$$d(\Delta n)/dt = d(\Delta p)/dt = -\alpha_r(n_0 + p_0)\Delta n, \quad (5.4)$$

і таку рекомбінацію називають *лінійною*. Розв'язком (5.4) буде експоненціальна функція (рис. 5.1, а):

$$\Delta n(t) = \Delta n(0) \exp(-t/\tau), \quad (5.5)$$

швидкість спадання якої визначається константою

$$\tau = 1/\alpha_r (n_0 + p_0), \quad (5.6)$$

яку називають *часом життя нерівноважних носіїв заряду*.

У другому випадку, коли рівень збудження високий, $\Delta n = \Delta p \gg (n_0 + p_0)$ і основну вагу має третій доданок в (5.3):

$$d(\Delta n)/dt = d(\Delta p)/dt = -\alpha_r\Delta n^2. \quad (5.7)$$

Таку рекомбінацію, за якої швидкість спадання надлишкової концентрації пропорційна її квадрату, називають *квадратичною* й описують гіперболічною функцією (рис. 5.1, б)

$$\Delta n(t) = \Delta n(0)/[1 + \Delta n(0)\alpha_r t]. \quad (5.8)$$

Для цього процесу природної характерної константи не виникає, але часто за аналогією з лінійною рекомбінацією використовують дещо штучне поняття *миттєвого часу життя*

$$\tau = 1/\alpha_r \Delta n, \quad (5.9)$$

яке випливає з формального порівняння (5.4) та (5.7) і залежить від значення Δn у певний момент часу.

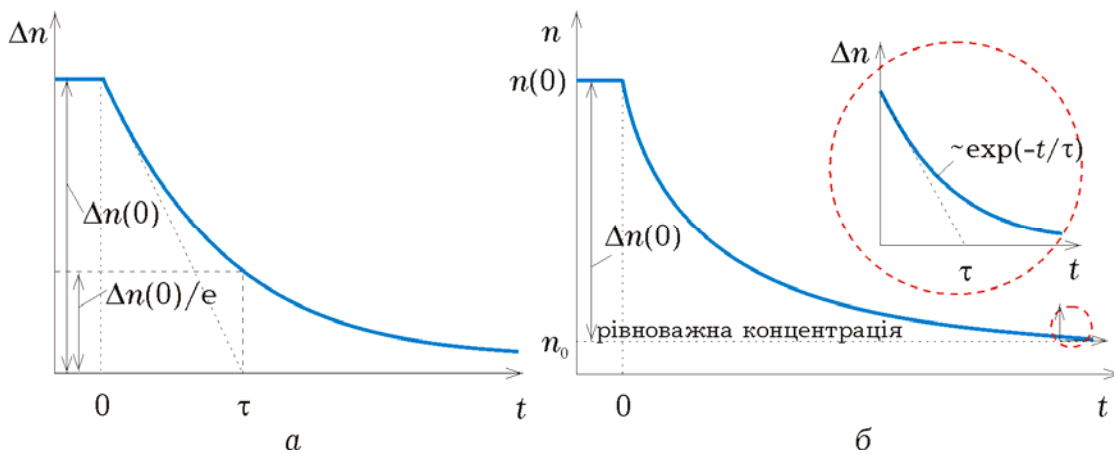


Рис. 5.1. Релаксація нерівноважних носіїв із вимиканням джерела збудження: *а* – за малої концентрації надлишкових носіїв (лінійна рекомбінація) вона спадає експоненціально; *б* – за великих рівнів надлишкових носіїв (квадратична рекомбінація) спад відбувається за гіперболічним законом, що переходить в експоненціальний із наближенням до рівноважного стану

Для газорозрядної плазми поняття надлишкових носіїв заряду зазвичай не використовують, але й у ній відбуваються процеси лінійної і квадратичної рекомбінації.

Оскільки під час рекомбінації рухливих носіїв заряду вивільнюється надлишок енергії, то основна ознака, за якою класифікують механізми рекомбінації, – це спосіб виділення цієї енергії. У зв'язку із цим розрізняють *випромінювальну* рекомбінацію, за якої відбувається вихід кванта електромагнітного випромінювання, і *безвипромінювальну*, коли вивільнювана енергія передається іншим часткам. Одна з важливих особливостей випромінювальної рекомбінації пов'язана з тим, що імпульс

випромінюваного кванта знехтувально малий, тому сумарний імпульс часток, що рекомбінують, залишається без зміни.

5.2. ЕЛЕКТРОННО-ДІРКОВА РЕКОМБІНАЦІЯ У НАПІВПРОВІДНИКАХ

За видами електронних переходів на енергетичній діаграмі (рис. 5.2) рекомбінацію електронів і дірок у напівпровіднику можна поділити на два види: міжзонну рекомбінацію (переходи 1 – 3) і рекомбінацію через локальні рівні (переходи 4 – 5).

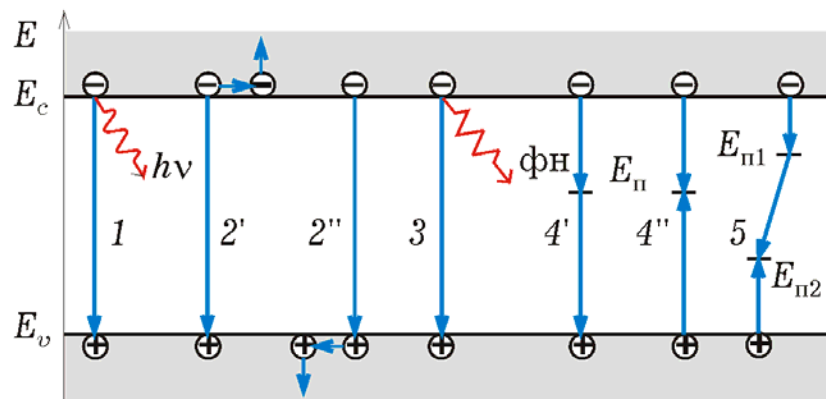


Рис. 5.2. Енергетичні переходи у разі міжзонної рекомбінації (1 – випромінювальної, 2' і 2'' – ударної, 3 – фононної) і рекомбінації через локальні рівні (4', 4'' і 5)

Міжзонна електронно-діркова рекомбінація означає перехід електрона із зони провідності у вільний стан валентної зони. Природно, що під час такого переходу виділяється енергія, рівна різниці початкового й кінцевого стану, а мінімальне її значення дорівнює ширині забороненої зони. Ця енергія може випромінюватися у вигляді електромагнітного кванта $h\nu \geq \Delta E$ (випромінювальна рекомбінація) або передаватися іншим часткам, наприклад електронам або діркам (ударна або Оже-рекомбінація), а також фононам кристалічної ґратки (фононна рекомбінація).

Особливість міжзонної випромінювальної рекомбінації легко зрозуміти, аналізуючи (5.6). У власних напівпровідниках $n_0 = p_0 = n_i$, тому $\tau_i = 1/2\alpha_r n_i$. Як впливає з (3.55), $n_i \sim \exp(-\Delta E/2k_B T)$, а отже, із

зменшенням ширини забороненої зони зменшується час життя нерівноважних носіїв заряду або збільшується ймовірність їхньої рекомбінації. У домішкових (наприклад, донорних) напівпровідниках з досить великим умістом домішки ($n_0 \gg n_i \gg p_0$) час життя нерівноважних електронів і дірок визначається концентрацією основних носіїв $\tau_n = 1/\alpha_r n_0$ й істотно менший τ_i .

Друга особливість міжзонної випромінювальної рекомбінації випливає із закону збереження квазіімпульсу або пропорційного йому хвильового вектора. Оскільки імпульс випромінюваного кванта знехтувально малий, то квазіімпульс електрона під час переходу у валентну зону практично не змінюється. Такі переходи, називані прямими, найімовірніші в прямозонних напівпровідниках, коли екстремуми дисперсійних кривих відповідають центру зони Бріллюена.

Розглянуті особливості пояснюють, чому випромінювальна рекомбінація істотна лише в прямозонних напівпровідниках із шириною забороненої зони $\Delta E < 0,3 \dots 0,5$ еВ, а швидкість рекомбінації зростає з температурою.

У разі ударної міжзонної рекомбінації (переходи $2'$ і $2''$, рис. 5.2) у кожному акті бере участь рекомбінувальна пара – електрон і дірка, а також третя частка – електрон ($2'$) або дірка ($2''$), якій передається вивільнювана енергія і яка в результаті переходить у збуджений стан. Імовірність потрійного зіткнення пропорційна концентраціям часток, що беруть участь у рекомбінації, тобто величині $n^2 p$ у донорних напівпровідниках, де третьою часткою зазвичай є електрон, і, відповідно, величині $n p^2$ в акцепторних напівпровідниках, тому що й тут більш ймовірна зустріч рекомбінувальної пари з основним носієм – діркою. Оскільки ймовірність зіткнення трьох часток у загальному випадку менша, ніж двох, то швидкість міжзонної ударної рекомбінації має бути меншою за швидкість випромінювальної, але зростання її з температурою відбувається швидше, пропорційно третій степені концентрації. Електронні переходи під час ударної рекомбінації необов'язково є прямими, тому що третя частка може прийняти практично будь-який імпульс. Ці властивості приводять до того, що в деяких напівпровідниках ударна рекомбінація може переважати над випромінювальною.

Специфічним різновидом ударної рекомбінації є *плазмова* рекомбінація, за якої енергія, що виділяється, передається всьому

колективу рухливих носіїв і витрачається на утворення *плазмонів*. Цей вид рекомбінації виявляється лише за великої концентрації носіїв, наприклад, у вироджених напівпровідниках.

Під час *фононної* міжзонної рекомбінації енергія електрона, що повертається у валентну зону, передається кристалічній ґратці й викликає збудження фононів. Оскільки енергія фононів ($\ll 10^{-2}$ eV) у багато разів менша, ніж ширина забороненої зони, то в кожному акті рекомбінації має брати участь велика кількість фононів. Відомо, що багаточасткові ефекти малоімовірні, тому фононну рекомбінацію можна помітити лише у вузькозонних напівпровідниках.

Отже, усі описані механізми міжзонної рекомбінації мають дуже малу швидкість у напівпровідниках із широкою забороненою зоною. Але експерименти показують, що спостережуваний час життя у таких напівпровідниках у багато разів менший за розрахунковий для кожного з видів міжзонної рекомбінації. Це пояснюється тим, що в широкозонних напівпровідниках механізм рекомбінації інший: вона відбувається за участю неминуче існуючих, а іноді створюваних спеціально, різного роду дефектів кристалічної решітки. Створюючи локальні енергетичні рівні в забороненій зоні, ці дефекти, називані *центрами рекомбінації*, забезпечують можливість дво- або навіть триступінчастого переходу електрона у валентну зону (переходи 4 й 5, рис. 5.2). Переходи 4' й 4'' означають, що спочатку рекомбінувальний електрон захоплюється рекомбінаційним центром й енергетично переходить із зони провідності на локальний енергетичний рівень E_{π} , створюваний цим центром. Наступний етап – це перехід 4' електрона з рівня E_{π} у валентну зону й рекомбінація його з однією з наявних там дірок.

Те саме, але з використанням поняття зв'язаної дірки означає перехід 4'' – дірка з валентної зони захоплюється рекомбінаційним центром, на якому й відбувається рекомбінація електронно-діркової пари. Аналогічно можна описати ступінчастий процес рекомбінації через два локальні проміжні стани $E_{\pi 1}$ й $E_{\pi 2}$ (перехід 5). Багатоступінчаста рекомбінація набагато імовірніша, ніж міжзонна, оскільки енергія рекомбінувального електрона виділяється у два або три етапи, а частину цієї енергії може приймати рекомбінаційний центр.

Для зменшення часу життя нерівноважних носіїв ефективні центри рекомбінації часто створюються спеціально введенням у напівпровідник домішок, які утворюють у забороненій зоні *глибокі*

локальні рівні. Річ у тім, що роль домішкових центрів може бути неоднаковою залежно від положення відповідних їм рівнів у забороненій зоні. Якщо енергетичний рівень розташований близько до межі дозволеної зони, то він відіграє роль *рівня прилипання*, тому що електрон або дірка, захоплені таким центром, легко можуть повернутися у відповідну зону в результаті теплового збудження. Типовим прикладом таких центрів є мілкі донорні (аналогічно й акцепторні) стани, які, незважаючи на їхню близькість до зони провідності й більшу ймовірність переходу в них електронів, виявляються в результаті теплового збудження практично повністю вільними навіть за температури, істотно менших за кімнатну.

Глибокі енергетичні рівні, розташовані ближче до середини забороненої зони, служать ефективними центрами рекомбінації, оскільки ймовірність зворотного теплового збудження в зону провідності захопленого електрона в цьому випадку менша. Чим довше електрон утримується у локальному стані E_n , тим більша ймовірність його переходу у валентну зону, або, що те саме, зустрічі з діркою в стані E_p . От чому центри рекомбінації часто називають *пастками*.

Теорію рекомбінації через локальний рівень розробили Холл, Шоклі й Рід, суть її викладено в численній навчальній літературі з фізики напівпровідників. Не зупиняючись на деталях цієї теорії, наведемо лише її найважливіший практичний висновок, придатний для домішкових напівпровідників з концентрацією надлишкових носіїв $\Delta n = \Delta p$, великою порівняно з концентрацією пасток N_p : *час життя нерівноважних носіїв заряду визначається часом життя неосновних носіїв і обернено пропорційний концентрації пасток*. Для донорних напівпровідників

$$\tau \approx \tau_n = 1/\alpha_{rp}N_p, \quad (5.10a)$$

для акцепторних

$$\tau \approx \tau_p = 1/\alpha_{rn}N_p, \quad (5.10b)$$

де τ_n , τ_p – часи життя електронів і дірок щодо захоплення їх пастками; α_{rn} , α_{rp} – відповідні коефіцієнти рекомбінації (захоплення) електронів і дірок на цих пастках.

Зміст цього висновку легко зрозуміти на прикладі сильно легованого напівпровідника. У донорному напівпровіднику рівень

Фермі (імовірність його заповнення електронами дорівнює $1/2$) розташований близько до зони провідності й, відповідно, вище рівня пасток $E_{\text{п}}$. Оскільки функція Фермі – Дірака змінюється в інтервалі кількох $k_{\text{Б}}T$ поблизу $E_{\text{ф}}$ майже від 0 до 1, то основна маса домішкових центрів буде заповнена електронами, і швидкість рекомбінації визначатиметься лише швидкістю переходу їх з рівня $E_{\text{п}}$ у валентну зону. Інакше кажучи, швидкість рекомбінації залежить від швидкості переходу дірок (неосновних носіїв) з валентної зони на рівень пасток. В акцепторному напівпровіднику майже всі пастки вільні (заповнені дірками) і час життя нерівноважних носіїв складається переважно із часу, необхідного електронам зони провідності для переходу на рівні пасток, тому що подальший перехід їх у валентну зону відбувається практично миттєво через велику кількість дірок, що містяться там.

Наведене пояснення є не дуже строгим, оскільки замість рівня Фермі треба розглядати електронний і дірковий квазірівні, але зміст цього висновку теорії Холла, Шоклі й Ріда залишається незмінним: швидкість двоступінчастого процесу рекомбінації визначається швидкістю захоплення пастками неосновних носіїв, тобто швидкістю менш імовірного (через малу концентрацію останніх) процесу. Таке уявлення про рекомбінацію електронів і дірок на пастках дозволяє формально розглядати акт рекомбінації як деяке зіткнення їх з пасткою і характеризувати його *ефективним перерізом захоплення* $S_{\text{п}}$, а час життя розглядати як середній час пробігу неосновних носіїв між такими зіткненнями:

$$\tau_p = 1 / S_{\text{п}} v_p N_{\text{п}}; \quad (5.11a)$$

$$\tau_n = 1 / S_{\text{п}} v_n N_{\text{п}}, \quad (5.11б)$$

де v_n , v_p – середні швидкості неосновних носіїв відносно центрів рекомбінації.

Переріз захоплення – важливий і широко використовуваний параметр процесу рекомбінації. Він може змінюватися в межах від 10^{-16} до 10^{-26} м² залежно від виду напівпровідника і домішки.

За способом передачі вивільнюваної енергії рекомбінація через локальні рівні, як і міжзонна, може бути випромінювальною і безвипромінювальною. Частина енергії може приймати центр рекомбінації і передавати її потім кристалічній ґратці.

Окремий вид – *поверхнева рекомбінація*, що може відбуватися за участю тампівських поверхневих станів, адсорбованих атомів або інших дефектів на поверхні. Залежно від їхнього впливу на швидкість рекомбінації розрізняють *швидкі* й *повільні* поверхневі стани. Час обміну електронами між об'ємними і швидкими поверхневими станами становить 10^{-6} с і менше. Повільні стани локалізовані не безпосередньо на самій поверхні напівпровідника, а на поверхні оксидної або іншої діелектричної плівки, що практично завжди покриває поверхню напівпровідника. Зрозуміло, що через малу провідність цієї плівки час переходу електрона між об'ємними й поверхневими станами може бути дуже великим (порядку 10^{-2} с і більше). Тому на час життя нерівноважних носіїв повільні стани безпосередньо не впливають.

У тих напівпровідниках, де концентрація рекомбінаційних центрів в об'ємі мала, швидкість поверхневої рекомбінації через швидкі стани може бути набагато більшою за швидкість об'ємної. Це особливо істотно для тонких зразків напівпровідника, де рекомбінація відбуватиметься здебільшого через поверхню. Переважання рекомбінації на поверхні викликає спрямований потік нерівноважних носіїв з об'єму. Швидкість цього потоку дорівнює відношенню кількості r_s рекомбінуючих за 1 с на одиниці поверхні електронно-діркових пар до їхньої надлишкової концентрації Δn_s у поверхневому шарі:

$$v_s = r_s / \Delta n_s. \quad (5.12)$$

Величина v_s має розмірність метр за секунду і називається *швидкістю поверхневої рекомбінації*. Слід зазначити, що v_s – це не швидкість окремих носіїв, а середня швидкість дифузійного потоку нерівноважних носіїв, спрямованого до поверхні й викликаного відмінністю в їхній концентрації на поверхні та в об'ємі.

На швидкість поверхневої рекомбінації істотно впливає кількість не тільки рекомбінаційних центрів, а й інших станів (у тому числі й повільних), які в результаті електронного обміну з об'ємом можуть заряджатися позитивно або негативно. Незалежно від знака цього заряду поле, що виникає, буде зменшувати потік до поверхні носіїв того чи іншого типу, що приведе до зниження швидкості рекомбінації. Природно, що швидкість поверхневої рекомбінації буде найбільшою, коли потенціал поверхні дорівнюватиме нулю.

5.3. РЕКОМБІНАЦІЯ ЕЛЕКТРОНІВ ТА ІОНІВ У ГАЗОВІЙ ПЛАЗМІ

Для плазми газового розряду найтипівішою є рекомбінація електрона з позитивним іоном. Як правило, вона відбувається за схемою



яка означає або безпосереднє перетворення іона в нейтральний атом, або ступеневе (через проміжні рівні збудження). Швидкість цього процесу пропорційна ймовірності зустрічі електронів й іонів, тобто добутку їхньої концентрації:

$$dn/dt = -\alpha_r n n_i. \quad (5.14)$$

Зазвичай концентрації електронів та іонів у нейтральній плазмі рівні між собою, і така рекомбінація є квадратичним процесом.

Надлишкова енергія ΔE , що виділяється під час рекомбінації і дорівнює сумі енергії іонізації атома й кінетичної енергії електрона, може випромінюватися як квант електромагнітної енергії. Тому через розкид електронних швидкостей спектр цього випромінювання має бути безперервним з мінімальною граничною частотою, що відповідає енергії іонізації. Для ступінчастих процесів спектр має складатися з декількох смуг.

За низьких і середніх тисків газу коефіцієнт об'ємної випромінювальної рекомбінації α_r дуже малий. Це пояснюється не тільки малою ймовірністю зустрічі електронів з іонами через низькі значення їхньої концентрації, але й короткочасністю взаємодії. Час взаємодії електрона й іона має відповідати часу випромінювання кванта, що становить $\sim 10^{-8}$ с. Але для швидкості електронів, характерної для плазми низького й середнього тиску, час їхнього прольоту в електричному полі іона зазвичай на кілька порядків менший. Тому випромінювальна рекомбінація в об'ємі можлива практично в газорозрядній плазмі лише високого тиску в результаті взаємодії іонів з повільними електронами.

Інший можливий результат процесу (5.13) – передача енергії ΔE , що виділяється, третій частці. Такими частками, зазвичай є нейтральні атоми, концентрація яких на два-три порядки вища за

концентрацію електронів та іонів. Але й цей вид рекомбінації відіграє істотну роль лише в плазмі високого тиску, оскільки ймовірність потрібного зіткнення за низьких густин газу мала.

У газових розрядах, що відбуваються за низького і середнього тиску, рекомбінації в об'ємі практично немає й основною є рекомбінація на стінках, що оточують плазму. Стінки відіграють роль третьої частки, якій можуть бути передані будь-яка енергія й будь-який імпульс рекомбінуючих часток. Потрапляння зарядів на стінки відбувається переважно в результаті хаотичного руху. Однак електрони мають більші, ніж іони, швидкість і довжину вільного пробігу, тому потрапляння їх на стінки більш імовірне. У результаті стінка заряджається негативно відносно плазми, а електричне поле, що виникає при цьому, сприяє витягуванню іонів із пристіночного шару плазми. Тому швидкість рекомбінації залежить лише від швидкості приходу іонів на стінки, і така рекомбінація є лінійною.

Ефективною може бути взаємна рекомбінація додатного й від'ємного іонів завдяки малим швидкостям і, відповідно, великому часу взаємодії. Хоча для роботи більшості газорозрядних приладів від'ємні іони небажані, іноді для збільшення швидкості рекомбінації у розрядний проміжок вводять домішку газу, атоми якого мають велику електронну спорідненість й можуть утворювати від'ємні іони. Швидкість рекомбінації визначається здебільшого швидкістю захоплення електрона домішковими атомами. Цей спосіб використовують, коли потрібно забезпечити швидку деіонізацію розрядного проміжку після вимикання живлячої напруги.

Контрольні запитання та вправи

5.1. Порівняйте динаміку рекомбінаційних процесів за низького і високого рівнях збудження нерівноважних носіїв. Якими параметрами характеризується швидкість цих процесів?

5.2. Опишіть основні види рекомбінації у напівпровідниках і коротко охарактеризуйте вплив основних факторів на відповідні часи життя.

5.3. Які основні механізми і особливості електронно-іонної рекомбінації у газорозрядній плазмі?

6. ОПТИЧНА ГЕНЕРАЦІЯ

Під *оптичною генерацією* розуміють утворення рухливих носіїв заряду всередині речовини або емісію електронів з речовини у вакуум під час освітлення поверхні. До електромагнітного випромінювання оптичного діапазону зазвичай відносять видиме світло з довжинами хвиль 0,38...0,76 мкм, а також інфрачервоний та ультрафіолетовий діапазони, що примикають до нього. Енергія квантів цього діапазону

$$E \text{ (eV)} = 1,24/\lambda \text{ (мкм)} \quad (6.1)$$

становить величини порядку частин та одиниць електрон-вольтів і відповідає характерним параметрам енергетичної структури твердих тіл. Тому електрони, що поглинули такі кванти, здатні порушувати валентні зв'язки й викликати зміну провідності речовини (*внутрішній* або *фоторезистивний ефект*), а також, виконуючи роботу виходу, емітувати з речовини (*зовнішній фотоефект* або *фотоемісія*). До фотоелектричних відносять також ефекти при освітленні контакту двох речовин (фотогальванічні ефекти), які буде розглянуто далі.

6.1. ОСОБЛИВОСТІ ПОГЛИНАННЯ СВІТЛА

Макроскопічна модель поглинання. Нехай поверхня речовини освітлюється монохроматичним світлом частоти ν з інтенсивністю I_ν . Частина енергії, рівна коефіцієнту відбивання R , відбивається від поверхні. Енергія, пропорційна $(1 - R)I_\nu$, проходить у речовину й поглинається. Кількість квантів світла, що проходять через одиницю поверхні за одиницю часу, дорівнює відношенню інтенсивності до енергії кванта

$$n_\phi(0) = (1 - R)I_\nu/h\nu. \quad (6.2)$$

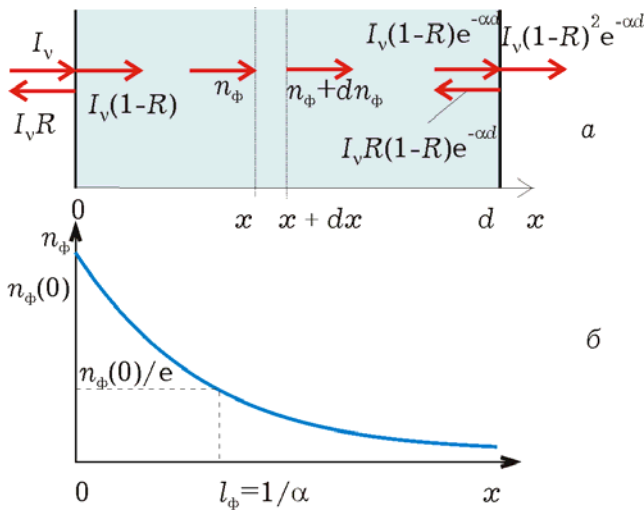


Рис. 6.1. До встановлення закону поглинання квантів світла в шарі речовини (а) і залежність кількості квантів від відстані (б)

Для встановлення закону поглинання розглянемо шар речовини завтовшки dx (рис. 6.1). Зменшення кількості квантів у цьому шарі через поглинання пропорційне кількості вхідних квантів і товщині шару:

$$dn_{\phi}(x) = -\alpha n_{\phi}(x)dx. \quad (6.3)$$

Знак «-» указує на поглинання, а коефіцієнт пропорційності α , називаний *коефіцієнтом поглинання*,

показує, яка частина квантів поглинається на одиниці довжини. Розв'язок диференціального рівняння (6.3)

$$n_{\phi}(x) = n_{\phi}(0)e^{-\alpha x} = (1 - R)I_v e^{-\alpha x}/h\nu \quad (6.4)$$

називають *законом Бугера – Ламберта*. Він показує, що інтенсивність світла всередині речовини спадає за експоненціальним законом (рис. 6.1). Швидкість спадання залежить від коефіцієнта поглинання. На відстані $l_{\phi} = 1/\alpha$, називаній *середньою глибиною поглинання* («довжиною вільного пробігу») квантів, інтенсивність світла спадає в $e \approx 2,7$ разів.

Якщо шар має скінченну товщину d , то до другої межі доходить світло з інтенсивністю, рівною $(1 - R)I_v e^{-\alpha d}$, від якої відбивається $R(1 - R)I_v e^{-\alpha d}$, а частина $(1 - R)^2 I_v e^{-\alpha d}$ проходить. Звідси коефіцієнт проходження через шар завтовшки d

$$T = I_v(d)/I_v(0) = (1 - R)^2 e^{-\alpha d}. \quad (6.5)$$

Цей вираз справедливий, якщо поглинання в шарі велике, тобто $e^{\alpha d} \gg 1$. Інакше необхідно враховувати багаторазове

відбивання світла від меж шару й відповідний коефіцієнт проходження

$$T = [(1 - R)^2 e^{-ad}] / (1 - R^2 e^{-2ad}). \quad (6.6)$$

Залежність коефіцієнта поглинання α від частоти називають *спектральною характеристикою поглинання*. Вигляд такої характеристики, наведений для Ge, Si й GaAs на рис. 6.2, залежить від особливостей зонної структури речовин і механізмів поглинання світла в них.

Зміна енергії та імпульсу електрона під час поглинання кванта.

Одну з найважливіших особливостей механізмів поглинання світла можна зрозуміти, проаналізувавши закони збереження енергії й імпульсу. Позначивши

індексом 1 – початковий, а індексом 2 – кінцевий стан електрона при поглинанні кванта світла, можна записати очевидне на перший погляд співвідношення для енергії

електрона: $E_2 - E_1 = h\nu$. Для вільного електрона збільшення кінетичної енергії на енергію кванта $h\nu$, наприклад близьку до 1 еВ, має привести до помітного збільшення швидкості й імпульсу. Але за законом збереження імпульсу $p_2 - p_1 = h/\lambda$, і поглинання кванта не повинне помітно змінити швидкість, тому що імпульс кванта h/λ дуже малий. Ця суперечність означає, що вільний електрон не може поглинути квант світла, і в акті взаємодії електронів із квантами обов'язково бере участь кристалічна гратка. Тому в рівняннях балансу енергії й імпульсу потрібно враховувати участь квантів теплових коливань гратки – фононів:

$$E_2 - E_1 = h\nu \pm E_{\text{фн}} \approx h\nu, \quad (6.7a)$$

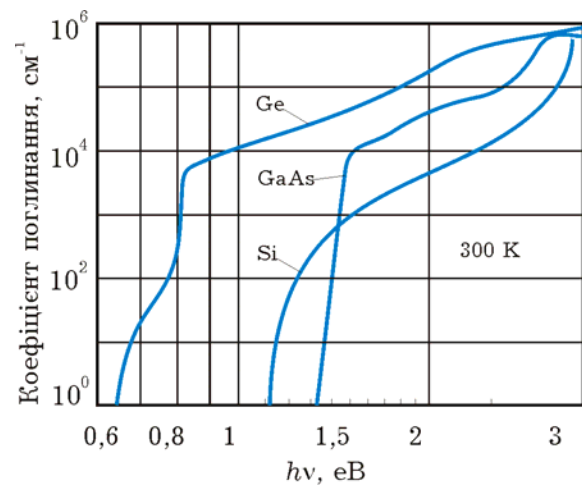


Рис. 6.2. Спектральні характеристики поглинання для основних напівпровідникових матеріалів: германію, кремнію й арсеніду галію

$$p_2 - p_1 = h/\lambda \pm \hbar k_{\text{фн}} \approx \pm \hbar k_{\text{фн}}. \quad (6.76)$$

Знаки « \pm » тут відповідають утворенню (збудженню) фонона і його поглинанню. У першому рівнянні можна знехтувати енергією фононів (соті й тисячні частки електрон-вольта), а в другому – імпульсом кванта світла, тому що імпульс фонона може досягати значень $\hbar k_{\text{фн}} \leq h/a$, які набагато більші за імпульс кванта світла і можуть змінювати хвильовий вектор електрона практично в межах усієї зони Бріллюена.

Участь фононів у поглинанні підтверджується експериментальною спектральною характеристикою поглинання для германію (див. рис. 6.2). Поглинання починається з частоти $\nu_0 = \Delta E/h$, кванти якої здатні перевести електрони з валентної зони лише в ті стани дна зони провідності, які відповідають абсолютному мінімуму дисперсійної кривої (L-долині), що лежить на краю зони Бріллюена. Такі переходи, називані *непрямими*, супроводжуються істотною зміною хвильового вектора електрона, що можливе лише за участі фононів. І тільки коли енергія квантів досягає значень, рівних ширині забороненої зони в центрі зони Бріллюена, починаються прямі переходи, що викликають новий підйом на спектральній характеристиці.

Ці два прості приклади не вичерпують всіх можливих варіантів участі фононів у процесі поглинання світла електронами. Зазначимо лише, що на експериментальних кривих спектрів поглинання участь фононів проявляється у згладжуванні границі поглинання.

Види поглинання. Існує багато причин, що викликають поглинання світла. Тут розглянемо лише ті, які приводять до утворення вільних зарядів. Серед них насамперед слід назвати *власне* або *фундаментальне* поглинання, за якого енергія поглиненого кванта витрачається на розрив валентного зв'язку. У результаті утворюється вільний електрон і дірка в системі ковалентних зв'язків, які розглядають як пару рухливих носіїв.

На енергетичній діаграмі (рис. 6.3, а) цей процес виглядає як перехід електрона з валентної зони в зону провідності. Природно, що квант із енергією, меншою за ширину забороненої зони, не зможе порушити валентний зв'язок, тому речовина для таких

квантів буде прозорою, якщо немає інших видів поглинання. Відповідно спектральна крива коефіцієнта поглинання (рис. 6.3, б) починається із частоти $\nu_0 = \Delta E/h$, названої *граничною частотою власного поглинання*.

На частотах, вищих за граничну, коефіцієнт поглинання зростає, тому що в міжзонних переходах беруть участь електрони дедалі глибших станів валентної зони. Результатом фундаментального поглинання є утворення електронно-діркових пар або *біполярна генерація*.

Іншим важливим видом поглинання з погляду утворення носіїв заряду є *домішкове поглинання*. У цьому випадку кванти поглинаються домішковими атомами (наприклад, донорними), а

енергія поглиненого кванта витрачається на розрив валентного зв'язку між атомом домішки і його «зайвим» електроном. У результаті утворюється вільний електрон і додатно іонізований донорний атом. На енергетичній діаграмі (рис. 6.4, а) цей процес зображено як перехід електрона з донорного стану в зону провідності. Природно, що іонізація донорів можлива лише за умови, що енергія кванта перевищує енергію іонізації, тобто на частотах $\nu > \nu_0 = \Delta E_d/h$. Однак на експериментальних спектральних кривих спостерігаються піки поглинання (рис. 6.4, б) на частотах, нижчих за граничну.

Це пов'язано з тим, що «зайвий» електрон донорного атома крім основного стану E_d має ще й стани збудження. Кванти світла з енергією, меншою за ΔE_d , не можуть викликати іонізацію домішки, але можуть викликати її збудження в різні дискретні стани (E_1, E_2, \dots на енергетичній діаграмі рис. 6.4, а). Таке поглинання не приводить до утворення рухливих носіїв і є «фотоелектрично неактивним».

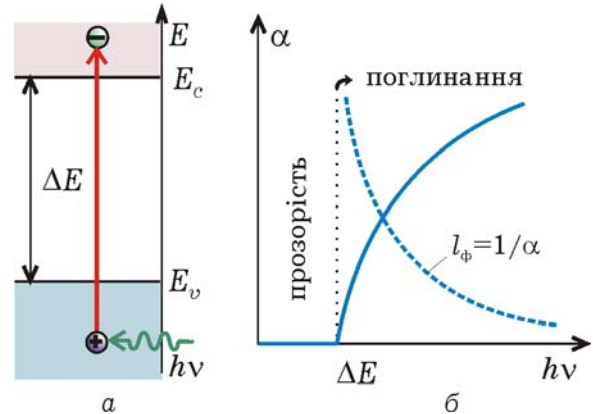


Рис. 6.3. Енергетичні переходи електронів при фундаментальному поглинанні й відповідна спектральна характеристика. Штрихова лінія показує залежність середньої глибини поглинання квантів від їхньої енергії

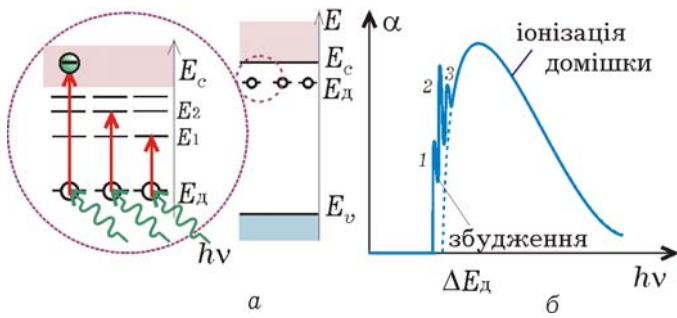


Рис. 6.4. Енергетичні переходи електронів для домішкового поглинання (а) й відповідна спектральна характеристика (б). Піки 1, 2, 3 на характеристиці відповідають переходам електронів на рівні збудження, показані на збільшеній вставці

в окремий вид *емісійного поглинання*, тобто поглинання таких квантів світла, які можуть викликати емісію електронів з речовини. Для напівпровідників його насправді можна розглядати як різновид власного (або домішкового) поглинання, що

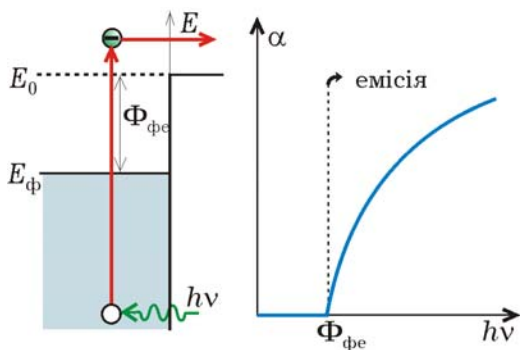


Рис. 6.5. Збудження електронів під час фотоемісії, яка можлива тільки у разі їхніх переходів у стани вище рівня вакууму

діаграмі від найвищого заповненого електронами стану до рівня вакууму. Для металів ця величина збігається з термодинамічною роботою виходу, але для напівпровідників – це різні поняття, тому що рівень Фермі (у невідроджених напівпровідниках) розташовується в забороненій для електронних станів зоні.

У результаті домішкове поглинання у донорних напівпровідниках може викликати утворення вільних електронів, а в акцепторних – утворення дірок у валентній зоні. Тому результатом домішкового поглинання є *монополярна генерація*, тобто утворення рухливих носіїв лише одного знака.

Спірним з погляду класифікації є виділення супроводжується збудженням електронів у стани зони провідності вище рівня вакууму, звідки можливий їхній вихід через потенціальний бар'єр. Однак для металів (рис. 6.5) це поглинання можна виділяти в окремий вид.

Емісія електронів неможлива, якщо енергія квантів світла менша, ніж *фотоелектронна робота виходу*, тобто відстань на енергетичній

6.2. ВЛАСТИВОСТІ ФОТОЕФЕКТІВ

6.2.1. Поведінка збуджених електронів

Поведінка збуджених електронів визначає й пояснює як спільні риси, так і розбіжності між внутрішнім і зовнішнім фотоэффектами, а це, у свою чергу, визначає технічні параметри фотокатодів для вакуумних фотоелементів і фоторезисторів.

Одним з основних тверджень, які привели Ейнштейна до успіхів у поясненні теорії фотоэффекту, була гіпотеза про те, що світло поглинається у речовині квантами і в кожному одиничному акті має поглинатися енергія не менша, ніж $h\nu$. Якщо такий квант поглинається електроном, то останній збуджується, тобто переходить у стан (рис. 6.6), який на величину $h\nu$ вищий за початковий (з точністю до малих значень енергій фононів, що беруть участь у процесі поглинання).

У початковий момент поведінка оптично збудженого електрона може істотно відрізнитися від поведінки рівноважних електронів, які перебувають у зоні провідності за рахунок термічного збудження. Оптично збуджені електрони – нерівноважні, і їхня енергія в момент утворення може значно перевищувати середню енергію рівноважних, рівну $3k_B T/2$ у невідроджених напівпровідниках. Через це такі електрони називаються «гарячими».

Якщо енергії поглиненого кванта достатньо, то «гарячі» електрони можуть опинитися в станах зони провідності, що лежать вище рівня вакууму. Однак унаслідок зіткнень із кристалічною ґраткою, а точніше з фононами, «гарячий» електрон дуже швидко розсіює свою енергію і переходить на вільні, більш низькі дозволені рівні. За сприятливих умов, рухаючись хаотично до поверхні, «гарячий» електрон може подолати потенціальний бар'єр, якщо його енергія залишилася вищою за рівень вакууму. Саме такі «гарячі» електрони в початковий момент свого існування зазвичай і зумовлюють фотоелектронну емісію.

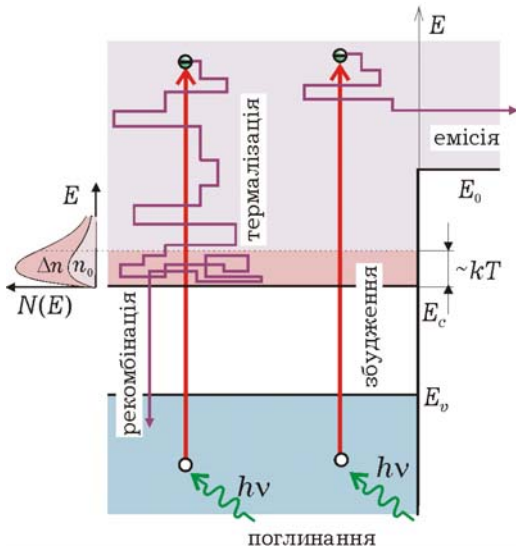


Рис. 6.6. Енергетична діаграма поведінки електронів під час їхнього збудження, емісії, термалізації і рекомбінації. Фрагмент праворуч показує вигляд функції розподілу нерівноважних носіїв

електрон розгубить за $10^{-11} \dots 10^{-12}$ с. Цей час релаксації надлишкової енергії й відповідає часу знаходження збудженого електрона в «гарячому» стані до його термалізації.

Енергія термалізованих електронів дорівнює, в середньому, кільком $k_B T$ і відповідає тим самим станам зони провідності, у яких перебувають термічно збуджені рівноважні електрони, концентрацію яких n_0 можна обчислити на основі рівноважної статистики:

$$n_0 = N_c \exp [(E_F - E_c)/k_B T]. \quad (6.8a)$$

Час життя термалізованих електронів у зоні провідності на кілька порядків вищий, ніж час термалізації, і може помітно перевищувати 10^{-8} с, після чого відбувається їхня рекомбінація з дірками валентної зони. Отже, практично весь час свого існування збуджені електрони проводять біля дна зони провідності й практично не відрізняються від рівноважних. Відповідно їхній енергетичний розподіл можна описати функцією розподілу (рис. 6.6), подібною до максвеллівської (для невиродженого колективу),

Інші електрони дуже швидко переходять у нижні стани зони провідності, або *термалізуються*. Тривалість процесу термалізації можна оцінити, виходячи з таких міркувань. Для розсіювання енергії, наприклад, в 1 еВ електрону потрібно кілька десятків зіткнень, тому що під час кожного зіткнення він може передати гратці, збуджуючи фонон, енергію, що не перевищує кількох сотих долей електрон-вольта. Оскільки середній час між двома послідовними зіткненнями в типових випадках становить $10^{-13} \dots 10^{-14}$ с, то всю свою надлишкову енергію «гарячий»

але в якій замість енергії E_{Φ} потрібно підставити положення квазірівня Фермі $E_{\Phi n}$. Тому концентрацію нерівноважних електронів можна визначити за співвідношенням

$$n = N_c \exp [(E_{\Phi n} - E_c)/k_B T], \quad (6.8б)$$

аналогічним (6.8а). Порівнявши ці вирази, одержимо

$$n = n_0 \exp [(E_{\Phi n} - E_{\Phi})/k_B T]. \quad (6.9)$$

Останнє співвідношення використовують для визначення положення квазірівня Фермі через надлишкову концентрацію електронів $\Delta n = n - n_0$, оптично збуджуваних у зону провідності:

$$E_{\Phi n} = E_{\Phi} + k_B T \ln(1 + \Delta n/n_0). \quad (6.10)$$

Звідси видно, що істотної зміни положення квазірівня можна досягти тільки у разі високого рівня освітленості, коли $\Delta n \gg n_0$. Якщо ж концентрація надлишкових носіїв порівнянна з рівноважною, то зсув квазірівня Фермі становить величину порядку $k_B T$. За малих рівнів освітленості зміною квазірівня можна знехтувати.

6.2.2. Інерційність фотоєфектів

З наведеного опису поведінки збуджених світлом електронів випливає, що емісія можлива лише поки вони перебувають у «гарячому» стані, не встигнувши термалізуватися, тому їхній вихід «затримується» після початку освітлення не більш ніж на $10^{-11} \dots 10^{-12}$ с. Це дозволяє вважати, що фотовідгук при зовнішньому фотоєфекті практично безінерційний. Крім того, з огляду на значення швидкості хаотичного руху електронів (порядку 10^5 м/с), можна зробити висновок, що вихід електронів відбувається із поверхневого шару, товщина якого не перевищує, в середньому, десятих або сотих часток мікрометрів.

Інша ситуація спостерігається у разі внутрішнього фотоєфекту. Оскільки процес термалізації відбувається практично миттєво, у провідності беруть участь електрони, які затримуються

біля дна зони провідності на час, необхідний для їхньої рекомбінації. Тому з увімкненням освітлення спостерігається їх «накопичення» і відповідне зростання провідності доти, доки не встановиться стаціонарний стан, у якому швидкість рекомбінації електронів зрівнюється зі швидкістю їхнього збудження.

Для кількісного аналізу часових залежностей фотовідгуку із вмиканням й вимиканням світла можна скористатися рівнянням неперервності

$$dn/dt = G_n - R_n. \quad (6.11)$$

де $G_n = G_v + G_0$ – швидкість генерації, що містить як оптичну, так і рівноважну теплову складові; R_n – швидкість рекомбінації.

Припускаючи, що і рекомбінація, і збудження нерівноважних носіїв відбуваються за рахунок міжзонних переходів ($\Delta n = \Delta p$), можна вважати $R_n = \alpha_r n p = \alpha_r (n_0 + \Delta n) \times (p_0 + \Delta p)$, а $G_0 = R_0 = \alpha_r n_0 p_0$, де α_r – коефіцієнт рекомбінації. Швидкість оптичної генерації можна оцінити у такий спосіб. Оскільки на одиницю поверхні падає $I_v/h\nu$ квантів світла за секунду, з яких $(1 - R)I_v/h\nu$ поглинається у поверхневому шарі, товщина якого приблизно дорівнює середній довжині їх «вільного пробігу» до поглинання $l_\phi = 1/\alpha$, і кожен квант збуджує в середньому β' електронів, то концентрація надлишкових електронів за одиницю часу зростає на $G_v = \beta'(1 - R)I_v/l_\phi h\nu$. Коефіцієнт пропорційності β' має сенс середньої кількості електронів, збуджуваних у зону провідності одним поглиненим квантом світла, і його називають *квантовим виходом*. Часто користуються поняттям квантового виходу β , розраховуваного на один *падаючий квант*. Очевидно, що $\beta = \beta'(1 - R)$, тоді $G_v = \beta \alpha I_v/h\nu$. У підсумку рівняння неперервності можна подати у формі

$$d(\Delta n)/dt = \beta \alpha I_v/h\nu - \Delta n/\tau_n - \alpha_r \Delta n^2, \quad (6.12)$$

де $\tau_n = 1/\alpha_r(n_0 + p_0)$ – час життя нерівноважних електронів. Звідси очевидно, що в стаціонарному стані ($d/dt = 0$) встановлюється значення концентрації надлишкових носіїв

$$\Delta n_{\text{ст}} = \frac{n_0 + p_0}{2} \left(\sqrt{1 + \frac{4\Delta n_{\tau}}{n_0 + p_0}} - 1 \right), \quad (6.13)$$

де $\Delta n_{\tau} \equiv \tau_n \beta \alpha I_{\nu} / h\nu$ – кількість збуджуваних світлом нерівноважних електронів за інтервал часу τ_n .

Рівняння (6.12) відрізняється від подібного (5.4) лише наявністю генераційного доданка у правій частині. Не розв'язуючи його повністю, обмежимося аналізом двох крайніх, але характерних ситуацій.

Перша. За низької освітленості ($\Delta n \ll n_0 + p_0$) рекомбінацію називають *лінійною*, тому що третім доданком у правій частині (6.12) можна знехтувати, і рівняння стає лінійним

$$d(\Delta n)/dt = \beta \alpha I_{\nu} / h\nu - \Delta n / \tau_n. \quad (6.14)$$

У стаціонарному стані, коли швидкості генерації й рекомбінації рівноважують одна одну,

$$\Delta n_{\text{ст}} = \Delta n_{\tau} \equiv \tau_n \beta \alpha I_{\nu} / h\nu. \quad (6.15)$$

Із вмиканням і вимиканням світла в моменти t_0^+ й t_0^- відповідно (рис. 6.7, а) перехідні процеси описують типовими виразами:

$$\Delta n(t) = \Delta n_{\text{ст}} \{1 - \exp[-(t - t_0^+) / \tau_n]\}, \quad (6.15а)$$

$$\Delta n(t) = \Delta n_{\text{ст}} \exp[-(t - t_0^-) / \tau_n]\}, \quad (6.15б)$$

які відповідають лінійній рекомбінації і графіки яких наведено на рис. 6.7, б. Релаксація фотопровідності в цьому випадку характеризується часом життя нерівноважних носіїв τ_n .

Друга. За великої інтенсивності світла ($\Delta n \gg n_0 + p_0$) переважаючим у правій частині (6.12) стає третій доданок

$$d(\Delta n)/dt = \beta \alpha I_{\nu} / h\nu - \alpha_r \Delta n^2, \quad (6.16)$$

і таку рекомбінацію називають *квадратичною*. Концентрація надлишкових носіїв (6.13) у стаціонарному стані встановлюється на рівні

$$\Delta n_{\text{ст}} = [\Delta n_{\tau} (n_0 + p_0)]^{1/2} \equiv [\beta \alpha I_{\nu} / (\alpha_r h \nu)]^{1/2}, \quad (6.17)$$

який нелінійно залежить від інтенсивності світла. Розв'язок (6.16)

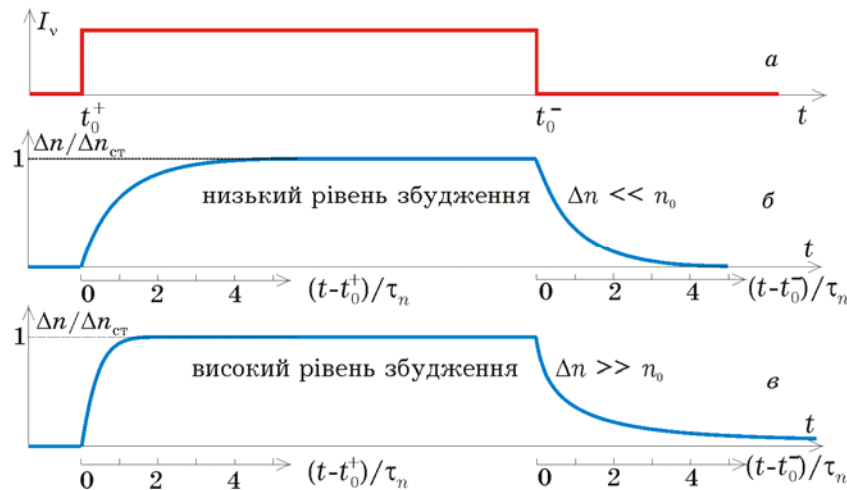


Рис. 6.7. Релаксаційні процеси відносної зміни надлишкової концентрації електронів із вмиканням і вимиканням світла (а) для різних рівнів освітленості: б – для малого рівня ($\Delta n_{\tau} \ll n_0 + p_0$), в – для великого ($\Delta n_{\tau} = 10(n_0 + p_0)$). Додаткові осі відповідають нормованим на час життя значенням, відліченим від характерного моменту перемикавання

з увімкненням світла матиме вигляд:

$$\Delta n(t) = \Delta n_{\text{ст}} \text{th}[\Delta n_{\text{ст}} \alpha_r (t - t_0^+)], \quad (6.18a)$$

а при вимиканні зі стаціонарного стану

$$\Delta n(t) = \Delta n_{\text{ст}} / [1 + \Delta n_{\text{ст}} \alpha_r (t - t_0^+)]. \quad (6.18б)$$

Графіки цих функцій наведено на рис. 6.7, в. Незважаючи на їхні розбіжності й відмінність від попереднього випадку, швидкість релаксаційних процесів тут також визначається часом порядку часу життя τ_n . Наочніше ця інерційність фотопровідності спостерігається після вимкнення освітлення: фотовідгук

продовжує існувати ще якийсь час до зникнення носіїв унаслідок їхньої рекомбінації. Цей ефект знижує швидкодію фоторезистивних елементів і її підвищення можливе за умови зменшення часу життя нерівноважних носіїв.

6.2.3. Лінійність світлової характеристики

Світлова характеристика фотоефектів, тобто залежність фотовідгуку від інтенсивності світла, також має особливості, зумовлені поведженням збуджуваних носіїв. Як зазначалося, у разі опромінення монохроматичним світлом кількість квантів, які поглинаються одиницею поверхні за 1 секунду, визначається інтенсивністю світла I_ν і дорівнює $(1 - R)I_\nu/h\nu$. Якщо вважати, що енергія кванта поглинається цілком в одиничному акті взаємодії і кожен електрон поглинає не більше одного кванта, то кількість емітованих з одиниці поверхні за 1 с електронів n_e буде пропорційна кількості поглинутих квантів:

$$n_e = \beta'(1 - R)I_\nu/l_\phi h\nu, \quad (6.19)$$

де β' – коефіцієнт пропорційності, що показує, скільки електронів емітується у середньому одним поглиненим квантом, і його також називають *квантовим виходом*. Часто під квантовим виходом β розуміють кількість емітованих електронів у перерахунку на один падаючий квант. Очевидно, що $\beta = \beta'(1 - R)$. У результаті для густини струму фотоемісії з поверхні можна одержати

$$j_\phi = en_e = e\beta'(1 - R)I_\nu/h\nu = e\beta I_\nu/h\nu = S_\nu(\nu) I_\nu. \quad (6.20)$$

Коефіцієнт пропорційності $S_\nu(\nu)$, що має тут розмірність А/Вт, називають *спектральною чутливістю*, а його залежність від частоти – *спектральною характеристикою* фотоемітера. Факт пропорційності між інтенсивністю падаючого світла й густиною фотоструму є змістом закону, встановленого 1888 року О. Г. Столетовим. Цей закон діє і в тому випадку, коли поверхня опромінюється не монохроматичним світлом, а світлом складного,

але постійного спектрального складу. У цьому випадку коефіцієнт пропорційності називають *інтегральною чутливістю*.

Подібний за формою закон спостерігається і для внутрішнього фотоефекту, але тільки за низької освітленості, коли концентрація збуджених електронів пропорційна інтенсивності. Для постійного освітлення встановлюється динамічна рівновага між процесами генерації та рекомбінації, при цьому стаціонарне значення надлишкової концентрації фотоелектронів визначається співвідношенням (6.15). Ці електрони викличуть зміну питомої електропровідності речовини на величину

$$\Delta\sigma = e\Delta n_{\text{ст}} \mu_n = e\mu_n\tau_n\beta\alpha I_\nu/h\nu = S_\nu(\nu) I_\nu. \quad (6.21)$$

Подібну зміну електропровідності створюють і дірки, які генеруються світлом у валентній зоні. Співвідношення (6.21), подібно до закону Столетова, характеризує зв'язок між інтенсивністю падаючого світла і фотопровідністю. Коефіцієнт пропорційності, як й у випадку фотоемісії, називають спектральною чутливістю, вираженою в сименс-метрах на ват.

Однак співвідношення (6.21) лінійне лише за низького рівня збудження, коли нерівноважних носіїв генерується набагато менше, ніж рівноважних, і, відповідно, значення фотопровідності менше так званої *темної провідності*. Але коли швидкість оптичного збудження стане порівнянною зі швидкістю теплового, час життя залежатиме від інтенсивності падаючого світла і лінійність порушиться. Якщо скористатися загальним виразом (6.13) для $\Delta n_{\text{ст}}$, то можна одержати для фотопровідності

$$\Delta\sigma = \frac{\sigma_0}{2} \left(\sqrt{1 + \frac{4\Delta n_\tau}{n_0 + p_0}} - 1 \right). \quad (6.22)$$

Цю залежність наведено на рис. 6.8 як функцію від рівня збудження, який зручно характеризувати параметром $\Delta n_\tau/(n_0 + p_0) \equiv \tau_n\beta\alpha I_\nu/h\nu(n_0 + p_0)$, пропорційним інтенсивності. З неї випливає, що провідність фоторезисторів нелінійно залежить від інтенсивності світла практично в усьому робочому діапазоні, коли фотовідгук стає порівнянний з темновим струмом. Лінійність

фотоемісії також може порушуватися, але за дуже великої інтенсивності, характерної для лазерного випромінювання.

6.2.4. Чутливість фотоелементів

Ефект накопичення електронів у зоні провідності за рахунок скінченного часу життя, що приводить до підвищеної інерційності внутрішнього фотоелементу, відіграє й позитивну роль – «за інших рівних умов» чутливість резистивних фотоелементів вища, ніж вакуумних, тому що в провідності бере участь більша кількість носіїв, ніж в емісії.

Якщо розглянути фотокатод і фоторезистор з однаковими освітлюваними площами $S = ab$ (рис. 6.9), то з (6.20) випливає, що величина емісійного фотоструму пропорційна потужності світлового потоку P_v :

$$i_{\text{ф}} = e\beta I_v S / h\nu = e\beta P_v / h\nu. \quad (6.23)$$

Для внутрішнього фотоелементу в шарі, товщина якого $l_{\text{ф}}$ дорівнює середній глибині поглинання квантів, збуджувані електрони забезпечують фотопровідність, яку можна знайти з (6.21). В електричному полі з напруженістю \mathcal{E} значення відповідного струму через поперечний переріз $al_{\text{ф}} = a/\alpha$

$$i_{\text{ф}} = al_{\text{ф}} \Delta\sigma \mathcal{E} = (\tau_n / t_{\text{пр}}) e\beta I_v S / h\nu = (\tau_n / t_{\text{пр}}) e\beta P_v / h\nu. \quad (6.24)$$

Тут вважається, що $\mu \mathcal{E} = v_{\text{др}} = b / t_{\text{пр}}$, де $t_{\text{пр}}$ – час прольоту електронів через зразок завдовжки b із дрейфовою швидкістю $v_{\text{др}}$.

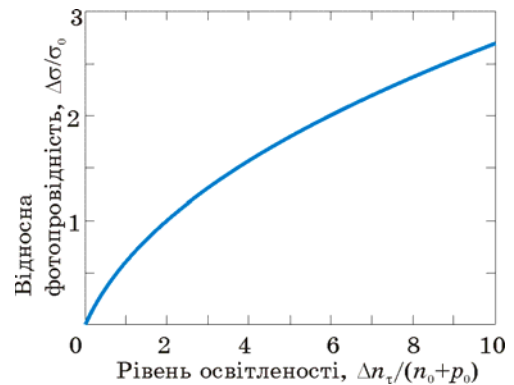


Рис. 6.8. Світлова характеристика внутрішнього фотоелементу в нормованих координатах

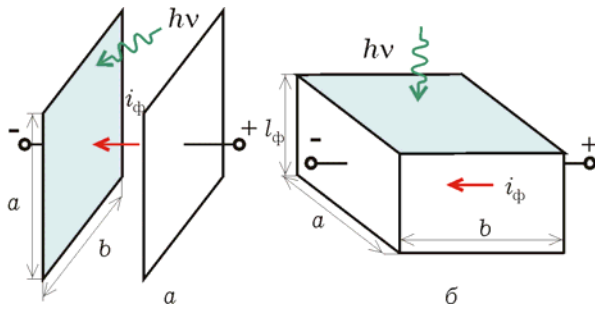


Рис. 6.9. Ескізи до порівняння чутливості зовнішнього (а) і внутрішнього (б) фото ефектів. Освітлювані поверхні (затушовані) мають однакову площу

провідності на час τ_n , рівний часу життя. Коефіцієнт підсилення може бути отриманий досить великим і становити кілька порядків.

Слід звернути увагу, що підвищення чутливості фоторезисторів за рахунок збільшення часу життя приводить до зменшення швидкодії.

6.2.5. Гранична частота фото ефектів

Гранична частота поглинання й відповідного фотовідгуку значною мірою визначає спектральні властивості фото елементів. Вона залежить від параметрів енергетичної структури речовини і може істотно відрізнитися для внутрішнього і зовнішнього фото ефектів.

Припустимо, що електрон, перебуваючи в стані з енергією E_1 , поглинає квант енергії $h\nu$ і, розсіявши частину δE своєї енергії, емітується з поверхні. Залишок його енергії у вакуумі проявляється у вигляді кінетичної енергії, яку знаходять з очевидного (див. рис. 6.10) балансу енергії:

$$E_k = E_1 + h\nu - \delta E - E_0. \quad (6.25)$$

Різні значення початкової енергії E_1 для електронів і випадковий характер втрат δE приводять до розкиду швидкостей виходу електронів. Але в енергетичному спектрі при фото емісії, викликаній монохроматичним світлом, спостерігатиметься характерна риса: наявність у функції розподілу деякого

Порівняння останніх виразів показує, що за інших рівних умов фотострум при фоторезистивному ефекті більший, ніж при фото емісії в $(\tau_n/t_{пр})$ разів. Це своєрідне підсилення викликане «затримкою» електронів у зоні

найбільшого значення $E_{\text{кmax}}$. Це значення матимуть ті електрони, які, по-перше, до поглинання кванта мали енергію $E_1 = E_m$, найбільшу для даної речовини, і, по-друге, емітувалися без втрат енергії ($\delta E = 0$):

$$E_{\text{кmax}} = h\nu - (E_0 - E_m). \quad (6.26)$$

Другий доданок у дужках називають *фотоелектронною роботою виходу* $\Phi_{\text{фе}}$. Він відповідає тій мінімальній енергії кванта $h\nu_0 = \Phi_{\text{фе}}$, за якої ще можливий вихід електронів з нульовою швидкістю. Тут треба звернути увагу на те, що поняття фотоелектронної і термодинамічної робіт виходу в загальному випадку не збігаються. З використанням поняття граничної частоти ν_0 вираз (6.26) можна навести у вигляді

$$E_{\text{кmax}} = h(\nu - \nu_0). \quad (6.27)$$

Це – найлаконічніша форма запису *закону Ейнштейна*: **максимальна** енергія емітованих фотоелектронів пропорційна частоті падаючого світла й не залежить від його інтенсивності. Останні слова вказують на неспроможність класичної теорії фотоэффекту, відповідно до якої, чим більша інтенсивність світла, тим сильніше електромагнітна хвиля «розгойдує» електрони в кристалічній ґратці й тим з більшою швидкістю вони зрештою вилітають.

Поняття граничної частоти існує й для внутрішнього фотоэффекту. Навіть без детального енергетичного аналізу зрозуміло, що збудження електронів з валентної зони у зону провідності можливе лише у тому випадку, коли $h\nu_0 \geq \Delta E$ (з точністю до енергії фононів, які трохи «розмивають» границю поглинання). Тому у власному напівпровіднику

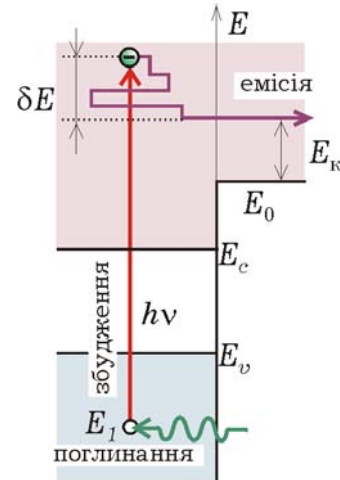


Рис. 6.10. Енергетична діаграма, що показує баланс енергії електронів під час фотоемісії, для виведення закону Ейнштейна

$$\nu_0 = \Delta E/h. \quad (6.28)$$

У донорному напівпровіднику для збудження носіїв у зону провідності потрібна ще менша енергія кванта – рівна енергії іонізації домішки, тому

$$\nu_0 = \Delta E_d/h. \quad (6.29)$$

Порівняння енергетичних параметрів речовин і відповідна оцінка граничних частот показують, що створити фотоелементи в довгохвильовій ділянці оптичного діапазону набагато простіше з використанням внутрішнього фотоелектричного ефекту. Діаграма рис. 6.11 ілюструє чутливість різних речовин до різних діапазонів електромагнітних хвиль. Так, наприклад, для створення вакуумних (емісійних) фотоелементів, чутливих навіть у ближній частині інфрачервоного діапазону, потрібна речовина з фотоелектронною роботою виходу близько 1 еВ. Створити фотокатоди з такою або меншою роботою виходу досить складно.

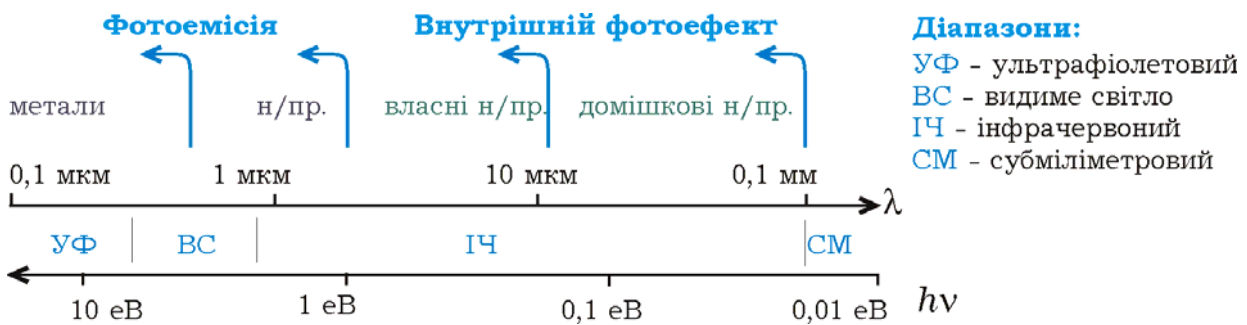


Рис. 6.11. Діаграма, що характеризує типові діапазони застосовності різних фотоелектричних ефектів: н/пр – напівпровідник

Водночас напівпровідник із шириною забороненої зони в кілька десятків часток електрон-вольта може служити фоторезистивним елементом у широкій частині інфрачервоного діапазону. Використовуючи донорні напівпровідники з різною енергією іонізації домішки, можна створити елементи, чутливі не тільки до інфрачервоного, але й до субміліметрового діапазону.

6.3. ФОТОЕЛЕМЕНТИ

6.3.1. Фотоемітери

Спектральною характеристикою фотокатода називають залежність чутливості до монохроматичного світла від його частоти. На вигляд спектральної характеристики впливають оптичні властивості речовини (залежність коефіцієнта відбивання від частоти), енергетичні параметри зонної структури, конструктивні особливості фотоелемента та зовнішні умови (насамперед, температура).

Розглянемо спочатку основні особливості типової спектральної характеристики металевих емітерів (рис. 6.12). Гранична частота фотоemisії відповідає фотоелектронній роботі виходу, яка для металів збігається з термодинамічною. При $T > 0$ границі немає, тобто на частотах, нижчих за граничну, фотострум існує за рахунок електронів, термічно збуджених вище рівня Фермі, які поглинули квант світла. Для більшості металів, що мають роботу виходу 4...5 еВ, границя фотоелектричного ефекту лежить в ультрафіолетовій області і лише для лужних і лужноземельних металів, у яких робота виходу становить 1,6...2 еВ, гранична частота потрапляє у видиму область.

На частотах, близьких до граничної, емісія забезпечується електронами, що мали до поглинання енергію, близьку до енергії Фермі. Збільшення частоти приводить до того, що в емісії можуть брати участь електрони дедалі глибших рівнів. Це, природно, викликає ріст спектральної кривої доти, доки енергія кванта не стане рівною повній роботі виходу $U_0 = E_0 - E_c$. У такому разі будь-

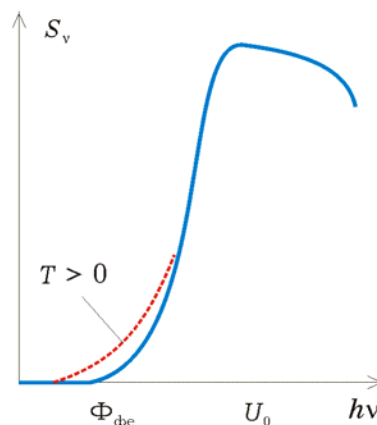


Рис. 6.12. Типовий вигляд спектральної характеристики фоточутливості металів. Для $T > 0$ границя «розмивається», тобто фотострум існує на частотах, нижчих за граничну, за рахунок додаткового термічного збудження

який електрон, що навіть перебував коло дна зони провідності, поглинувши квант, має можливість емітувати, а на характеристиці спостерігається «насичення» щодо кількості вихідних електронів. Така ситуація спостерігається вже в ультрафіолетовій області, тому що повна робота виходу для більшості металів становить 10 еВ і більше.

Фотоефект, що характеризується монотонним зростанням спектральної чутливості із підвищенням частоти, називають *нормальним*. Розрізняють ще *селективний* (вибірковий) фотоефект, що характеризується наявністю максимумів (одного або декількох) спектральної чутливості.

Подальше підвищення частоти приводить до зниження чутливості, тому що зростає енергія квантів і їхня кількість за постійної інтенсивності світла зменшується. Формально це враховано наявністю величини $h\nu$ у знаменнику виразів для кількості поглинутих квантів (див. підрозд. 6.1). Крім того, зі збільшенням частоти ослаблюється взаємодія фотонів з електронами, що виражається у збільшенні прозорості металу, тобто й у збільшенні довжини «пробігу» фотонів. «Гарячі» електрони утворюються у такому разі далі від поверхні, і їхній вихід з металу утруднюється.

Слід також мати на увазі, що на спектральну характеристику фотоелектронних приладів впливає спектральна характеристика вікна, через яке освітлюється фотокатод. Навіть для найкращих вікон прозорість падає на хвилях, коротших за 0,1...0,2 мкм.

В області нормального фотоефекту чутливість металевих катодів дуже мала, квантовий вихід становить усього $10^{-3} \dots 10^{-6}$. Це пояснюється насамперед високою відбивною здатністю металевих поверхонь – лише сота частина всіх падаючих фотонів проникає у метал і поглинається. Малому квантовому виходу сприяє й сильне розсіювання енергії «гарячих» електронів у результаті їхньої взаємодії з електронами зони провідності, концентрація яких у металі дуже велика.

Фотоелектронна емісія з напівпровідників має ряд особливостей порівняно з емісією із металів. Фоточутливість напівпровідників помітно вища, ніж металів. Квантовий вихід у найефективніших напівпровідникових емітерів може досягати

0,2...0,4 і близький до теоретичної межі, рівної 0,5. Квантовий вихід істотно залежить від співвідношення між шириною забороненої зони й енергією електронної спорідненості. Наприклад, за однакової фотоелектронної роботи виходу (рис. 6.13), квантовий вихід у Cs_3Sb ($\Delta E/\chi \approx 4$) приблизно на порядок вищий, ніж у Cs_3Bi ($\Delta E/\chi \approx 1/3$). Пояснюється це тим, що для $\Delta E < \chi$ виявляється ще один механізм розсіювання: збуджені електрони мають достатню енергію, щоб у свою чергу збудити «вторинні» електрони з валентної зони. Це, звичайно, підсилює розсіювання енергії «гарячих» електронів і знижує ймовірність їхньої емісії.

Спектральна характеристика напівпровідникових емітерів зазвичай має яскраво виражений селективний характер і часто містить не один, а кілька максимумів. Це може бути пов'язане з багатошаровою структурою катода з різними оптичними властивостями шарів. Через внутрішні відбивання у шарі, що емітує, утворюються вузли або пучності стоячої світлової хвилі, що знижує або підвищує чутливість залежно від частоти.

В однорідних напівпровідниках селективність пов'язана насамперед з особливостями зонної структури. Так, до появи максимуму найчастіше приводить наявність донорної домішки. Енергія кванта, необхідна для збудження електронів донорних атомів вище рівня вакууму, зазвичай набагато менша фотоелектронної роботи виходу з основної речовини. Тому на спектральній характеристиці донорного напівпровідника з'являється максимум на частотах нижчих, ніж у власних напівпровідниках.

Фотоемісійні властивості напівпровідників сильно залежать від стану поверхні. Якщо на поверхні є заповнені електронами енергетичні стани, що потрапляють у заборонену зону, можлива

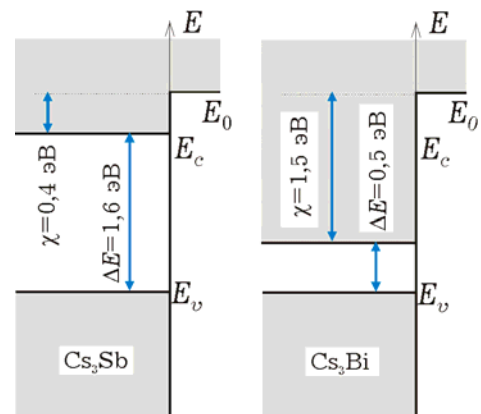


Рис. 6.13. Енергетичні діаграми напівпровідників Cs_3Sb й Cs_3Bi , що мають однакоvu фотоелектронну роботу виходу, але різні відношення $\Delta E/\chi$

емісія із цих станів на частотах, нижчих за граничну.

Однак роль поверхневих станів важлива не тільки з огляду їхньої прямої участі в емісії. Не менш важливим є обмін зарядами між поверхневими й об'ємними станами, що приводить до появи електричного поля у поверхневому шарі, який істотно впливає на фотоемісію. Із цього погляду напівпровідники *p*-типу є кращими емітерами порівняно з *n*-типом.

В акцепторних напівпровідниках електрони з поверхневих станів переходять до акцепторних атомів й утворюється подвійний електричний шар із від'ємно іонізованих акцепторів і додатних зарядів поверхневих станів. Електричне поле викривляє енергетичну діаграму приповерхневого шару донизу, знижуючи енергію електронної спорідненості. У донорних напівпровідниках спостерігається протилежна ситуація – поле від'ємного заряду на поверхні й додатно іонізованих донорів у поверхневому шарі підвищує потенціальний бар'єр, підвищуючи енергію електронної спорідненості.

Описаний прийом використовують для створення *фотокатодів з від'ємною електронною спорідненістю* (ВЕС). Для цього поверхню напівпровідника з великою концентрацією акцепторів покривають моноатомним шаром електропозитивних

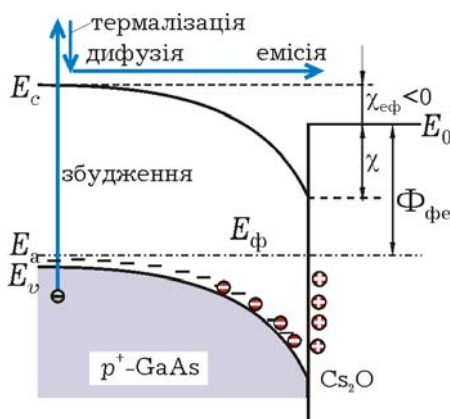


Рис. 6.14. Енергетична діаграма біля поверхні фотокатода з ВЕС

атомів. Обмін електронами приводить до появи тонкого (порядку 100 \AA) подвійного електричного шару, електричне поле якого викликає крутий вигин зон і зниження енергії електронної спорідненості до від'ємних значень (рис. 6.14).

Особливістю таких катодів є те, що в них полегшується вихід електронів насамперед із глибоких шарів з області плоских зон. Для електронів, збуджених біля поверхні, потенціальний бар'єр залишається колишнім. «Гарячі» електрони,

збуджені в глибині, практично миттєво термалізуються й протягом часу життя в зоні провідності рухаються хаотично. Тому рух їх до поверхні має дифузійний характер, а глибина, з якої можлива

емісія, відповідає дифузійній довжині.

З огляду спектральної характеристики особливістю фотокатодів з ВЕС є, по-перше, те, що гранична частота визначається, як і для внутрішнього фотоефекту, шириною забороненої зони використовуваного напівпровідника й знаходиться переважно в інфрачервоній області. По-друге, зниження електронної спорідненості підвищує їхню чутливість насамперед поблизу границі фотоефекту, де коефіцієнт фундаментального поглинання малий й кванти поглинаються у глибині речовини. Із підвищенням частоти зменшується довжина «вільного пробігу» фотонів і вони збуджують електрони ближче до поверхні, де ефект зниження електронної спорідненості виявляється слабкіше.

Катоди з ВЕС використовують як фотоприймачі насамперед інфрачервоного випромінювання, зокрема у приладах нічного бачення.

6.3.2. Фоторезистивні матеріали

Спектральні характеристики деяких фоторезистивних матеріалів наведено на рис. 6.16.

Для виготовлення фоторезисторів використовують сульфід і селенід кадмію й свинцю, телурид свинцю тощо у вигляді монокристалів, полікристалічних плівок, спресованих і спечених з порошкоподібного матеріалу таблеток. Сульфід кадмію за кімнатної температури має ширину забороненої зони 2,4 еВ, тому границя власного поглинання знаходиться у видимому діапазоні ($\lambda \approx 0,5$ мкм). Введення до складу CdS домішок хлору (донори), міді або срібла (акцептори) сприяє збільшенню часу життя носіїв заряду, тобто приводить до підвищення чутливості. У середньому чутливість CdS становить 6 мА/В·лм, а в деяких плівкових зразках 400 мА/В·лм, темновий опір $10^7 \dots 10^{10}$ Ом, поріг чутливості $\sim 10^{-10}$ лм, однак вони інерційні ($\tau = 1 \dots 140$ мс).

Чутливість селенідо-кадмійових фоторезисторів ще вища (30...600 мА/В·лм), вони менш інерційні, мінімальний світловий потік, що виявляється, має порядок 10^{-11} лм. Ширина забороненої зони PbS за кімнатної температури становить 0,37 еВ, що

відповідає порогу фотопровідності $\lambda_0 \approx 3,3$ мкм; в PbSe $\Delta E = 0,27$ еВ й $\lambda_0 \approx 4,6$ мкм. Ці фоторезистивні матеріали призначені для роботи в середній інфрачервоній області спектра. При охолодженні до температури рідкого азоту характеристики суттєво зміщуються у довгохвильову область.

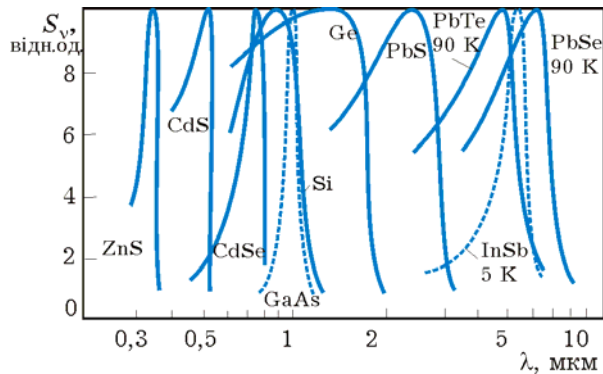


Рис. 6.16. Спектральні характеристики фоточутливості деяких фоторезистивних напівпровідникових сполук

поширення через малу чутливість.

Для реєстрації випромінювання далекої інфрачервоної області спектра ($\lambda_0 > 6$ мкм) використовують, головним чином, домішкові фоторезистори на основі германію, кремнію і їхніх сплавів. Як домішки для легування германію використовують золото, цинк, сурму, мідь, кадмій і ртуть. Zn, Cu, Cd, Hg утворюють акцепторні рівні, Au – три акцепторні й один донорний, Sb застосовують як компенсвальну домішку.

У субміліметровому діапазоні використовують фотоприймачі з InSb з донорною фотопровідністю, охолоджувані рідким гелієм. Значення домішкової фотопровідності залежить від концентрації домішки й від ступеня термічної іонізації за цієї температури. Для кімнатної температури домішки з $\Delta E_{\text{дом}} \approx 0,1 \dots 0,01$ еВ практично повністю термічно іонізовані. Тому для одержання домішкового поглинання світла необхідно охолоджувати напівпровідник – і тим сильніше, чим менша енергія іонізації домішки.

Для виготовлення фоторезисторів на основі власної фотопровідності використовують також сульфід талію, сульфід і селенід цинку, арсенід галію, антимонід й арсенід індію, а також германій з домішкою марганцю й кремній з домішкою цинку. Фоторезистори із власного Ge й Si не одержали

Контрольні запитання та вправи

6.1. Сформулюйте закон Бугера-Ламберта. Яким параметром визначається швидкість поглинання світла? Що таке довжина «вільного пробігу» фотонів?

6.2. Перерахуйте фотоелектрично активні види поглинання та пов'яжіть їх граничні частоти з характерними параметрами зонної структури.

6.3. Аналізуючи закони збереження енергії та імпульсу, опишіть особливості взаємодії фотонів і електронів. Яка роль фононів у цій взаємодії? Як це проявляється у прямозонних та непрямозонних напівпровідниках?

6.4. Опишіть відмінність у поведінці збуджених електронів під час внутрішнього та зовнішнього фотоефектів.

6.5. Дайте порівняльну характеристику внутрішнього і зовнішнього фотоефектів і відповідних фотоприймачів по швидкодії, чутливості, лінійності світлової характеристики і граничній частоті.

6.6. Опишіть фотоелектронні процеси, що відбуваються при освітленні електронно-діркового переходу. Для яких цілей можна використовувати різні режими роботи переходу?

7. ТВЕРДОТІЛЬНІ СТРУКТУРИ І ПРОЦЕСИ

7.1. ЕФЕКТ ПОЛЯ

З перерозподілом носіїв у електричному полі (див. 3.6.2) пов'язаний один з найважливіших для прикладної мікроелектроніки ефектів. Він полягає у зміні поверхневої концентрації носіїв заряду, тобто і поверхневої провідності, під впливом зовнішнього поперечного електричного поля. Це відбувається перш за все тому, що поле, діючи на рухливі носії, змінює їх концентрацію і викликає накопичення у поверхневому шарі просторового заряду. Його роль багатогранна: екранування електричного поля, зміна поверхневої провідності, вплив на емісійні властивості поверхні, тощо.

Для з'ясування причин утворення та властивостей шару поверхневого заряду розглянемо деякий зразок напівпровідника, розташований в однорідному електричному полі (рис. 7.1, а). Проміжок між поверхнею і електродом потрібен, щоб забезпечити відсутність електричного струму в колі за стаціонарного стану і тим самим – рівноважний розподіл рухливих носіїв. Якщо зразок є напівпровідником з великою кількістю рухливих носіїв, електричне поле викликати їх переміщення до границь, доки створюване ними поле не буде повністю компенсувати поле зовнішнього джерела всередині напівпровідника. Це явище зветься *екрануванням* зовнішнього електричного поля поверхневим шаром просторового заряду. Хоча поле всередині напівпровідника відсутнє, все ж силові лінії частково проникають вглиб речовини на певну відстань. Носії не накопичуються безпосередньо на поверхні, а дифундуючи від поверхні створюють рівноважний розподіл, який можна описати співвідношеннями (3.83) або (3.84). Стаціонарний стан встановлюється таким, що дифузійний потік від поверхні врівноважується дрейфовим потоком, спрямованим поверхневим електричним полем.

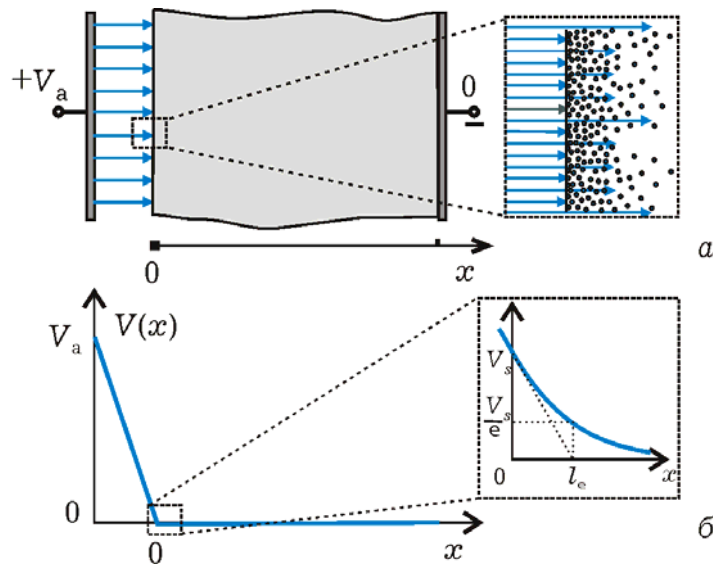


Рис. 7.1. Силві лінії зовнішнього електричного поля закінчуються на шарі поверхневого заряду, не проникаючи усередину провідної речовини. У кружках показано фрагменти розподілу поля і заряду в поверхневому шарі у збільшеному масштабі

Розподіл рівноважної концентрації носіїв i , відповідно, величину просторового заряду біля поверхні можна знайти, використовуючи співвідношення (3.84) в одновимірному випадку для невідродженого напівпровідника

$$n(x) = n_0 \exp[eV(x)/k_B T], \quad (7.1)$$

$$p(x) = p_0 \exp[-eV(x)/k_B T],$$

якщо відомий розподіл потенціалу біля поверхні. У свою чергу розподіл $V(x)$ можна знайти з рівняння Пуассона

$$\frac{d^2 V}{dx^2} = -\frac{\rho(x)}{\epsilon \epsilon_0} \quad (7.2)$$

через розподіл густини об'ємного заряду $\rho(x)$. Оскільки через електричну нейтральність напівпровідника за відсутності поля $\rho = 0$, то просторовий заряд створюється перерозподілом електронів і дірок, а його величина пропорційна перевищенню їх концентрацій над значеннями n_0 і p_0 :

$$\rho(x) = -e[n(x) - n_0] + e[p(x) - p_0]. \quad (7.3)$$

Останні три вирази являють собою систему, самоузгоджений розв'язок якої за заданих граничних умов дозволяє знайти розподіл потенціалу $V(x)$, густини заряду $\rho(x)$, рівноважних концентрацій $n(x)$ і $p(x)$ а також напруженості електричного поля $\mathcal{E}(x) = -dV/dx$.

Розглянемо такий розв'язок для власного напівпровідника. Оскільки в ньому $n_0 = p_0 = n_{i0}$, то після відповідних підстановок рівняння Пуассона набуде вигляду:

$$\frac{d^2V}{dx^2} = \frac{2en_{i0}}{\epsilon\epsilon_0} \operatorname{sh} \frac{eV}{k_B T}. \quad (7.4)$$

Для окремого випадку малих значень потенціалу $V \ll k_B T/e$ і $\operatorname{sh} \alpha \approx \alpha$ цей розв'язок є елементарним. Якщо вважати відомим потенціал на поверхні $V(0) = V_s$ і прийняти потенціал в об'ємі за початок відліку $V(\infty) = 0$, то розв'язок матиме вигляд функції

$$V(x) = V_s \exp(-x/l_e), \quad (7.5)$$

яка експоненціально спадає з віддаленням від поверхні (рис. 7.1, б). Швидкість спадання визначається величиною

$$l_e = \sqrt{\epsilon\epsilon_0 k_B T / 2e^2 n_{i0}}, \quad (7.6)$$

яку називають *довжиною екранування*. Розподіл густини заряду знайдемо з (7.1) та (7.3)

$$\rho(x) = \rho_s \exp(-x/l_e), \quad (7.7)$$

де $\rho_s = 2e^2 n_{i0} V_s / k_B T$. Інтегрування (7.7) по всій товщині зразка дає добуток $\rho_s l_e$, що зручно спрощено уявляти, нібито весь поверхневий заряд розподілений рівномірно з густиною ρ_s у шарі скінченної товщини $\rho_s l_e$. З цим і пов'язаний фізичний зміст і назва величини l_e .

Вираз (7.6) показує, що довжина екранування залежить від концентрації рухливих носіїв і може складати $\sim 10^{-7}$ м і більше. Якщо оцінити її значення для металів, хоча строго кажучи це невірно через використання тут статистики невиродженого електронного газу, то отримаємо, що довжина екранування буде співмірна з періодом кристалічної ґратки. Наявність температури в (7.6) пов'язана з дифузійними процесами, які відбуваються біля поверхні через градієнт концентрації рухливих носіїв.

Отримані розв'язки описують досить просту фізичну картину поведінки рухливих носіїв. Заряд біля поверхні утворюється електронами і дірками, густина якого на поверхні визначається додатковими концентраціями з (7.3)

$$\rho_s = -e[n_s - n_0] + e[p_s - p_0], \quad (7.8)$$

а поверхневі концентрації залежать від поверхневого потенціалу відповідно до (3.88)

$$n_s = n_0 \exp[eV_s/k_B T], \quad p_s = p_0 \exp[-eV_s/k_B T]. \quad (7.9)$$

Додатний потенціал V_s викликає підвищення концентрації електронів у поверхневому шарі і виштовхування з нього дірок. Якщо $V_s < 0$, спостерігається зворотна картина – збагачення шару дірками та збіднення електронами. Природньо, що зміна концентрації носіїв біля поверхні під дією зовнішнього електричного поля викликає зміну електричної провідності поверхневого шару (рис. 7.2). Це явище і зветься *ефектом поля*.

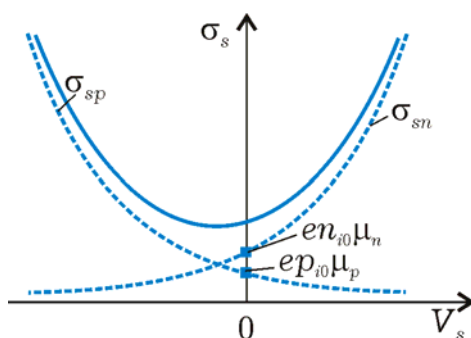


Рис. 7.2. Залежність поверхневої провідності від поверхневого потенціалу для власного напівпровідника

Для кількісного аналізу цього ефекту необхідно самоузгоджено, як систему, розв'язати рівняння Пуассона (7.89) разом із (7.88) та (7.90), не використовуючи наближення малого потенціалу. Для формальних спрощень зручно ввести безрозмірний потенціал, нормований на тепловий, $eV(x) / k_B T \equiv Y$, нормовану координату $x / l_e \equiv z$, та деякий параметр $\lambda \equiv p_0 / n_i = n_i / n_0$, що характеризує співвідношення між концентраціями носіїв у домішкових та власних напівпровідниках. За таких позначень рівняння Пуассона можна звести до вигляду

$$d^2Y / dz^2 = (e^Y - 1) / \lambda - (e^{-Y} - 1)\lambda. \quad (7.10)$$

Спростити його можна, понизивши порядок рівняння, якщо помножити всі його члени на $2dY/dz$ та зробивши тотожні перетворення:

$$\int_0^{dY/dz} d(dY/dz)^2 = \int_0^Y 2[(e^Y - 1) / \lambda - (e^{-Y} - 1)\lambda] dY. \quad (7.11)$$

Тут у нижніх межах інтегрування використано граничні умови – потенціал та його похідна в глибині напівпровідника обрано нульовими. Після інтегрування отримаємо для першої похідної

$$dY/dz = \pm \sqrt{2[(e^Y - Y - 1) / \lambda + (e^{-Y} + Y - 1)\lambda]} \equiv F(Y). \quad (7.12)$$

Подальше інтегрування в аналітичному вигляді неможливе і просторовий розподіл потенціалу можна знайти лише чисельним інтегруванням

$$\int_0^z dz = \int_{Y_s}^Y dY / F(Y). \quad (7.13)$$

Тут також нижньою межею інтегрування виступає умова $Y(0) = Y_s$, яка вважає заданим поверхневий потенціал V_s .

Результатом інтегрування є знаходження просторового розподілу потенціалу і подальшого розподілу носіїв, визначення надлишкової концентрації і їх внеску в поверхневу провідність.

Залежність додаткової провідності від поверхневого потенціалу показано на рис. 7.3 для донорних та акцепторних напівпровідників.

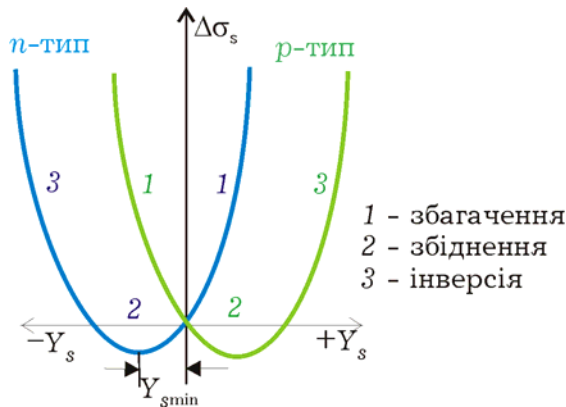


Рис. 7.3. Залежність додаткової поверхневої провідності від поверхневого потенціалу для донорних (ліва крива) та акцепторних (права крива) напівпровідників

Особливості фізичних процесів та характерних режимів, які визначають хід цієї залежності розглянемо на прикладі донорного напівпровідника. Вираз (7.3) для густини просторового заряду зручно, використовуючи рівняння електронейтральності (3.58), подати у вигляді

$$\rho(x) = -en(x) + ep(x) + eN_d^+. \quad (7.14)$$

Безпосередньо на поверхні просторовий заряд з врахуванням співвідношень (7.9 96) можна обчислити через поверхневий потенціал наступним чином:

$$\rho_s = -en_n \exp[eV_s/k_B T] + ep_n \exp[-eV_s/k_B T] + eN_d^+. \quad (7.15).$$

Цей вираз найбільш наочно показує вплив поверхневого потенціалу на співвідношення зарядів різного знаку на поверхні. За додатних потенціалів $V_s > k_B T/e$ концентрація електронів (рис. 7.4, а) стає більшою за концентрацію іонізованих донорів. Концентрація дірок стає ще меншою, хоча і так вона була на кілька порядків меншою, ніж концентрація електронів. Відповідно такий режим зветься *режимом збагачення*. На енергетичній діаграмі за рахунок додатного потенціалу це виглядає як викривлення енергетичних зон донизу і накопичення основних носіїв у западині потенціального рельєфу біля поверхні. Треба мати на увазі, якщо $eV_s \geq E_c - E_\Phi$,

рівень Фермі потрапляє у зону провідності і поверхневий шар стає виродженим.

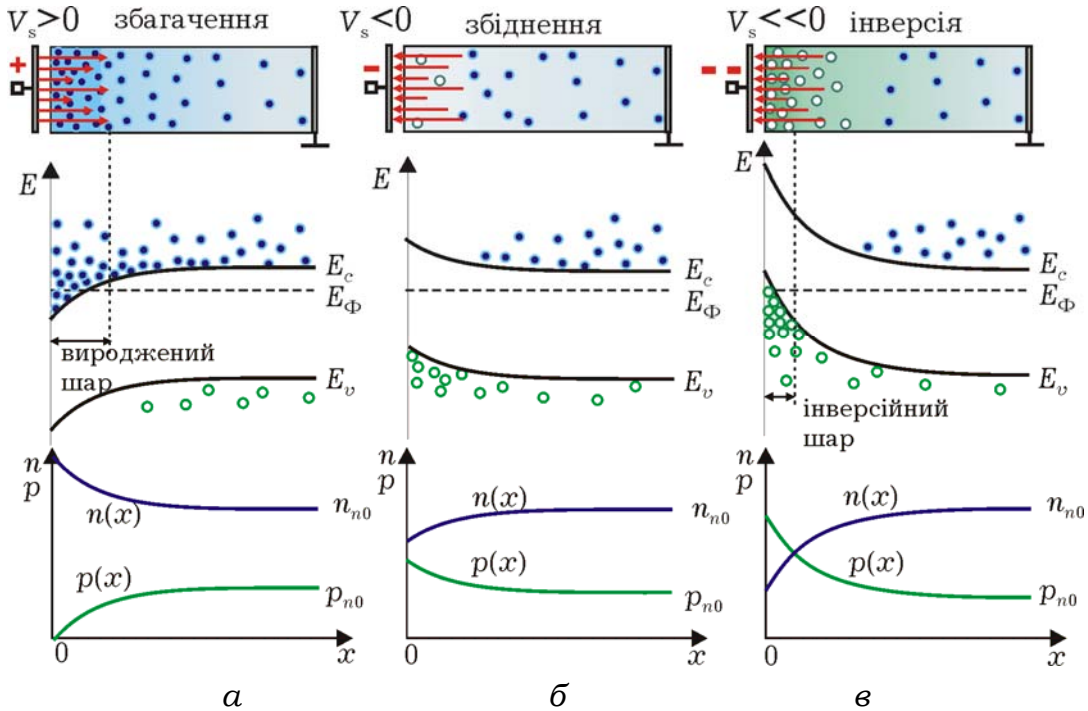


Рис. 7.4. Викривлення енергетичних зон та зміна концентрацій електронів та дірок у поверхневому шарі донорного напівпровідника. Для різних поверхневих потенціалів характерні режими: збагачення (а), збіднення (б) та інверсії (в)

За від'ємних потенціалів V_s зони вигинаються догори (рис. 7.4, б), поверхневий шар збіднюється електронами і їх концентрація стає меншою за концентрацію іонізованих донорів. Таку ситуацію називають *режимом збіднення*. Не зважаючи на те, що поверхневий шар збагачується дірками, за умови $p_n \exp[e|V_s|/k_B T] < N_d^+$ їх поки залишається менше, ніж електронів, а в (7.14) найбільшу вагу має третій доданок. Найменше рухливих носіїв у збідненому шарі буде коли концентрації електронів і дірок вирівнюються (перші два доданки) і стають рівними n_i . У такому разі поверхнева провідність дорівнюватиме провідності власного напівпровідника

Подальше зменшення потенціалу (рис. 7.4, в) викривлює зони догори ще більше, а дірок на поверхні стає більше за концентрацію електронів, тому провідність починає зростати. Такий режим називають *інверсією* типу провідності. Для від'ємних потенціалів, коли виконується нерівність $p_n \exp[e|V_s|/k_B T] > N_d^+$, дірок стає

більше, ніж електронів в об'ємі, тобто в режимі «плоских зон». Слід звернути увагу на те, що інверсійний шар відділений від глибинного об'єму шаром, збідненим електронами.

Детальніше розглянемо найпоширеніший **режим збіднення**, коли заряд у поверхневому шарі визначається практично тільки іонізованими донорами і переважає третій доданок в (7.14). Якщо до того ж прийняти, що всі домішки іонізовані і рівномірно розподілені у просторі, то у правій частині рівняння Пуассона залишається постійна величина:

$$\frac{d^2V}{dx^2} = -\frac{eN_D}{\epsilon\epsilon_0}. \quad (7.16)$$

Перша похідна, тобто напруженість електричного поля, є функцією, яка лінійно спадає з віддаленням від поверхні (рис. 7.5, б)

$$\mathcal{E}(x) = -\frac{dV}{dx} = -\frac{eN_D}{\epsilon\epsilon_0}(d-x), \quad (7.17)$$

а потенціал є параболічною функцією (рис. 3.21, в)

$$V(x) = -\frac{eN_D}{2\epsilon\epsilon_0}(d-x)^2. \quad (7.18)$$

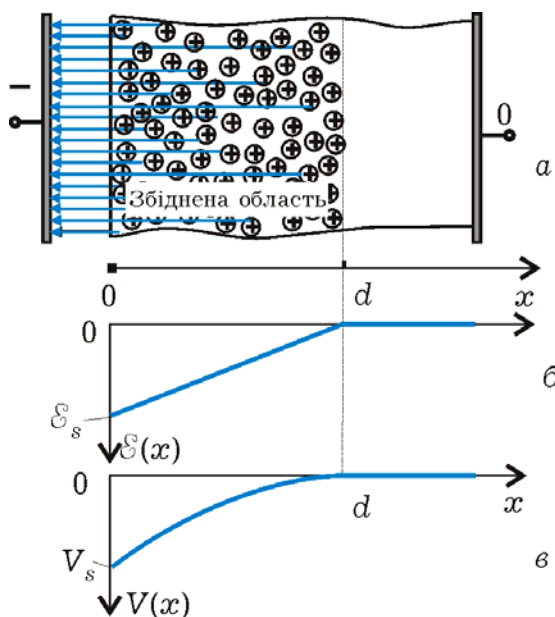


Рис. 7.5. Заряд у поверхневому шарі донорного напівпровідника у режимі збіднення (а) визначається практично лише іонізованими донорними атомами Розподіл напруженості електричного поля (б) є лінійною функцією, а потенціалу (в) – параболічною. За товщину збідненої області обирається відстань, де апроксимоване значення поля стає рівним нулю

Тут граничними умовами обрано: відомий потенціал $V_s < 0$ на поверхні та нульовий потенціал в об'ємі далеко від поверхні. Для експоненціального розподілу (7.5) такою була умова $V(\infty) = 0$, але у даному випадку зручно використати наближення різкої границі збідненого шару і прийняти, що потенціал дорівнюватиме нулю там, де напруженість поля (7.17) стає рівною нулю, тобто на відстані $d = \sqrt{2\varepsilon\varepsilon_0 |V_s| / eN_d}$ від поверхні. Ця товщина тим менша, чим більша концентрація домішки, а також залежить від поверхневого потенціалу. Цей факт використовується, наприклад, для керування провідністю поверхневого шару у польових транзисторах.

Режим збагачення – це ще один з часто використовуваних режимів у мікроелектронних компонентах. Як приклад розглянемо особливості збагаченого шару в напівпровіднику з концентрацією донорів N_d , на поверхні якого створено додатний потенціал V_s . У рівнянні Пуассона в правій частині має стояти просторовий заряд, створений надлишком концентрації електронів над концентрацією іонізованих донорів:

$$\frac{d^2V}{dx^2} = \frac{e}{\varepsilon\varepsilon_0} [n(x) - N_d]$$

де в рівноважній ситуації $n(x) = n_0 \exp[eV(x) / k_B T]$.

Обираючи початок відліку потенціалу в глибині, де $n(x) = N_d$ (рис. 7.6), маємо $n_0 = N_d$. На поверхні ($x = 0$), де потенціал рівний V_s , $n_s = N_d \exp[eV_s / k_B T]$.

З використанням позначень: $l_e \equiv \sqrt{\varepsilon\varepsilon_0 k_B T / e^2 N_d}$ ¹, $Y \equiv eV / k_B T$, $z \equiv x / l_e$, рівняння Пуассона прийме вигляд:

$$d^2Y / dz^2 = \exp Y - 1 \quad (7.19)$$

Помножаючи обидві частини рівняння (7.19) на $2 dY / dz$

$$2 \frac{dY}{dz} \frac{d}{dz} \left(\frac{dY}{dz} \right) = 2(\exp Y - 1) \frac{dY}{dz},$$

¹ Слід відрізняти це позначення від (7.6) для власного напівпровідника.

понизимо порядок похідної

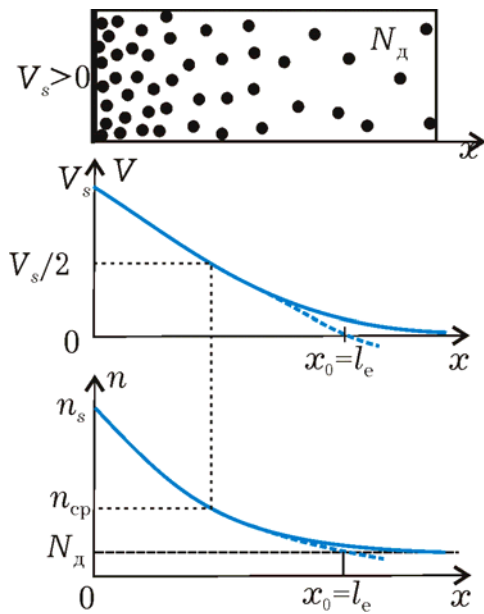


Рис. 7.6. Розподіли потенціалу і електронної концентрації поблизу поверхні донорного напівпровідника в режимі збагачення. Штрихові лінії відповідають наближеним розв'язкам.

виконуючи чисельне інтегрування

$$z \approx - \int_{Y_s}^Y \frac{dY}{\sqrt{2(\exp Y - Y - 1)}}. \quad (7.22)$$

Спрощені аналітичні співвідношення можна отримати за умови більш-менш значного збагачення $n \gg N_d$. Наприклад, якщо потенціал вже втричі перевищує тепловий ($Y > 3$), то досить добре виконується $\exp Y \gg Y$. Вводячи нормований потенціал Y_s на поверхні і інтегруючи (7.109) від Y_s до Y , отримаємо наближений розв'язок для потенціалу:

$$Y \approx Y_s - 2 \ln \left[1 + z \exp(Y_s / 2) / \sqrt{2} \right]. \quad (7.23)$$

$$d \left(\frac{dY}{dz} \right)^2 = 2(\exp Y - 1)dY.$$

а

Інтегруючи це рівняння, отримаємо:

б

$$dY / dz = \pm \sqrt{2(\exp Y - Y + C)}, \quad (7.20)$$

в

де C – константа інтегрування. З умови відсутності електричного поля ($dY / dz \approx 0$) в глибині підкладки, де $V = 0$, маємо $C = -1$:

$$dY / dz \approx -\sqrt{2(\exp Y - Y - 1)}. \quad (7.21)$$

Знак перед коренем обрано з фізичних міркувань – спадання поля вглиб матеріалу. З цього рівняння можна отримати розподіл потенціалу у неявній формі,

Прийняте припущення приводить до того, що потенціал досягає нульового значення не при $z \rightarrow \infty$, а в точці

$$z_0 = \sqrt{2} (1 - \exp(-Y_s / 2)) \approx \sqrt{2}, \quad (7.24)$$

тобто можна вважати, що шар збагачення має товщину $\sqrt{2}l_e$, яка є довжиною екранування у напівпровіднику з концентрацією електронів, рівною N_d .

Розподіл носіїв у збагаченому шарі може бути описаний виразом:

$$n(z) = N_d e^Y = N_d e^{Y_s} / \left[1 + z \exp(Y_s / 2) / \sqrt{2} \right]^2, \quad (7.25)$$

причому в точці z_0 справедлива рівність $n(z_0) = N_d$, а на поверхні $n(0) \equiv n_s = N_d \exp(Y_s)$. Інтегрування останнього виразу в межах шару дає, що середня по товщині надлишкова концентрація електронів визначається за формулою

$$\Delta n_{\text{ср}} = N_d e^{Y_s/2} = \sqrt{N_d n_s}. \quad (7.26)$$

Цікавою є інтерпретація цього результату (див. рис. 7.6): середнє значення співпадає зі значенням в точці, де потенціал вдвічі менший за поверхневий (перша рівність), або воно є середнім геометричним між значеннями концентрацій на краях збагаченого шару (друга рівність).

Слід відзначити, що електричне поле біля поверхні може створюватись по-різному. Це може бути зовнішнє поле, що проникає всередину напівпровідника, або поле, створюване іонізованими адсорбованими атомами на поверхні, або викликане накопиченням заряду на тампівських поверхневих станах, або контактне поле, як наприклад, у електронно-дірковому переході (рис. 3.16). Ефект поля покладено в основу цілого класу польових приладів, найбільш поширених в інтегральній мікроелектроніці.

7.2. КОНТАКТ ДВОХ МЕТАЛІВ

Утворення контакту між двома твердими тілами обов'язково викликає взаємний обмін рухливими носіями. За однакової

температури тіл у стаціонарному стані встановлюється термодинамічна рівновага, яка характеризується постійним значенням рівня Фермі, єдиним для всього колективу рухливих носіїв. Встановлення рівноваги супроводжується переходом електронів, який здійснюється у результаті термоелектронної емісії або тунельного проходження, або дифузійних процесів в залежності від якості контакту та властивостей контактуючих тіл.

Розглянемо два метали, енергетичні діаграми яких подано на рис. 7.7, а. За достатньо малої відстані між ними можливий обмін електронами у результаті термоелектронної емісії. Якщо $\Phi_1 < \Phi_2$, то струм емісії j_1 буде перевищувати j_2 , і другий метал буде отримувати надлишковий електронний заряд. Зменшення його потенціалу викличе появу у проміжку електричного поля, яке заважатиме переходу електронів з першого металу у другий. На енергетичній діаграмі (рис. 7.7, б) це можна відобразити взаємним зміщенням енергетичних рівнів і появою додаткового потенціального бар'єру eV_k , який зменшує струм j_1 і практично не впливає на струм j_2 :

$$\begin{aligned} j_1 &= A_1 T^2 \exp[-(\Phi_1 + eV_k) / k_B T], \\ j_2 &= A_2 T^2 \exp(-\Phi_2 / k_B T). \end{aligned} \quad (7.27)$$

Дані вирази описують струми, які виникають при проходженні електронів через відповідні потенціальні бар'єри (див. додаток 3). Рівновага встановиться, коли зрівняються зустрічні потоки електронів. Значення рівноважного потенціального бар'єру знаходиться з умови $j_1 = j_2$ (для спрощення приймемо $A_1 = A_2$):

$$eV_k = \Phi_2 - \Phi_1 \quad (7.28)$$

Різниця потенціалів V_k , яка виникає між металами, зветься контактною. Зазначимо, що перехід електронів з одного металу в інший змінює біля їх поверхні концентрацію електронів, але настільки незначно, що значення енергій Фермі, які відраховується від дна зони провідності й залежать від концентрації електронів (див. розд. 3), практично не змінюються. Тому контактна різниця потенціалів визначається відстанню між початковими положеннями рівнів Фермі, або різницею між значеннями термодинамічних робіт виходу.

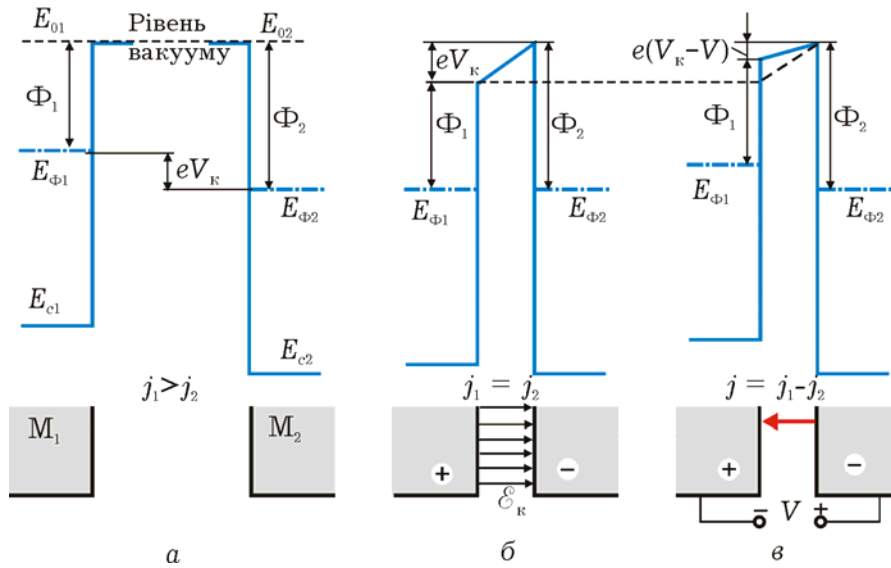


Рис. 7.7. Енергетичні діаграми двох металів до (а) та після (б) встановлення рівноваги, а також після порушення рівноваги зовнішньою напругою (в).

Рівновага контактуючих металів може бути порушена, якщо до них прикласти напругу V від зовнішнього джерела. Залежно від знаку цієї напруги (додатною будемо вважати напругу, яка включена назустріч контактній, як вказано на рис. 7.7, в) буде змінюватися величина бар'єру i , відповідно, величина струму j_1 :

$$j_1 = AT^2 \exp[-(\Phi_1 + eV_K - eV) / k_B T] \quad (7.29)$$

Результуючий струм j , що протікає через контакт, дорівнюватиме різниці $j_1 - j_2$. З урахуванням (7.28)

$$j = j_s [\exp(eV / k_B T) - 1], \quad (7.30)$$

де $j_s = AT^2 \exp(-\Phi_2 / k_B T)$ - так званий струм насичення.

Вираз (7.30) описує вольт-амперну характеристику контакту двох металів. Її вигляд подано на рис. 7.8 і свідчить про нелінійні властивості такого контакту, тобто про різний опір у прямому і зворотному напрямках. Але треба мати на увазі, що мова йде не про ідеальний контакт, а про контакт з достатньо малим вакуумним проміжком між металами.

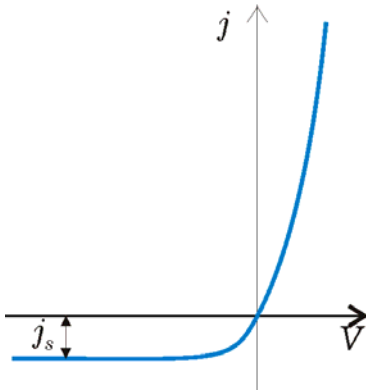


Рис. 7.8. Ідеалізована вольт-амперна характеристика контакту двох металів, розділених тонким вакуумним прошарком.

Розглянуті явища мають місце у приладах, що використовують структури типу МДМ (метал – діелектрик – метал), які можна розглядати як контакт двох металів з вакуумним проміжком, заміненим тонкою діелектричною плівкою. Вольт-амперна характеристика МДМ-структур відрізняється від (7.30), оскільки тут не врахована можливість тунельного проходження електронів через потенціальний бар'єр, а також сили дзеркального відображення, які суттєво змінюють форму бар'єру і викликають ефект Шоттки (розд. 3). Зазначимо, що ці структури не знайшли широкого вжитку в твердотільній електроніці.

Проте співвідношення (7.28) зберігає свою силу і є достатньо універсальним для інших видів контактів.

7.3. ІНЖЕКЦІЯ НЕРІВНОВАЖНИХ НОСІЇВ

У багатьох діодних та транзисторних структурах використовують так звані *інжекційні контакти*. Серед них є такі, що інjektують у напівпровідник неосновні носії, рух яких забезпечується процесом *дифузії нерівноважних носіїв* заряду. Покладемо, що на поверхні донорного напівпровідника якимось чином, наприклад, за рахунок інжекції ззовні, створюється надлишкова концентрація Δp_0 дірок, які є нерівноважними неосновними носіями для даного напівпровідника (рис. 7.9). Математична модель цього процесу може бути створена на основі рівняння неперервності (2.88). Для спрощення задачі будемо її вважати одновимірною ($\partial/\partial y = \partial/\partial z = 0$) та стаціонарною ($\partial/\partial t = 0$). Крім того вважатимемо, що процеси нерівноважної генерації відсутні ($g_p = 0$), рекомбінація дірок з рівноважними електронами пропорційна поточній надлишковій концентрації

$[r_p = (p - p_0) / \tau_p = \Delta p / \tau_p]$, а рівень легування напівпровідника достатньо високий і вважати його опір настільки малим, щоб знехтувати електричним полем ($\vec{\mathcal{E}} = 0$). Останнє припущення означає, що дрейфовий рух є несуттєвим відносно дифузійного. Обрані припущення суттєво спрощують рівняння неперервності

$$D_p \frac{d^2(\Delta p)}{dx^2} - \frac{\Delta p}{\tau_p} = 0, \quad (7.31)$$

яке можна отримати, якщо \vec{j}_p з (2.81) підставити в (2.88) та врахувати, що $\text{div grad } p = \nabla^2 p = d^2(\Delta p)/dx^2$. Розв'язок останнього рівняння, який задовольняє граничним умовам $\Delta p(0) = \Delta p_0$ та $\Delta p(\infty) = 0$, має вигляд

$$\Delta p(x) = \Delta p_0 \exp(-x / L_{D_p}), \quad (7.32)$$

де введено позначення $L_{D_p} \equiv \sqrt{D_p \tau_p}$.

Отриманий результат показує, що інжектовані дірки дифундують в об'єм і там рекомбінують з електронами, яких в донорному напівпровіднику є в достатній кількості. Величина L_{D_p} , яку називають *дифузійною довжиною*, відповідає відстані, на якій надлишкова концентрація спадає в $e \approx 2,71$ разів, і характеризує середню глибину, на яку дифундують дірки за час їх життя τ_p . Кількісно ця величина становить соті або десяті долі міліметра залежно від матеріалу та рівня його легування.

Рух дірок створює дифузійний струм (рис. 7.9), величина якого визначається з (2.81) і який спадає при віддаленні від поверхні:

$$j_p(x) = \frac{e D_p \Delta p_0}{L_{D_p}} \exp(-x / L_{D_p}). \quad (7.33)$$

Треба мати на увазі, що повний струм у колі створюється ще і електронами, які рухаються назустріч діркам і поповнюють втрати електронів через їх рекомбінацію з дірками. Для розрахунку цього струму треба розв'язати рівняння неперервності для електронів, але зручно знайти $j_n(x)$ як електронну складову, що доповнює діркову

до повного струму (рис. 7.9, в), з умови $j_n(x) + j_p(x) = j = \text{const}(x)$. Якщо вважати, що електрони не перетинають поверхню, то

$$j_n(x) = j - j_p(x) = j_p(0) - j_p(x) = \frac{eD_p \Delta p_0}{L_{Dp}} [1 - \exp(-x / L_{Dp})]. \quad (7.34)$$

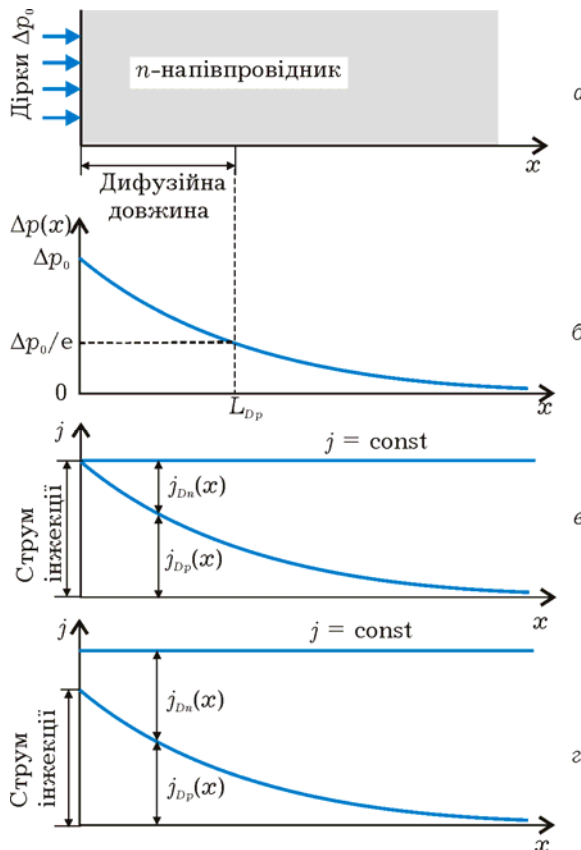


Рис. 7.9. Дифузія нерівноважних дірок в акцепторному напівпровіднику. Діаграми для струмів відповідають різним коефіцієнтам інжекції: рівному 1 (в) та меншому 1 (г).

її швидкість буде визначатися коефіцієнтом дифузії неосновних носіїв.

Зустрічаються ситуації, коли струм через поверхню створюється не тільки інжекованими дірками, але й електронами, що виходять. У такому випадку електронна складова матиме вигляд:

$$j_n(x) = j[1 - \gamma \exp(-x / L_{Dp})]. \quad (7.35)$$

де $\gamma = j_p(0) / j$ – коефіцієнт інжекції, що показує, яка частина повного струму створюється інжекованими носіями, у даному випадку дірками.

Подібне явище дифузії спостерігається, якщо нерівноважні носії створюються під час освітлення поверхні. Не зважаючи на те, що у цьому випадку в рівній кількості створюються і надлишкові електрони, їх дифузія від місця освітлення буде *амбіполярною*, а

7.4. ЕЛЕКТРОННО-ДІРКОВИЙ ПЕРЕХІД

Електронно-дірковий перехід виникає у місці контакту двох напівпровідників з різним типом провідності, енергетичні діаграми

яких наведені на рис. 7.10, а. Не розглядаючи різних технологічних способів його виготовлення, відзначимо лише, що такий перехід не може бути отримано механічним з'єднанням двох напівпровідників через сильний вплив на його властивості поверхневих станів пов'язаних з дефектами контактуючих поверхонь. Описаний у даному розділі перехід є єдиним монокристалом зі змінним профілем легування донорними та акцепторними домішками.

7.4.1. Рівноважний стан переходу

За відсутності будь-якого зовнішнього збудження та за постійної температури електронно-діркові колективи p - та n -областей напівпровідника знаходяться між собою у рівновазі, що характеризується єдиним рівнем Фермі (рис. 7.10, б) для обох частин. Це стає можливим у результаті обміну рухливими носіями зарядів та утворенню контактної різниці потенціалів.

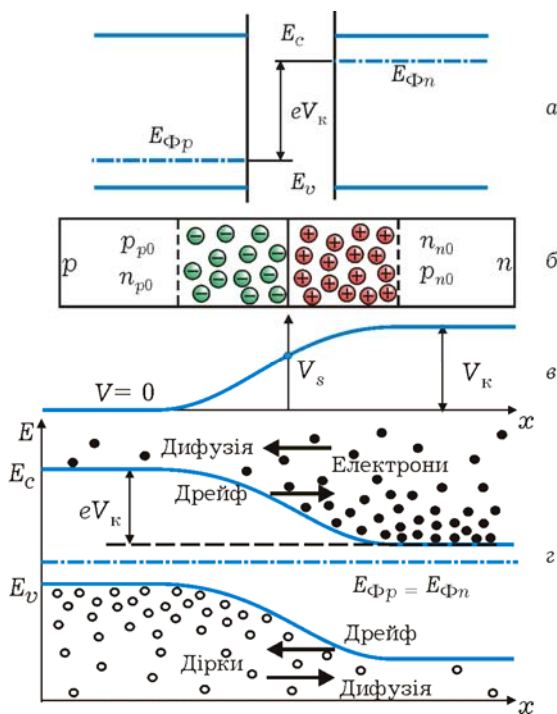


Рис. 7.10. Енергетичні діаграми вихідних акцепторного та донорного напівпровідникових матеріалів (а) та структура приконтактної області (б).

Контактне електричне поле (в) викривляє енергетичну діаграму (г), вирівнюючи дифузійні та дрейфові потоки як для електронів, так і для дірок.

Незалежно від конкретного виду функції розподілу потенціалу в області контакту значення контактної різниці потенціалу можна знайти через відстань рівнів Фермі (рис. 7.10, а) вихідних матеріалів, положення яких, у свою чергу, визначається співвідношеннями, отриманими у розділі 3.5. Однак для цього зручніше скористатися співвідношеннями (3.84), записаними для одновимірного випадку

$$\begin{aligned} n(x) &= n_0 \exp[eV(x) / k_B T], \\ p(x) &= p_0 \exp[-eV(x) / k_B T]. \end{aligned} \quad (7.36)$$

Приймаючи нульовим потенціал p -області, де концентрації основних (p_{p0}) і неосновних (n_{p0}) носіїв залишаються постійними, з (7.36) отримуємо, що $n_0 = n_{p0}$ і $p_0 = p_{p0}$. Крім того, записуючи їх для точок n -області, де відсутнє контактне поле, потенціал $V(x) = V_k$, а рівноважні концентрації дорівнюють $n(x) = n_{n0}$ і $p(x) = p_{n0}$, маємо:

$$\begin{aligned} n_{n0} &= n_{p0} \exp(eV_k / k_B T), \\ p_{n0} &= p_{p0} \exp(-eV_k / k_B T). \end{aligned} \quad (7.37)$$

Звідси можна отримати два рівноправних вирази для висоти контактного потенціального бар'єру

$$eV_k = k_B T \ln(n_{n0} / n_{p0}) = k_B T \ln(p_{p0} / p_{n0}). \quad (7.38)$$

Можна отримати й інші форми запису цього виразу використовуючи закон діючих мас $n_{n0} p_{n0} = n_{p0} p_{p0} = n_{i0}^2$. Результат показує, що незалежно від вигляду просторового розподілу легуючих домішок контактна різниця потенціалів визначається відношенням концентрацій однойменних носіїв в n - та p -областях.

Незважаючи на співмножник T в (7.38), залежність контактної різниці потенціалів є досить складною і визначається переважно температурним ходом співвідношення між концентраціями основних і неосновних носіїв. Проте її легко зрозуміти, якщо звернутись до температурних змін рівня Фермі (рис. 7.11) для донорних та акцепторних напівпровідників (див. розд. 3.5). Враховуючи, що висота контактного потенціального бар'єру дорівнює відстані між вихідними значеннями рівнів до контакту, бачимо, що вона спадає з підвищенням температури, і за кімнатної температури становить декілька десятих долів електрон-вольта. Також варто мати на увазі, що вона збільшується з підвищенням рівня легування.

Утворення контактної різниці потенціалів є результатом та необхідною умовою встановлення рівноважного стану. Подібний процес був описаний у розділі 2.5, де розглядалися особливості дифузії заряджених частинок. Важливо пам'ятати, що дана рівновага є динамічною, і її треба розглядати не як відсутність

струму через перехід, а як рівність двох зустрічних процесів і відповідних струмів: дрейфового та дифузійного. Дійсно, оскільки встановлюється контактне електричне поле, то обов'язковим є існування дрейфового руху носіїв заряду, а оскільки відмінність у концентраціях носіїв у p - і n -областях зберігається, то неминучим є протіканням дифузійних процесів, причому $\vec{j}_D = -\vec{j}_{др}$. Більш того, рівновага є *детальною*, що в даному випадку означає рівність дифузійних і дрейфових струмів окремо для електронів і для дірок (рис. 7.9, з):

$$\vec{j}_{Dn} + \vec{j}_{дрn} = 0, \quad \vec{j}_{Dp} + \vec{j}_{дрp} = 0, \quad (7.39)$$

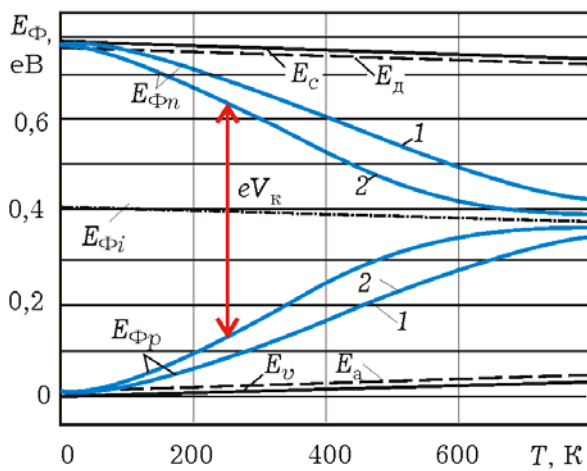


Рис. 7.11. Температурні залежності рівня Фермі для донорного та акцепторного германію ($\Delta E = 0,785$ eВ) з концентрацією домішок 10^{16} см^{-3} (1) і 10^{17} см^{-3} (2), взяті з рис. 3.12 та 3.14 та побудовані на тлі енергетичної діаграми з врахуванням температурної зміни ширини забороненої зони. Величина контактного бар'єру знаходиться як відстань між рівнями Фермі та спадає з підвищенням температури.

Розглянемо ці процеси більш детально. Неосновні носії p -області з концентрацією n_{p0} у своєму хаотичному русі попадають в область контактного поля, яке є для них прискорювальним, і вони переходять в n -область напівпровідника. Якщо в області переходу їх кількість не змінюється в результаті, наприклад, процесів генерації або рекомбінації, то величина струму, який вони створюють, визначається лише кількістю електронів, які, хаотично рухаючись з середньою тепловою швидкістю v_T , попадають на межу прискорювального поля і практично не залежить від напруженості електричного поля. Тому струм, створюваний неосновними носіями p -області, дорівнює:

$$j_{np} = en_{p0}v_T / 4 \quad (7.40)$$

і спрямований назустріч їх потоку, який вказано на рис. 7.10, з.

Основні носії n -напівпровідника мають концентрацію n_{n0} , на декілька порядків вищу за концентрацію неосновних носіїв p -

області, і будуть переходити з n - в p -область. Густина їх потоку за 1 секунду на границю контактної області, як і в попередньому випадку, визначатиметься хаотичним рухом і становитиме $n_{p0}v_T / 4$. Але ці електрони будуть рухатися у гальмівному полі контактної різниці потенціалів, і подолати його зможуть лише ті електрони, кінетична енергія яких більша за висоту потенціального бар'єру eV_K . Тому струм, який створюють основні носії n -області, зважаючи на їхній максвеллів розподіл, дорівнює:

$$j_{n_n} = e(n_{n0}v_T / 4) \exp(-eV_K / k_B T). \quad (7.41)$$

Беручи до уваги зв'язок (7.37) між концентраціями електронів, легко впевнитися, що $j_{n_n} = j_{n_p}$. А оскільки ці струми створюються зустрічними потоками, то результуючий електронний струм через перехід дорівнюватиме нулю.

Подібні явища спостерігаються і для дірок. Неосновні носії n -області у своєму хаотичному русі попадають у контактне електричне поле, яке прискорює їх, переносячи в p -область створюючи струм

$$j_{p_n} = ep_{n0}v_T / 4 \quad \text{точка} \quad (7.42)$$

Назустріч йому направлений струм, який створено лише тими дірками p -області, які можуть подолати потенціальний бар'єр eV_K :

$$j_{p_p} = e(p_{p0}v_T / 4) \exp(-eV_K / k_B T). \quad (7.43)$$

Як і для електронів, діркові струми знаходяться у динамічній рівновазі і взаємно компенсуються.

Застереження, зроблене при виведенні формули (7.40), фактично означає, що в області контактного поля не відбувається зіткнень електронів, і тому необхідною умовою їх переходу через контактний шар є попадання на межу цього шару. Ясно, що така ситуація можлива лише в тонкому електронно-дірковому переході, товщина якого набагато менша за довжину вільного пробігу. Рухливі носії проскакують тонкий перехід як вакуумний проміжок, не відчуваючи зіткнень. Тому розглянутий підхід відповідає так званій *діодній теорії* електронно-діркового переходу. Строго кажучи, у

такому разі терміни «дрейф» та «дифузія» не мають сенсу, бо вони передбачають процеси з масовими зіткненнями.

Рух носіїв через «товстий» перехід, де можливі багаторазові зіткнення, описується *дифузійною теорією*. Відповідно до неї, струм через перехід має дві складові – дифузійну та дрейфову, які описуються співвідношеннями (2.81). Оскільки розподіл рівноважної концентрації електронів у переході є функцією координати (7.31), то існує електронна складова дифузійного струму, пропорційна градієнту концентрації:

$$j_{Dn}(x) = eD_n \frac{dn}{dx} = en_{p0} \left(\frac{eD_n}{k_B T} \right) \frac{dV}{dx} \exp[eV(x) / k_B T]. \quad (7.44)$$

Дрейфова складова електронного струму пропорційна концентрації $n(x)$ і напруженості електричного поля $\mathcal{E}(x) = -dV / dx$ в області переходу:

$$j_{дрn}(x) = en(x)\mu_n \mathcal{E}(x) = -en_{p0}\mu_n \frac{dV}{dx} \exp[eV(x) / k_B T]. \quad (7.45)$$

Беручи до уваги співвідношення Ейнштейна (2.84) між рухливістю і коефіцієнтом дифузії, переконуємося у динамічній рівновазі між дифузійними та дрейфовими електронними процесами.

Аналогічно, враховуючи розподіл (7.36) для діркової концентрації і співвідношення Ейнштейна (2.84) для дірок, можна переконатися, що дрейфовий струм

$$j_{дрp}(x) = ep(x)\mu_p \mathcal{E}(x) = -ep_{p0}\mu_p \frac{dV}{dx} \exp[-eV(x) / k_B T]. \quad (7.46)$$

і дифузійний струм

$$j_{Dp}(x) = -eD_p \frac{dp}{dx} = ep_{p0} \left(\frac{eD_p}{k_B T} \right) \frac{dV}{dx} \exp[-eV(x) / k_B T]. \quad (7.47)$$

рівні за величиною і протилежні за напрямом.

7.4.2. Товщина електронно-діркового переходу

Розглянемо *різкий*, або *ступінчастий*, перехід, у якому концентрація донорів залишається постійною і рівною N_d у всій n -

області до місця контакту (точка $x = 0$ на рис. 7.12). Аналогічно і p -область ($x < 0$) має однорідний розподіл акцепторної домішки з концентрацією N_a . Для несиметричного переходу $N_d \neq N_a$ і припустимо, що $N_d < N_a$.

При встановленні контактної різниці потенціалів V_k (рис.7.12, з) на межі між напівпровідниками буде існувати деякий потенціал V_s , додатний для об'єму p -області ($V_s > 0$) і від'ємний для n -області. Його значення (порядку десятих долів вольт за кімнатних температур та концентрацій N_d і N_a , що не дуже відрізняються одна від одної) достатнє для того, щоб викликати збіднення приконтактних шарів основними носіями (див. розд. 7.1). У результаті біля границі p -області в шарі товщиною d_p буде існувати від'ємний просторовий заряд іонізованих акцепторних атомів з постійною об'ємною густиною

При встановленні контактної різниці потенціалів V_k (рис.7.12, з) на межі між напівпровідниками буде існувати деякий потенціал V_s , додатний для об'єму p -області ($V_s > 0$) і від'ємний для n -області. Його значення (порядку десятих часток вольт за кімнатних температур та концентрацій N_d і N_a , що не дуже відрізняються одна від одної) достатнє для того, щоб викликати збіднення приконтактних шарів основними носіями (див. розд. 7.1). У результаті біля границі p -області в шарі товщиною d_p буде існувати від'ємний просторовий заряд іонізованих акцепторних атомів з постійною об'ємною густиною

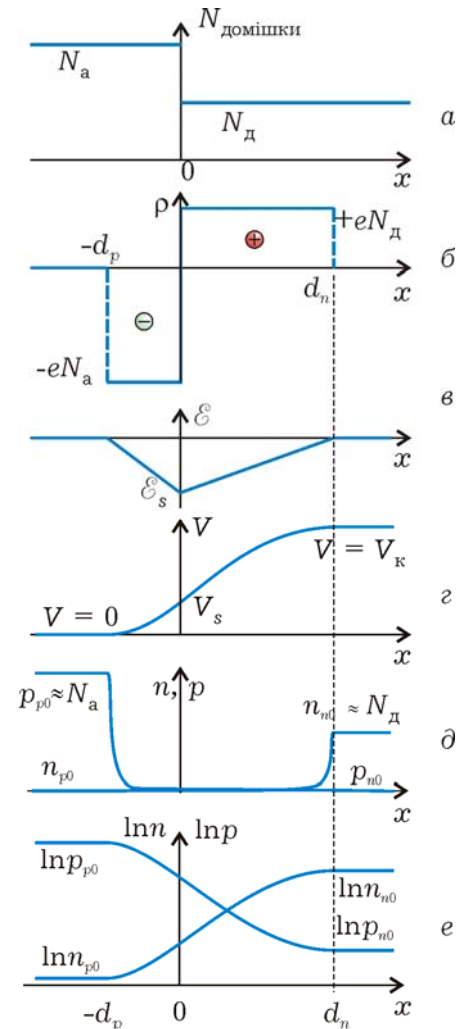


Рис. 7.12. Розподіл концентрації домішки (а), густини просторового заряду (б), напруженості електричного поля (в), потенціалу (г) та концентрації носіїв у лінійному (д) та логарифмічному (е) масштабах

збіднення приконтактних шарів основними носіями (див. розд. 7.1). У результаті біля границі p -області в шарі товщиною d_p буде існувати від'ємний просторовий заряд іонізованих акцепторних атомів з постійною об'ємною густиною

$$\rho(x) = -eN_a^- \approx -eN_a = \text{const}(x).$$

Аналогічно в шарі n -області товщиною d_n - додатний просторовий заряд іонізованих донорів з густиною $\rho(x) = -eN_d^+ \approx -eN_d = \text{const}(x)$ (рис. 7.12, б). Для знаходження розподілу потенціалу $V(x)$, тобто і розподілу концентрації носіїв, а також товщини збіднених шарів необхідно розв'язати рівняння Пуассона із заданою густиною просторового заряду $\rho(x)$. Але для n -області воно тотожно (7.16), тому зручно скористатися готовими результатами (7.17) та (7.18). Необхідно тільки замінити початок відліку потенціалу на V_κ , використовуючи вираз для товщини збідненої області у вигляді

$$d_n = \sqrt{2\varepsilon\varepsilon_0 |V_s - V_\kappa| / eN_d}. \quad (7.48)$$

Відповідно до рівняння (7.43) перша похідна визначатиме лінійний розподіл напруженості електричного поля в n -області

$$\mathcal{E}(x) = -\frac{dV}{dx} = -\frac{eN_d}{\varepsilon\varepsilon_0} (d_n - x), \quad (7.49)$$

а потенціал є параболічною функцією (рис. 7.12, з)

$$V(x) = V_\kappa - \frac{eN_d}{2\varepsilon\varepsilon_0} (d_n - x)^2. \quad (7.50)$$

Для p -області ($x < 0$ на рис. 7.12) подібне рівняння Пуассона матиме аналогічні розв'язки:

$$\mathcal{E}(x) = -\frac{dV}{dx} = -\frac{eN_a}{\varepsilon\varepsilon_0} (d_p + x), \quad (7.51)$$

$$V(x) = \frac{eN_a}{2\varepsilon\varepsilon_0} (d_p + x)^2. \quad (7.52)$$

Оскільки потенціальна функція має бути неперервною і плавною ці вирази необхідно «зшити». Прирівнюючи напруженості поля в точці $x = 0$, отримаємо співвідношення

$$N_d d_n = N_a d_p, \quad (7.53)$$

яке показує, що товщина збідненої області обернено пропорційна концентрації домішки, а сумарний просторовий заряд збіднених областей дорівнює 0 (площі прямокутників на рис. 7.12, б однакові). З рівності потенціалів на межі слідує, що

$$N_d d_n^2 + N_a d_p^2 = 2\varepsilon\varepsilon_0 V_k / e. \quad (7.54)$$

Останні два вирази дозволяють визначити значення товщини збіднених областей:

$$d_n = \sqrt{\frac{2\varepsilon\varepsilon_0 V_k N_a}{e N_d (N_a + N_d)}}; \quad d_p = \sqrt{\frac{2\varepsilon\varepsilon_0 V_k N_d}{e N_a (N_d + N_a)}} \quad (7.55)$$

а також повну товщину електронно-діркового переходу

$$d_{pn} = d_n + d_p = \sqrt{\frac{2\varepsilon\varepsilon_0 V_k}{e} \cdot \frac{N_a + N_d}{N_a N_d}} \quad (7.56)$$

У несиметричному переході, що розглядається, $N_a > N_d$, тому $d_n > d_p$, а якщо $N_a \gg N_d$, то $d_{pn} \approx d_n$ і електронно-дірковий перехід (збіднена область) практично повністю знаходиться в області донорного напівпровідника. У типових випадках товщина переходу має порядок одиниць і десятих частин мікрометра.

Розподіл концентрацій рухливих носіїв у збідненому шарі подано на рис. 7.12, δ і описується виразами (7.36), де p_0 і n_0 відповідають концентраціям основних (p_{p0}) і неосновних (n_{p0}) носіїв p -області, потенціал якої прийнято рівним φ . У логарифмічному масштабі (рис. 7.12, e) їх розподіл пропорційний функції $V(x)$, у чому можна впевнитися логарифмуванням виразів (7.36). Для правильного розуміння останніх двох графіків треба зважити на те, що зміна потенціалу на десяті долі вольтів приводить до зміни концентрації на порядки.

7.4.3. Вольт-амперна характеристика переходу

Порушення рівноваги. Підключення до електронно-діркового переходу напруги V від зовнішнього джерела порушує рівновагу між дрейфовими та дифузійними складовими струму, а їх різниця визначає величину та напрям результуючого струму, що

протікає через перехід. Оскільки збіднена область має великий (порівняно з іншими ділянками кола) опір, наближено можна вважати, що вся зовнішня напруга прикладається безпосередньо до переходу. Електричне поле, створюване зовнішнім джерелом, додається до поля контактної різниці потенціалів і в залежності від полярності зовнішньої напруги зменшує (*пряме* включення) результуючу висоту потенціального бар'єру або збільшує її (*зворотне* включення).

Якісно картину явищ у разі зміни зовнішньої напруги можна уявити наступним чином. Пряме включення, зменшуючи висоту потенціального бар'єру, послаблює електричне поле, яке заважає переходу електронів з *n*-області в *p*-область і зустрічному переходу дірок. У результаті збільшується кількість основних носіїв, здатних подолати потенціальний бар'єр, і зростає створюваний ними струм. Цей струм за напрямом співпадає з зовнішнім полем і зветься *прямим*. *Зворотний* струм, який створюється неосновними носіями, які «скочуються» з потенціального бар'єру, несуттєво зменшується з послабленням поля.

Зворотне включення зовнішньої напруги підвищує потенціальний бар'єр і стримує потік основних носіїв, зменшуючи прямий струм. Зворотний струм неосновних носіїв, що прискорюються електричним полем, дещо збільшується із зростанням зворотної напруги. Якісна залежність прямого і зворотного струмів подана на рис. 7.13 штриховими лініями. Результуючий струм через перехід дорівнює різниці прямих і зворотних струмів, і вигляд вольт-амперної характеристики визначатиметься ступенем порушення рівноваги між потоками основних та неосновних носіїв.

Зміна концентрацій носіїв у збідненому шарі.

Кількісний розрахунок вольт-амперної характеристики проведемо для різкого переходу. Безпосереднім інтегруванням рівняння Пуассона можна впевнитися, що розподіл потенціалу, напруженості поля та густини просторового заряду, а також товщина збідненої області описуються такими ж виразами, як і в рівноважному стані, якщо в них замінити V_k на $V_k - V$, вважаючи $V > 0$ для прямого підключення та $V < 0$ – для зворотного. Але треба мати на увазі, що пряма напруга не може перевищувати контактну різницю потенціалів, зі збереженням режиму збіднення у приконтатному шарі.

Пряма напруга зменшує різницю потенціалів на межах збідненої області і зменшує її товщину (рис. 7.14, а, ліворуч). На енергетичній діаграмі переходу (б) це виглядає перш за все як зниження потенціального бар'єру від рівноважного значення eV_k (штрихові лінії) до значення $e(V_k - V)$.

Порушення рівноважного стану викликає і зміщення квазірівнів Фермі на величину eV . Для тонкого переходу зміна квазірівнів E_{Φ_n} та E_{Φ_p} у межах збідненого шару несуттєва, тому електронний квазірівень буде практично таким самим, як і в n -області, а дірковий – таким, як і в p -області. Ця особливість дозволяє записати вирази для розподілу концентрацій носіїв у збідненому шарі у формі (7.36), з яких встановлюються співвідношення між концентраціями носіїв на межах збідненого шару (7.37), які подамо у вигляді

$$\begin{aligned} n_{p0} &= n_{n0} \exp(-eV_k / k_B T), \\ p_{n0} &= p_{p0} \exp(-eV_k / k_B T). \end{aligned} \quad (7.57)$$

Напруга на переході змінює розподіл носіїв так, що концентрації основних носіїв залишаються такими саме, а неосновних – відповідно до заміни V_k на $V_k - V$:

$$\begin{aligned} n_p &= n_{n0} \exp[-e(V_k - V) / k_B T] = n_{p0} \exp(eV / k_B T), \\ p_n &= p_{p0} \exp[-e(V_k - V) / k_B T] = p_{n0} \exp(eV / k_B T). \end{aligned} \quad (7.58)$$

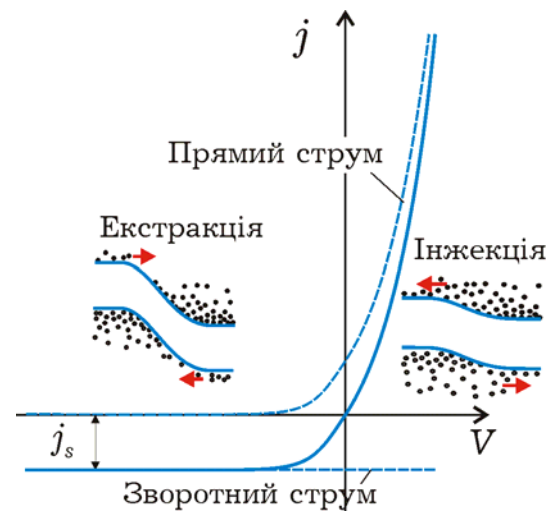


Рис. 7.13. Загальний вигляд вольт-амперної характеристики електронно-діркового переходу як результат суперпозиції прямого та зворотного струмів.

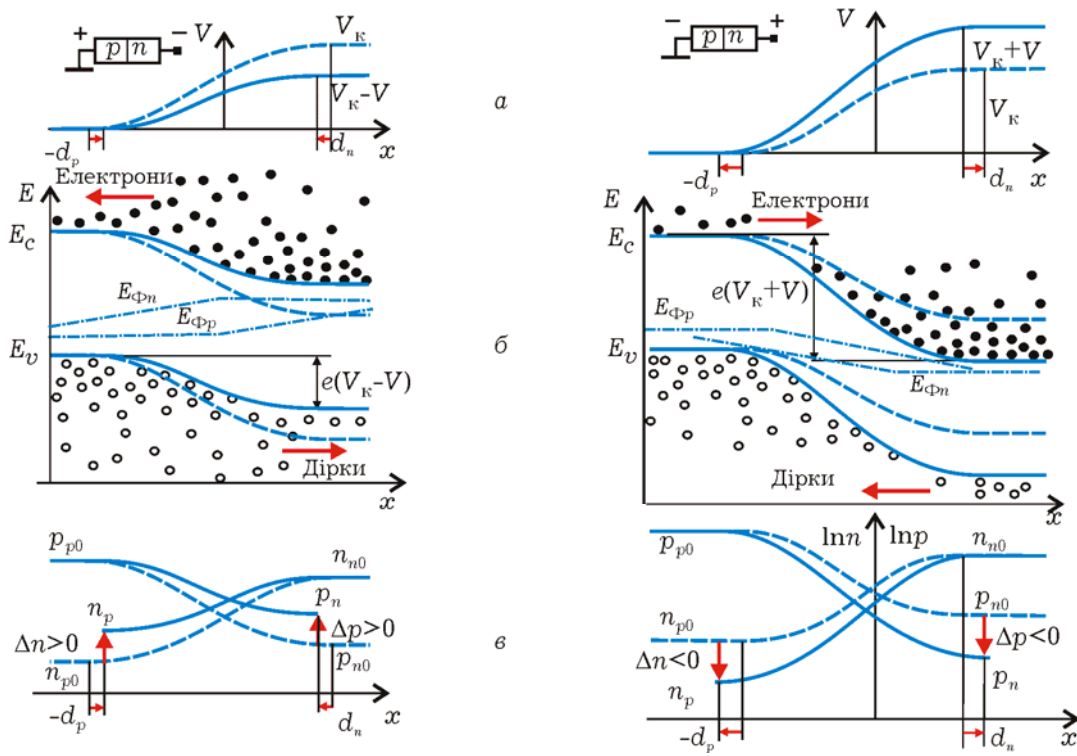


Рис. 7.14. Зміна розподілу потенціалу (а), енергетичної діаграми (б) і концентрації носіїв (в) для прямого (ліворуч) та зворотного (праворуч) включення електронно-діркового переходу. Штрихові лінії відповідають рівноважному стану.

Останні вирази показують, що в результаті дії прямої напруги і зниження потенціального бар'єру на межі *p*-області і збідненого шару (точка $x = -d_p$ на рис. 7.14, в) збільшується кількість електронів з надлишковою концентрацією

$$\Delta n(-d_p) = n_p - n_{n0} = n_{p0}[\exp(eV / k_B T) - 1], \tag{7.59}$$

а на межі збідненого шару з *n*-областю ($x = d_n$) підвищується концентрація дірок на величину

$$\Delta p(d_n) = p_n - p_{n0} = p_{n0}[\exp(eV / k_B T) - 1]. \tag{7.60}$$

Це явище називають *інжекцією*. Описані явища висвітлюють процеси регулювання потоків носіїв, які відбуваються в електричному полі при контактній області. Проте для повноти розуміння особливостей процесів протікання струму та інших властивостей електронно-діркового переходу та приладів на його

основі необхідно розглянути, що відбувається за межами власне електронно-діркового переходу.

Поведінка носіїв за межами збідненого шару.

Інжектовані носії є неосновними для тих областей, куди відбувається інжекція. Їх поведінка була розглянута в модельній задачі розд. 7.3 на прикладі інжекції дірок в n -напівпровідник. Дифундуючи вглиб напівпровідника, інжектовані носії рекомбінують з наявними там основними носіями. Просторова швидкість спадання концентрації і відповідного струму інжекції визначається дифузійними довжинами електронів $L_{D_n} \equiv \sqrt{D_n \tau_n}$ і дірок $L_{D_p} \equiv \sqrt{D_p \tau_p}$, причому на глибині у декілька дифузійних довжин струм інжекції неосновних носіїв практично повністю переходить у дрейфовий струм основних носіїв, які підходять з глибини напівпровідника для поповнення втрат рухливих носіїв через рекомбінацію.

Функції розподілу концентрацій інжектованих носіїв, а також струмів основних і неосновних носіїв можуть бути отримані на основі виразів (7.32) – (7.34) Зважаючи на напрям і початок відліку для відповідних областей електронно-діркового переходу (рис. 7.15, ліворуч) вони приймуть вигляд:

$$\begin{array}{ll} p\text{-область } (x < -d_p) & n\text{-область } (x > d_n) \\ \Delta n(x) = \Delta n(-d_p) \exp\left(-\frac{x + d_p}{L_{D_p}}\right); & \Delta p(x) = \Delta p(d_n) \exp\left(-\frac{x - d_n}{L_{D_p}}\right); \end{array} \quad (7.61)$$

$$j_{D_n}(x) = \frac{eD_n \Delta n(-d_p)}{L_{D_n}} \exp\left(-\frac{x + d_p}{L_{D_n}}\right); \quad j_{D_p}(x) = \frac{eD_p \Delta p(d_n)}{L_{D_p}} \exp\left(-\frac{x - d_n}{L_{D_p}}\right); \quad (7.62)$$

$$\begin{array}{ll} j_{дрn}(x) = \frac{eD_n \Delta n(-d_p)}{L_{D_n}} \times & j_{дрp}(x) = \frac{eD_p \Delta p(d_n)}{L_{D_p}} \times \\ \times \left[1 - \exp\left(-\frac{x + d_p}{L_{D_n}}\right) \right]; & \times \left[1 - \exp\left(-\frac{x - d_n}{L_{D_p}}\right) \right]. \end{array} \quad (7.63)$$

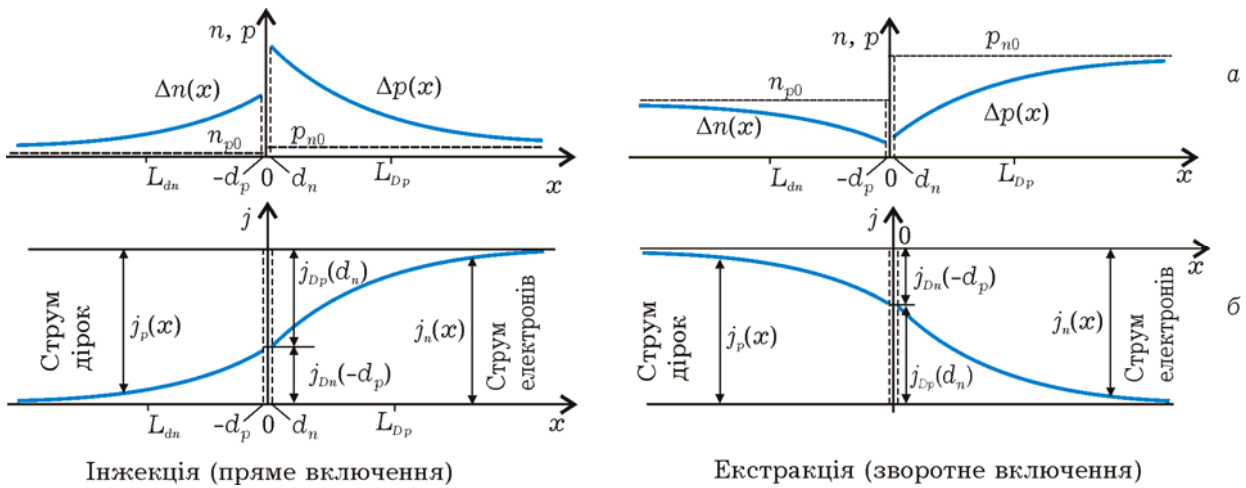


Рис. 7.15. Розподіл концентрації неосновних носіїв (а) і струмів (б) за межами збіднених шарів у разі інжекції (ліворуч) та екстракції (праворуч). Слід звернути увагу на різні масштаби та напрями струмів

Графіки просторового розподілу струмів, наведені на рис. 7.15, зображено так, що один струм доповнює інший до повного. Зручність такого подання у наочності процесу зміни електронної складової струму в n -напівпровіднику на дірковий струм у p -напівпровіднику. Сума дрейфових та дифузійних струмів, викликаних інжекцією, у кожній області залишається постійною і рівною струму інжекції через відповідну межу збідненої області:

$$j^n(x) = j_{Dn}(-d_p) = \frac{eD_n \Delta n(-d_p)}{L_{D_n}}; \quad j^p(x) = j_{Dp}(d_n) = \frac{eD_p \Delta p(d_n)}{L_{D_p}}. \quad (7.64)$$

Якщо електронно-дірковий перехід тонкий, тобто товщина збіднених областей багато менша за дифузійні довжини, що типово для багатьох pn -структур, носії проходять область контактного поля не рекомбінуючи. У такому випадку струми, що проходять через площину контакту, визначатимуться лише струмами інжекції. Тому повний струм через перехід і, відповідно, у всьому колі дорівнюватиме сумі виразів (7.64):

$$j = j^n + j^p = \frac{eD_n \Delta n(-d_p)}{L_{D_n}} + \frac{eD_p \Delta p(d_n)}{L_{D_p}} = \text{const}(x). \quad (7.65)$$

Підключення до переходу зворотної напруги, поле якої за напрямом співпадає з контактним, підвищує сумарну різницю

потенціалів на межах збідненої області і збільшує її товщину відповідно до (7.56)

$$d_{pn} = \sqrt{\frac{2\epsilon\epsilon_0(V_k + V)}{e} \cdot \frac{N_a + N_d}{N_a N_d}} \quad (7.66)$$

На енергетичній діаграмі (рис. 7.14, б, праворуч) це викликає підвищення потенціального бар'єру до значення $e(V_k + V)$. та зміщення квазірівнів Фермі на величину $e|V|$, як вказано на рис. 7.14, б, причому $E_{\Phi p} > E_{\Phi n}$. Зміна квазірівнів Фермі вздовж переходу така, що і в цьому випадку можна формально використовувати співвідношення (3.84). Зв'язок між концентраціями рухливих носіїв на межах збідненого шару можна описати тими самими виразами (7.58), що і для прямого включення, вважаючи в них V від'ємним. Те саме відноситься і до виразів (7.59) та (7.60) для надлишкових концентрацій. Однак при зворотному включенні величини $\Delta n(-d_p)$ і $\Delta p(d_n)$ є від'ємними, що означає зменшення концентрацій неосновних носіїв на межах p - та n -областей і зветься *екстракцією*. Всі кількісні співвідношення від (7.56) до (7.65) при екстракції можна залишити в силі, якщо змінити знак перед V . Але треба мати на увазі, що зміняться знаки і перед струмами, тому дифузійні потоки неосновних носіїв будуть направлені не вглиб напівпровідника, а до межі збідненої області (див. рис. 7.14, в, праворуч).

У результаті вольт-амперна характеристика може бути описана виразом (7.65), в який треба підставити (7.59) та (7.60). Формально її вигляд співпадатиме з (7.30) для контакту двох металів (рис. 7.8):

$$j = j_s [\exp(eV / k_B T) - 1], \quad (7.67)$$

але значення струму насичення виражається інакше:

$$j_s = \frac{eD_n n_{p0}}{L_{D_n}} + \frac{eD_p p_{n0}}{L_{D_p}}. \quad (7.68)$$

Зручним може бути подання (7.68) в іншій формі з урахуванням (7.37) й того, що за кімнатних температур зазвичай має місце виснаження домішки і $n_{n0} = N_d$ та $p_{p0} = N_a$

$$j_s = \left(\frac{eD_n N_d}{L_{D_n}} + \frac{eD_p N_a}{L_{D_p}} \right) \exp\left(-\frac{eV_k}{k_B T}\right) \equiv j_{s0} \exp\left(-\frac{eV_k}{k_B T}\right). \quad (7.69)$$

Останній вираз у явному вигляді показує, що величина струму насичення залежить від контактної різниці потенціалів і зростає з підвищенням температури досить інтенсивно (рис. 7.16).

Вираз (7.67) описує вольт-амперну характеристику ідеального електронно-діркового переходу. За великих зворотних напругах у тонкому збідненому шарі напруженість електричного поля стає настільки сильною, що можливе виникнення *електростатичного* (зінерівського) *пробу* (див розд 4.2), який супроводжується різким зростанням зворотного струму. Деяко подібне явище можливе і в разі *теплового пробую*, коли саморозігрівання струмом, що протікає, викликає інтенсивну генерацію неосновних носіїв і відповідне збільшення зворотного струму.

Вираз (7.67) є непридатним також для великих прямих струмів. Якщо зовнішня напруга повністю компенсує контактну різницю потенціалів, потенціальний бар'єр зникає, і вольт-амперна характеристика стає близькою до лінійної (омічної).

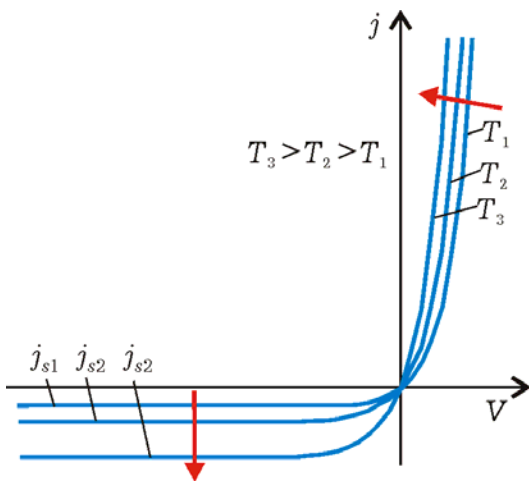


Рис. 7.16. Вплив температури на струм насичення та вигляд вольт-амперної характеристики.

У сильно несиметричному переході (наприклад, $N_a \gg N_d$) через те, що концентрації основних носіїв сильно різняться, величина прямого струму визначатиметься практично тільки струмом інжекції дірок з *p*- в *n*-область. Зазвичай високолегований інжектуючий шар у такому разі називають *емітером*, а шар з малою концентрацією – *базою*. Вольт-амперна характеристика такого переходу визначається переважно параметрами (коефіцієнт дифузії, концентрація, дифузійна довжина, час життя) носіїв одного виду, у нашому прикладі – дірок.

У товстих електронно-діркових переходах суттєвими можуть виявитися і генераційно-рекомбінаційні процеси у збідненій області, які під час виведення (7.67) не враховувались. Зворотна напруга підсилює збіднення контактного шару рухливими носіями,

порушуючи рівновагу між процесами генерації та рекомбінації на користь перших. Генеровані носії створюють зворотний струм, який залежить від величини зворотної напруги, збільшуючись з її зростанням. У широкозонних напівпровідниках (наприклад, у кремнії) більше значення контактної різниці потенціалів забезпечує дуже малий струм насичення (7.69). Тому струм генерації у контактному шарі може значно перевищувати j_s і вольт-амперна характеристика кремнієвих переходів не має режиму насичення за зворотної напруги. Пряма напруга підвищує концентрацію носіїв у переході і пришвидшує їх рекомбінацію. Однак додатковий рекомбінаційний струм є помітним лише за невеликих прямих струмах.

7.4.4. Ємності електронно-діркового переходу

Бар'єрна ємність. Контактна різниця потенціалів і зовнішня напруга викликають збіднення приконтактного шару рухливими носіями і утворення не скомпенсованого просторового заряду іонізованих домішкових атомів. Наявність електричного поля у подвійному електричному шарі переходу спонукає до аналогії з конденсатором, на «пластинах» якого з площею контакту S зосереджено додатний заряд $Q^+ = eSN_d d_n$ в n -області та від'ємний $Q^- = -eSN_a d_p$ в p -області. Звісно, в силу електричної нейтральності ці заряди за абсолютною величиною однакові. Для характеристики динамічних властивостей електронно-діркового переходу визначимо його ємність через реакцію заряду на зміну напруги, тобто через похідну $C_0 = d|Q^\pm| / dV$. Зважаючи на вирази (7.55) для товщини окремих шарів, у яких за наявності прикладеної напруги треба V_k замінити на $V_k - V$, після обчислення похідної та простих перетворень отримаємо

$$C_0 = S \sqrt{\frac{2\epsilon\epsilon_0 e}{(V_k - V)} \cdot \frac{N_d N_a}{N_a + N_d}} \quad (7.70)$$

Цікаво, що, беручи до уваги вираз (7.56) для повної товщини переходу, останній вираз можна подати у вигляді класичної формули для ємності плоского конденсатора $C_0 = \epsilon\epsilon_0 S / d_{pn}$.

Отриманий вираз характеризує ємність, яку називають *бар'єрною* або *зарядовою*. Залежність її від прикладеної напруги використовують для створення *варіканів*, які застосовують в електронних колах як ємність, керовану напругою, наприклад, у коливальному контурі для «електронного» перестроювання його резонансної частоти. Зазвичай для цього перехід включають у зворотному напрямку, знак зовнішньої напруги в (7.70) – від'ємний. Для прямого включення, коли прикладена напруга наближається до контактної, поняття різкої межі збідненої області і, відповідно, поняття бар'єрної ємності втрачає сенс.

Дифузійна ємність. Для прямого включення характерний інший прояв інерційності електронно-діркового переходу, пов'язаний з явищем інжекції. Не зважаючи на те, що інжектвані носії не створюють не скомпенсованого просторового заряду, дифузійні процеси забезпечують ємнісний характер перехідних процесів протікання струму. Дійсно, у разі миттєвого підключення до переходу прямої напруги (рис. 7.43) через перехід протікає дуже великий, обмежений лише опором зовнішнього кола, дифузійний струм. Він виникає за рахунок великого градієнта концентрації неосновних носіїв, які інжектуються через знижений потенціальний бар'єр і в момент часу t_1 поки що сконцентровані на межі збідненого шару. З часом вони дифундують в об'єм, градієнт концентрації зменшується і, відповідно, зменшується дифузійний струм. Це триває доти, поки не встановиться стаціонарний стан (t_3), коли розподіл носіїв і струмів відповідатимуть співвідношенням (7.61) – (7.63). Зображений на рис. 7.17 зв'язок між струмом і напругою є характерним для перехідних процесів за участі паралельного з'єднання ємнісних і резистивних компонентів. Тому стосовно до змін напруги інерційність дифузійного руху інжектованих носіїв можна змодельовати за допомогою ємності, яку називають *дифузійною*.

Величину дифузійної ємності C_d можна знайти, розрахувавши заряд інжектованих носіїв. Обмежуючись різко несиметричним переходом (наприклад, $N_a \gg N_d$) і використовуючи (7.61) з правої колонки, отримаємо

$$Q = eS \int_{d_n}^{\infty} \Delta p(x) dx = e\Delta p(d_n)L_{Dp}. \quad (7.71)$$

Беручи до уваги залежність (7.60) для $\Delta p(d_n)$ від прикладеної

напруги та позначення деяких величин, знайдемо дифузійну ємність через похідну

$$C_d = \frac{dQ}{dV} = \frac{Se^2 p_{n0} L_{Dp}}{k_B T} \exp\left(\frac{eV}{k_B T}\right) = \frac{eS}{k_B T} (j + j_s) \tau_p. \quad (7.72)$$

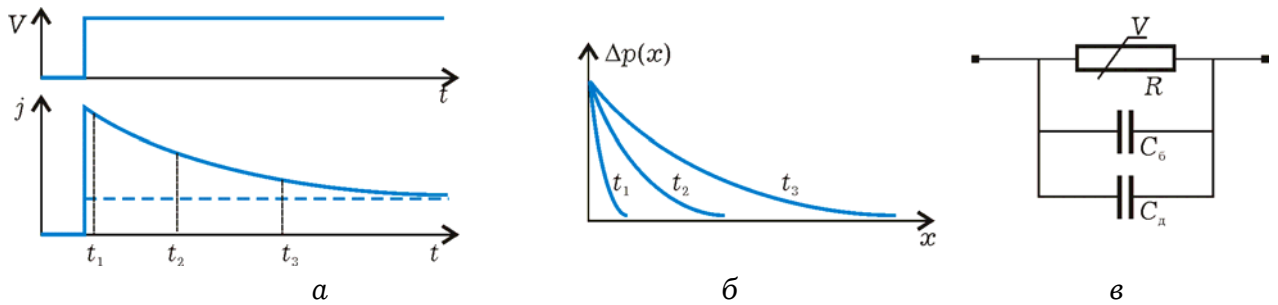


Рис. 7.17. Перехідні процеси зміни струму (а) у часі після включення напруги, просторовий розподіл нерівноважної концентрації дірок (б) у різні моменти часу та схемна модель електронно-діркового переходу (в) у вигляді паралельного з'єднання нелінійного і ємнісних компонентів.

Аналіз отриманих виразів показує, що дифузійна ємність є переважаючою у разі протікання прямого струму. За зворотної напруги, коли струм досягає режиму насичення, дифузійна ємність практично зникає і проявляється дія бар'єрної. Ці ємності суттєво обмежують використання електронно-діркових переходів на високих частотах. Гранична частота значною мірою визначається часом встановлення процесів дифузії і рекомбінації інжектованих носіїв, тобто часом їх життя. Більш високочастотними є асиметричні переходи з більш високим ступенем легування n -області (емітер), в яких переважає інжекція електронів у p -область (базу). Тому встановлення перехідних процесів буде відбуватися швидше, бо електрони, маючи зазвичай меншу ефективну масу, мають і менший час життя.

7.5. КОНТАКТИ МЕТАЛ – НАПІВПРОВІДНИК

Серед великої кількості можливих різноманітних пар контактів металів з напівпровідниками за функціональними властивостями можна виділити два типи. По-перше, це контакт з яскраво вираженою нелінійністю, вольт-амперна характеристика якого має типовий діодний характер. Це стосується перш за все випрямного контакту метал-напівпровідник, який називають бар'єром Шоттки.

По-друге, це так звані омичні контакти, вольт-амперна характеристика яких близька до лінійної, маючи малий опір.

Бар'єр Шоттки утворюється біля межі розділу металу з донорним напівпровідником (рис. 7.18, а), якщо робота виходу електронів з металу (Φ_m) більша, ніж з напівпровідника (Φ_n). У такому разі рівноважний стан встановлюється за контактної різниці потенціалів $V_k = (\Phi_m - \Phi_n) / e$, яка виникає у результаті переходу електронів з напівпровідника до металу. Їх просторовий заряд зосереджується практично на поверхні металу. Контактне поле, спрямоване, як показано на рис. 7.18, б, практично не буде проникати в метал, а локалізується у деякому поверхневому шарі напівпровідника. Від'ємному відносно глибини напівпровідника потенціалу поверхні (рис. 7.18, в) відповідає режим збіднення, подібний до того, який розглядався в ефекті поля (підрозд. 7.1). Товщина цього шару знаходиться з розв'язку рівняння Пуассона (7.163) і визначається контактною різницею потенціалів $d = \sqrt{2\epsilon\epsilon_0 |V_k| / eN_d}$. Аналогічний результат отримано в (7.74) з відповідною заміною потенціалу як товщину n -шару електронно-діркового переходу. Просторовий заряд у збідненому шарі створюється практично тільки іонізованими донорними атомами, іноді цей шар звать *запірним*. Саме цей шар з підвищеним опором забезпечує нелінійність бар'єру Шоттки, підтверджуючи правило: асиметрія конструкції викликає асиметрію вольт-амперної характеристики.

На рис. 7.18, г видно, що в зоні провідності для електронів напівпровідника утворюється потенціальний бар'єр висотою eV_k , а для електронів металу висота бар'єру становить $(\Phi_m - \chi)$. Саме ці бар'єри вирівнюють зустрічні потоки електронів, забезпечуючи рівноважний стан. З підключенням зовнішньої напруги у прямому напрямку поле джерела послаблює контактне поле у збідненому шарі і зменшує висоту потенціального бар'єру з боку напівпровідника (рис. 7.18, д). Електрони, які є основними носіями в n -напівпровіднику, переходячи до металу, створюють прямий струм, тим більший, чим нижче потенціальний бар'єр. У той же час, поле зовнішнього джерела, не проникаючи в метал, не змінює висоту бар'єру для електронів металу, їх струм, механізм якого можна розглядати як термоелектронну емісію через бар'єр висотою $(\Phi_m - \chi)$ залишається практично незмінним.

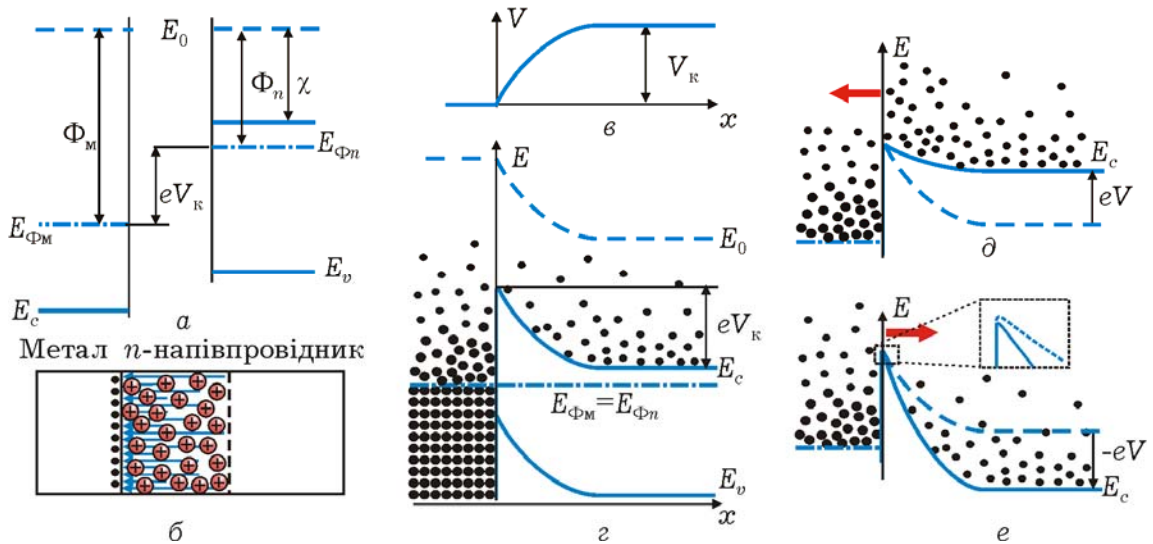


Рис. 7.18. Енергетичні діаграми метала та донорного напівпровідника до контакту (а) і шар просторового заряду іонізованих донорів (б) у разі $\Phi_M > \Phi_n$. Розподіл потенціалу (в) та енергетична діаграма електронів у контакті (г) в рівноважному стані. Зміна форми бар'єру для прямого (д) та зворотного підключення (е) зовнішньої напруги.

Підвищення потенціального бар'єру зворотною напругою (рис. 7.18, е) зменшує прямий струм, і через контакт буде протікати переважно зворотний струм, викликаний рухом термічно збуджених електронів металу в прискорювальному полі контактного шару. Таким чином, у створенні струму приймають участь лише електрони, і надбар'єрний перехід їх з напівпровідника до металу подібний до «інжекції» у електронно-дірковому переході, але він не супроводжується явищами дифузії та рекомбінації. «Інжектований» заряд практично миттєво розподіляється в об'ємі, створюючи дрейфовий струм. Саме відсутність відносно повільних процесів рекомбінації у контакті метал – напівпровідник, або на мові схемних моделей – відсутність дифузійної ємності, забезпечує його більшу швидкодю. Діоди, створені на основі такого контакту (діоди Шоттки), можуть працювати на частотах порядку 100 ГГц.

Вольт-амперна характеристика контакту Шоттки формально має практично такий самий вигляд, як і для електронно-діркового переходу. Однак за фізичною природою зворотний струм подібний до термоелектронної емісії і на його значення суттєво впливає форма потенціального бар'єру біля поверхні металу. Сили дзеркального відображення згладжують бар'єр, і зовнішня напруга зменшує роботу виходу з металу (рис. 7.18, е на вставці), викликаючи ефект, подібний до ефекту Шоттки під час

термоелектронної емісії електронів у вакуум (див. підрозд. 3.7.2). Це приводить до зростання струму зі збільшенням зворотної напруги.

Омічні контакти. Зовсім інші властивості має контакт металу та донорного напівпровідника, якщо $\Phi_M < \Phi_n$. (рис. 7.19).

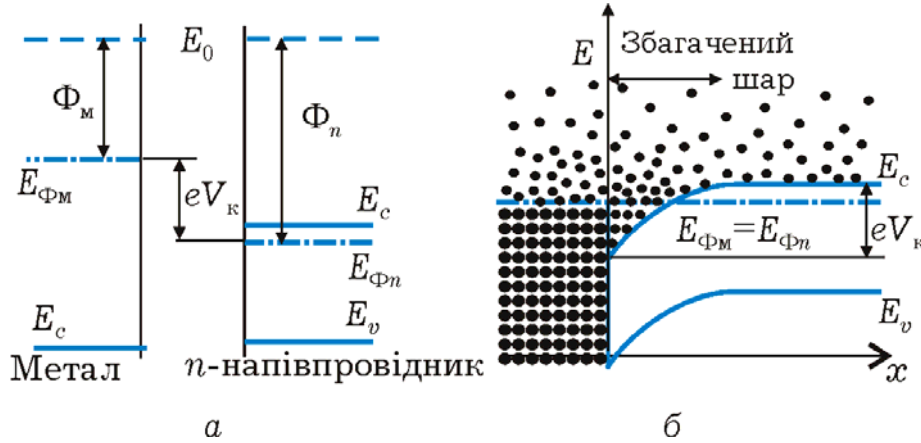


Рис. 7.19. Енергетичні діаграми металу та донорного напівпровідника у разі $\Phi_M < \Phi_n$ до контакту (а). У рівноважному стані утворюється збагачений електронами шар просторового заряду (б) у потенціальній ямі зони провідності

Процес встановлення рівноважного стану у такому контакті супроводжується переважаючим потоком електронів з металу і формуванням подвійного електричного шару. Шар додатного заряду створюється практично на поверхні металу, а просторовий заряд електронів утримується у приповерхневому шарі напівпровідника. Електричне поле просторового заряду викривлює енергетичні зони таким чином, що у зоні провідності утворюються потенціальна яма (рис. 7.19, б), де і концентруються надлишкові електрони. Розподіли потенціалу, напруженості електричного поля та концентрацій носіїв аналогічні таким, які розглянуто в розд. 7.1 для режиму збагачення.

Вольт-амперна характеристика такого контакту формально має бути асиметричною і мати вигляд 7.30, оскільки умови проходження електронів через бар'єр у різні боки неоднакові. Але збагачений електронами контактний шар напівпровідника має малий опір, тому нахил вольт-амперної характеристики поблизу рівноважного стану, який визначається диференціальним опором, буде відповідати високим значенням струму за малих падінь напруги на контакті. Такі контакти мають майже лінійну вольт-амперну характеристику і їх називають *омічними*.

7.6. n^+n -КОНТАКТИ

Схожі властивості має контакт між шарами напівпровідника з суттєво різними концентраціями домішки, причому один з них (n^+) є виродженим, або близьким до виродження. Такі контакти поряд з попередніми є невід'ємною частиною багатьох електронних компонентів.

Зонна структура n^+n -контакту подібна до такої в омичному контакті метал-напівпровідник, бо енергія Фермі збільшується з підвищенням концентрації донорів і $\Phi_{n^+} < \Phi_n$ (рис. 7.20, а).

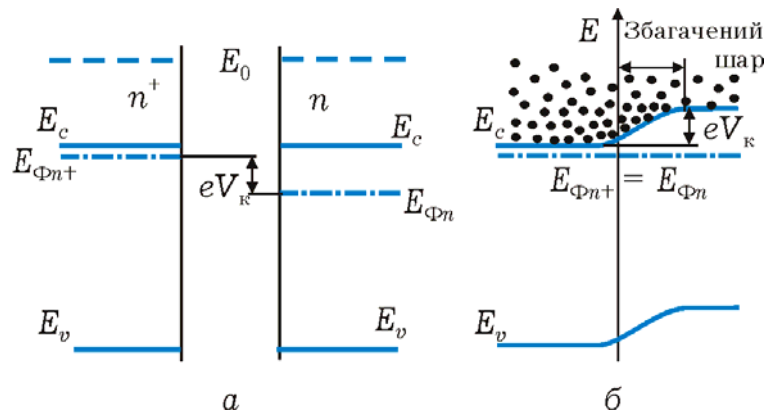


Рис. 7.20. Енергетичні діаграми n^+ - та n -напівпровідників до контакту (а). У рівноважному стані в n -області утворюється збагачений електронами шар просторового заряду (б) в потенціальній ямі зони провідності та збіднений в n^+ - області.

В рівноважному стані між шарами напівпровідника встановлюється контактна різниця потенціалів, яка визначається відношенням концентрацій електронів (домішки)

$$V_{\text{к}} = \frac{kT_{\text{Б}}}{e} \ln \frac{N_{\text{д}}^+}{N_{\text{д}}}. \quad (7.73)$$

Контактне електричне поле матиме такий саме напрямок, як і у попередній структурі метал-напівпровідник, викликаючи режим збагачення в n -області. Основна відмінність полягає у тому, що силові лінії поля на відміну від металу не закінчуються на поверхні n^+ -області, а заходять у неї на певну відстань, викликаючи там викривлення зон догори (рис. 7.20, б) і збіднення шару електронами. Враховуючи відмінність шарів у концентраціях домішки на декілька порядків, товщина шару збіднення буде набагато меншою за товщину шару збагачення.

Просторові розподіли, наведені на рис. 7.21 можна описати,

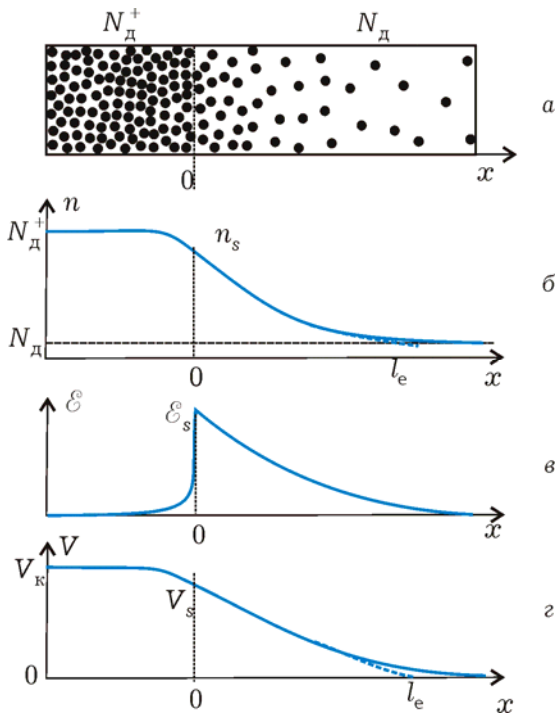


Рис. 7.21. Просторові розподіли концентрації електронів (а), напруженості поля (б) та потенціалу (в) в n^+n^- -структурі.

відповідною зміною позначень

використовуючи результати, отримані в розд. 7.1. Вважаючи, що режим збіднення в n^+ -шарі є слабким, можна використати вираз (7.5) для експоненціального розподілу потенціалу з відповідним врахуванням позначень напрямку та початку відліку

$$V(x) = (V_{\kappa} - V_s) \exp(x/l_e^+) \quad (7.74)$$

Відповідно, поверхневий заряд у шарі збіднення дорівнює

$$\Delta q^+ = e^2 N_{\text{д}}^+ (V_{\kappa} - V_s) l_e^+ / kT_{\text{Б}}. \quad (7.75)$$

У збагаченому шарі n^- області для розподілу потенціалу придатний вираз (7.23) з

$$V(x) \approx V_s - 2kT_{\text{Б}} \ln [1 + x \exp(eV_s / 2kT_{\text{Б}}) / ex_0], \quad (7.76)$$

де $x_0 \approx \sqrt{2}l_e^-$ – товщина шару збагачення

Розподіл носіїв у збагаченому шарі може бути описаний виразом (7.25):

$$n(x) = N_{\text{д}} e^{eV(x)/k_{\text{Б}}T} = N_{\text{д}} e^{eV_s/k_{\text{Б}}T} / [1 + x \exp(eV_s / 2k_{\text{Б}}T) / ex_0]^2, \quad (7.77)$$

причому в точці x_0 справедлива рівність $n(x_0) = N_{\text{д}}$, а на поверхні $n(0) \equiv n_s = N_{\text{д}} \exp(eV_s / k_{\text{Б}}T)$.

Надлишковий заряд електронів на одиниці поверхні збагаченого шару можна підрахувати через добуток середньої концентрації (7.26) та товщину (7.24) збагаченого шару

$$\Delta q = eN_{\text{д}} x_0 \exp(eV_s / 2k_{\text{Б}}T) \quad (7.78)$$

Прирівнюючи надлишкові заряди (7.75) та (7.78) по обидва боки, що є однією з граничних умов, можна отримати трансцендентне рівняння $s e^{s/2} = \sqrt{2}$ відносно величини $s \equiv e(V_k - V_s) / k_B T$, з якого легко визначити потенціал на межі шарів. Враховуючи розв'язок цього рівняння ($s \approx 0,9018$), маємо, що граничний потенціал $V_s \approx V_k - 0,902 k_B T / e$ менший від контактного на величину трохи меншу за тепловий потенціал. Відповідно концентрація на межі $n_s = N_d^+ / e^s \approx N_d^+ / 2,464$ зменшується. Слід зауважити про наближеність отриманих результатів, бо розподіл (7.23) отримано за умови більш-менш значного збагачення: $n \gg N_d$, а режим збіднення в n^+ -шарі є слабким

7.7. ГЕТЕРОПЕРЕХОДИ

Гетеропереходи утворюються біля межі контакту двох різних матеріалів. Найчастіше – це напівпровідники, які на відміну від гомопереходів мають різну ширину забороненої зони та можливо різну концентрацію або вид домішки. Вони лежать в основі більшості сучасних твердотільних приладів, які є *гетероструктурами*.

Поблизу межі поділу двох напівпровідників через різний атомний потенціал характерні електронні стани, наприклад, дно зони провідності (рис. 7.22), які поступово змінюються від одного матеріалу до другого. Просторово область локалізації цієї зміни складає кілька періодів ґратки, тобто має порядок одиниць нанометрів. Для потреб мікроелектроніки така відстань є нехтувально малою і можна вважати, що ця зміна має вигляд сходинок.

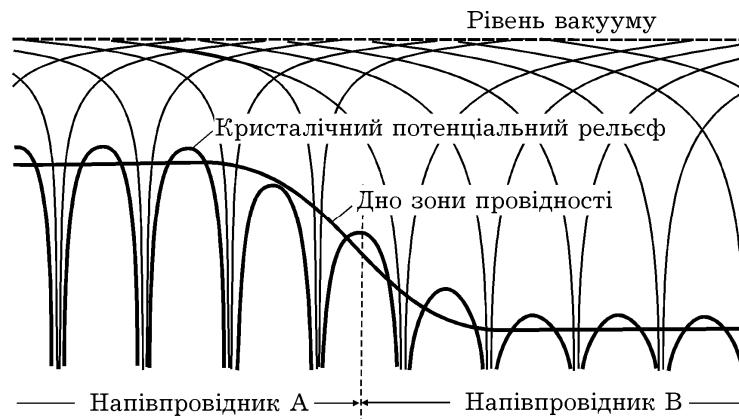


Рис. 7.22. Зміна атомарного потенціального рельєфу на межі поділу двох напівпровідників

Тому характерною особливістю «макроскопічних» енергетичних діаграм гетеропереходів є «розриви» ліній дна зони провідності та стелі валентної зони. Вигляд енергетичної діаграми і властивості гетеропереходу залежать як від значень ширини забороненої зони обох напівпровідників ΔE_1 та ΔE_2 , так і від їхньої енергії електронної спорідненості χ_1 та χ_2 (рис. 7.23). Якщо вважати, що спільним початком відліку енергії є рівень вакууму E_0 , то відповідно до рисунка величина розриву дна зони провідності становитиме $\Delta E_c = \chi_1 - \chi_2$, а величина розриву стелі валентної зони - $\Delta E_v = (\chi_2 + \Delta E_2) - (\chi_1 + \Delta E_1)$.

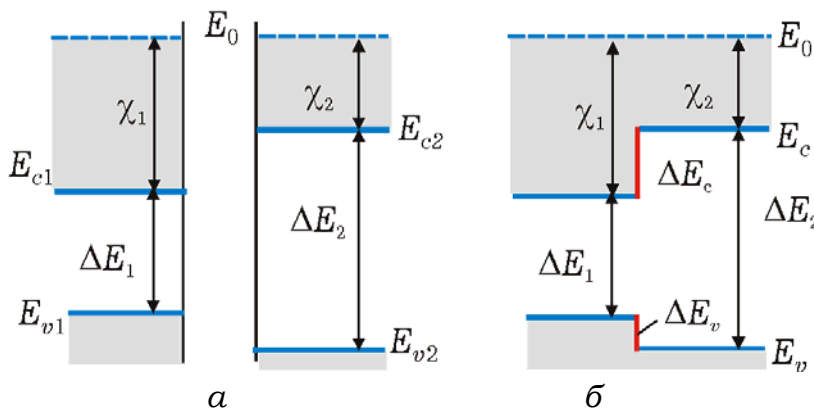


Рис. 7.23. Контакт двох напівпровідників з різними параметрами енергетичної структури викликає «розриви» ліній дна зони провідності та стелі валентної зони на макроскопічних енергетичних діаграмах

Серед можливих варіантів енергетичної структури розрізняють три типи гетеропереходів (рис. 7.24) в залежності від співвідношень між енергіями електронної спорідненості та між ширинами заборонених зон напівпровідників.

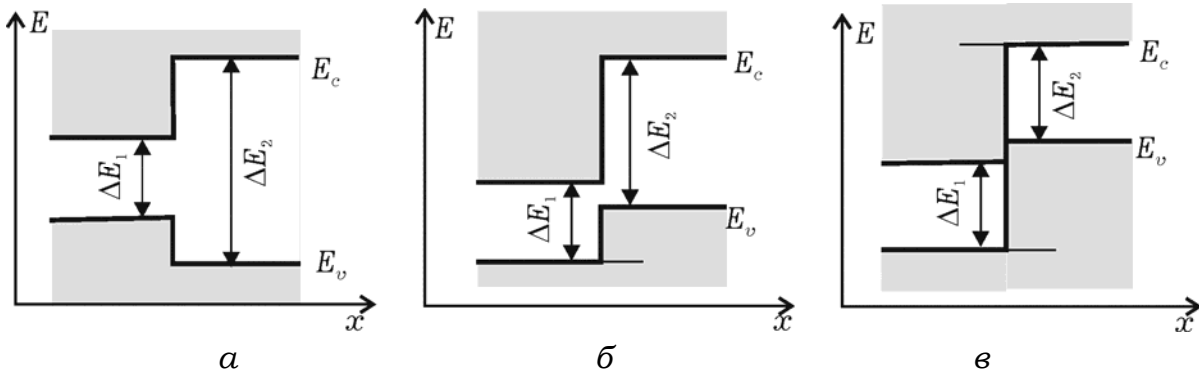


Рис. 7.24. Різні типи гетеропереходів щодо взаємного розміщення забороненої зони: а) перехід I типу (GaAs-AlGaAs; GaSb-AlSb; GaAs-GaP; InGaAs-InAlAs; InGaAs-InP); б) перехід II типу (InP-Al_{0,48}In_{0,52}As; InP-InSb; GaInP-GaAsP; Si-SiGe); в) перехід III типу (InAs-GaSb)

Важливою передумовою високої якості гетеропереходу є однаковість сталих кристалічної ґрати. Їх неспівпадання приводить до появи поверхневих дефектів, які можуть суттєво впливати на електричні властивості гетеропереходу. Втім в наноелектроніці це іноді використовують спеціально для створення, наприклад, квантових точок.

Під час проектування гетеропереходів зручно використовувати довідникові діаграми, подібні до рис. 7.25, для підбору як величин забороненої зони, так і міжатомної відстані. Наприклад, з наведеної діаграми видно, що для потрійної сполуки $Al_xGa_{1-x}As$, яку можна розглядати як сплав простих бінарних сполук $AlAs$ та $GaAs$ у співвідношенні, яке задається числом x , стала кристалічної ґратки залишається практично незмінною для будь-якого стехіометричного складу. Це є однією з причин, чому саме переходи $Al_xGa_{1-x}As/GaAs$ є найбільш дослідженими та технологічно відпрацьованими, з їх використанням отримано рекордну рухливість $2000 \text{ м}^2/\text{В}\cdot\text{с}$ (це відповідає довжині вільного пробігу електронів $1,5 \text{ мм!}$).

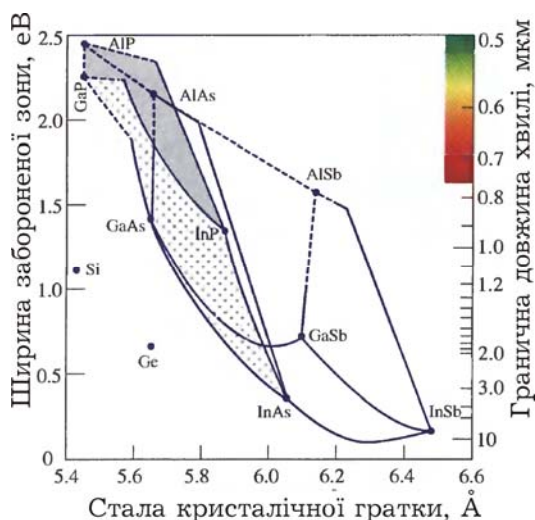


Рис.7.25. Діаграма залежностей між шириною забороненої зони та міжатомною відстанню для деяких матеріалів. Лінії, що поєднують дві точки, відповідають потрійним сполукам змінного складу, що є сплавами відповідних бінарних матеріалів. Права шкала – гранична частота фундаментального поглинання

На правій шкалі рис. 7.25 наведено граничну довжину хвилі фундаментального поглинання, яка відповідає ширині забороненої зони відповідної сполуки. Це зручно при проектуванні оптоелектронних компонентів, бо саме у фотоніці, мабуть найчастіше, використовують потрійні або більш складні сполуки.

Остаточний вигляд енергетичних діаграм гетероконтактів, на відміну від зображених з «плоскими зонами» на рис. 7.24, залежить не тільки від значень ширини забороненої зони і електронної спорідненості, але й від виду та концентрації домішки. Через різну роботу виходу матеріали, обмінюючись електронами, утворюють

контактну різницю потенціалів, яка викликає викривлення енергетичних зон поблизу контакту.

Розглянемо для прикладу контакт двох напівпровідників з різним типом провідності (рис. 7.26, а), так званий *анізотипний* гетероперехід. Через різні рівні Фермі і відповідно різну роботу виходу (рис. 7.26, б), як і у звичайному електронно-дірковому переході, дифузійно-дрейфова рівновага встановлюється за рахунок контактної електричного поля. Воно викликає зміну потенціальної енергії електрона, яка на енергетичній діаграмі виглядає як викривлення енергетичних зон (рис. 7.26, в).

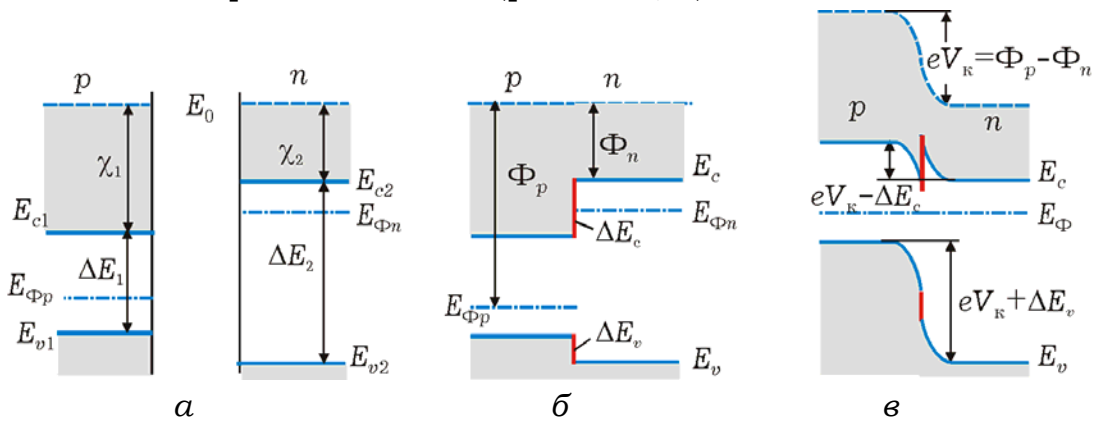


Рис. 7.26. Енергетичні діаграми акцепторного та донорного напівпровідників до контакту (а), після контакту, але без врахування контактної різниці потенціалів (б), після встановлення рівноважного стану і контактної різниці потенціалів (в)

Характерною особливістю такого гетероструктурного діоду є те, що різні знаки розривів зон ΔE_c та ΔE_v забезпечує різну висоту потенціального бар'єру для різнойменних носіїв: $eV_\kappa - \Delta E_c$ для електронів, що переходять з n -області до p -області, і $eV_\kappa + \Delta E_v$ - для дірок, які рухаються у зворотному напрямку. Така відмінність забезпечує практично односторонню інжекцію, у даному випадку електронів з n - в p -напівпровідник. Тут також мати на увазі, що на межі переходу утворюється потенціальний пік, який може бути достатньо вузьким і тому частково прозорим для електронів.

Властивість односторонньої провідності можуть мати також ізотипні nn - або pp -гетеропереходи. На рис. 7.27 зображена енергетична діаграма електронно-електронного nn -гетеропереходу. Як і у попередньому випадку, умови переходу носіїв у різних напрямках будуть різними, що викликано відмінністю у ширині заборонених зон. Але головною особливістю даного гетеропереходу є

те, що прями́й струм створюється практично тільки електронами, а процес дифузії неосновних носіїв відсутній: електрони з широкозонного напівпровідника «інжектуються» в також електронний напівпровідник. Відсутність перехідних процесів, пов'язаних з рекомбінацією зі скінченим часом життя неосновних носіїв, дозволяє використовувати такі гетеропереходи, як і переходи з бар'єром Шоттки, на надвисоких частотах.

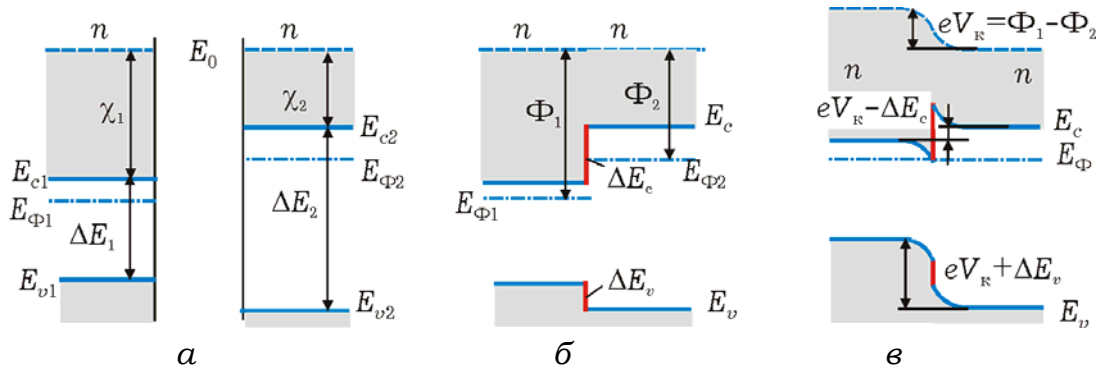


Рис. 7.27. Енергетичні діаграми двох донорних напівпровідників з різною шириною забороненої зони до контакту (а), після контакту, але без врахування контактної різниці потенціалів (б), після встановлення рівноважного стану і контактної різниці потенціалів (в)

Слід також відзначити ще один специфічний вид гетероструктур – так звані варизонні напівпровідники. Вони утворюються у разі епітаксiального вирощування напівпровідників

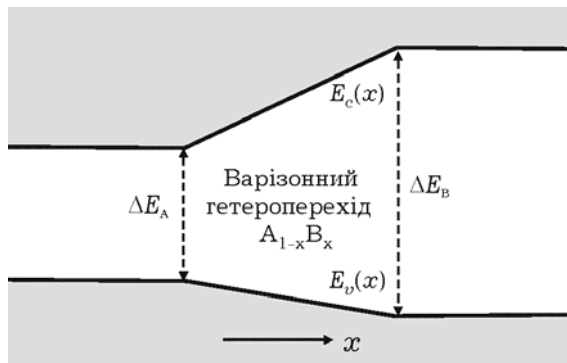


Рис. 7.28. Енергетична діаграма варизонного переходу для лінійної зміни ширини забороненої зони

для сполук $A^III B^V$, ширина забороненої зони змінюється непропорційно стехіометричному параметру x при утворенні твердих розчинів типу $AB_{1-x}C_x$. Крім того, у деяких

зі змінним стехіометричним складом (рис. 7.28). Поступова зміна співвідношення між матеріалами приводить до зміни ширини забороненої зони. У першому наближенні вважають, що

$$\Delta E_{AB} = x \Delta E_A + (1 - x) \Delta E_B$$

Тому її профіль $\Delta E(x)$ є лінійним для рівномірної зміни стехіометричного складу. Проте, як показує діаграма на рис. 7.25

напівпровідниках, наприклад, у такому поширеному матеріалі, як $\text{Al}_x\text{Ga}_{1-x}\text{As}$, змінюється положення долин у зоні провідності, що також приводить до порушення монотонності функції $\Delta E(x)$ від x .

Наведені приклади, звісно, не вичерпують усіх можливих типів гетеропереходів. Їх особливості, пов'язані з різною шириною забороненої зони та електронною спорідненістю, надають електронним технологам додаткову ступінь свободи для створення різноманітних сучасних твердотільних електронних компонентів.

7.8. ТРАНЗИСТОРНІ СТРУКТУРИ

7.8.1. Гетеробіполярні транзистори

Підвищення швидкодії біполярних транзисторів можливе, якщо для емітера використовувати матеріал з більшою шириною забороненої зони, ніж для матеріалу бази. Наприклад, перші гетеротранзистори мали емітер з $n\text{-GaAs}$, а базу з $p\text{-Ge}$. Пізніше почали використовувати гетеропереходи на основі з'єднань $\text{Al}_x\text{Ga}_{1-x}\text{As}$, $\text{Al}_x\text{Ga}_{1-x}\text{N}$ та ін.

На межі різних напівпровідників виникають електричні поля, які на енергетичній діаграмі (див. рис. 7.23) моделюються розривом дна зони провідності, рівним різниці електронної спорідненості $\Delta E_c = \chi_1 - \chi_2$, і розривом стелі валентної зони $\Delta E_v = (\Delta E_1 + \chi_1) - (\Delta E_2 + \chi_2)$. Разом з полем контактної різниці потенціалів $V_k = (\Phi_1 - \Phi_2)/e$ (рис. 7.29, а) ці поля зменшують висоту потенціального бар'єру для електронів і збільшують його для дірок валентної зони. Асиметрія умов проходження електронів і дірок через емітерний перехід приводить до того, що струм інжекції дірок j_p зменшується в $\exp(\Delta E_v/kT)$ раз, а струм інжекції електронів j_n збільшується в $\exp(\Delta E_c/kT)$ раз. В результаті можна отримати коефіцієнт інжекції електронів $\gamma = j_n/(j_n + j_p)$ практично рівним одиниці, що робить процеси рекомбінації дірок в емітері неістотними.

У звичайному гомопереході відношення електронної і діркової складових струму емітера визначається відношенням концентрації донорів в емітері і акцепторів в базі, що вимушувало для збільшення коефіцієнта інжекції обмежувати ступінь легування бази. У гетеропереході можливе сильне легування базової області, що призводить до зниження подовжнього опору бази і сталої часу

заряду колекторного кола $\tau_k = R_b C_{kб}$, яка істотно впливає на швидкодію транзистора при надтонких базах.

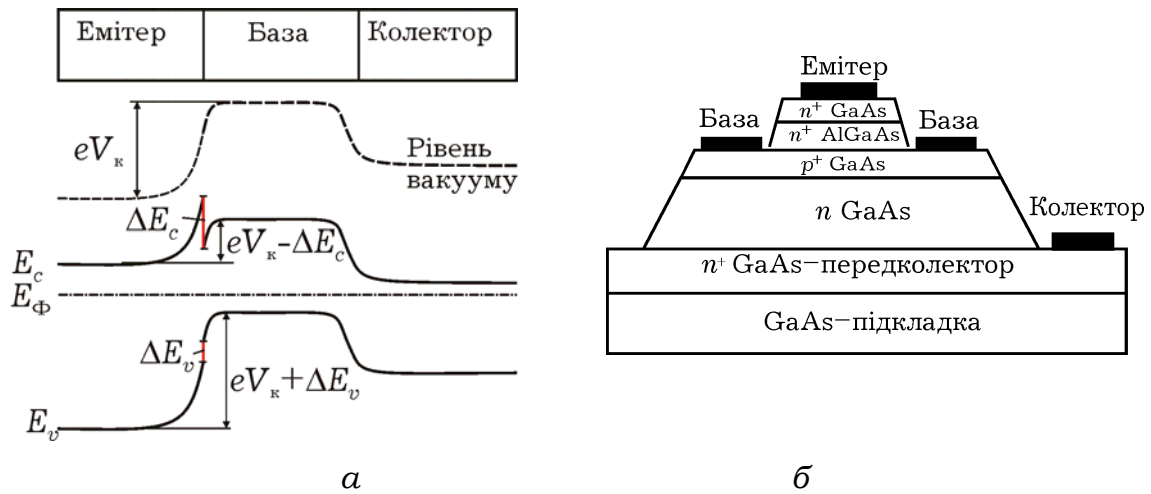


Рис. 7.29. Енергетична діаграма типового *pn*-гетеропереходу (а) та топологічна структура гетеробіполярного транзистора (б)

Крім того, на межі емітер-база електрони набувають додаткової кінетичної енергії, рівної ΔE_c , і при відповідній товщині бази може бути реалізований балістичний, або близький до його, режим перенесення в вузькозонному матеріалі. Це зменшує час прольоту в базі, що є ще одним чинником підвищення швидкодії.

Завдяки можливостям технології, використанню нових напівпровідникових матеріалів і оптимізації топології субмікронні гетеробіполярні транзистори мають граничні частоти посилення більше 200 ГГц і використовуються в широкому спектрі пристроїв міліметрового діапазону.

Ще одна перевага гетеробіполярних транзисторів, пов'язана з конструктивними особливостями (рис. 7.29, б). В них основне нагрівання відбувається у переході біля колектора, який набагато легше охолоджувати, ніж у польових транзисторах, де тепло виділяється вздовж всього каналу. Саме тому гетеробіполярним транзисторам віддають перевагу при конструюванні потужних пристроїв.

7.8.2. Субмікронні польові транзистори з бар'єром Шоттки

Надвисокочастотні польові транзистори із затвором у вигляді бар'єру Шоттки (ПТШ) виготовляють зазвичай з арсеніду галію, рідше з фосфіду індію або нітриду галію. Основою структури приладу служить напівізольююча підкладка (рис. 7.30, а), на якій вирощено активний епітаксіальний шар з високою концентрацією донорної домішки, а також тонкий n^+ -контактний шар, нанесений на шар з концентрацією домішки $\sim 1 \cdot 10^{19} \text{ см}^{-3}$. Такий шар необхідний для формування омичних контактів стоку і витоку.

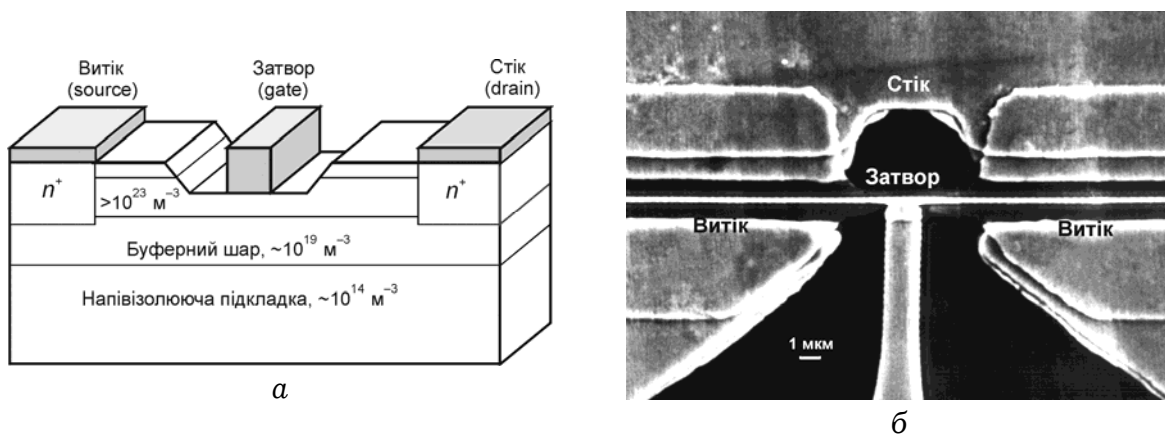


Рис. 7.30. Типова структура (а) і фотографія (б) польового транзистора із затвором у вигляді бар'єру Шоттки

Затвор, розташований між контактами стоку і витоку, утворює з матеріалом напівпровідника бар'єр Шоттки. У пластині в поверхневому n^+ -шарі витравлюється поглиблення до потрібної напруги відсічення і струму стоку насичення, в якому потім формується субмікронний затвор.

Контактна різниця потенціалів бар'єру разом з прикладеною зовнішньою напругою, від'ємною по відношенню до витоку потенціалом зміщення (транзистор зазвичай включається по схемі із спільним витоком), створюють під затвором збіднену область, що формує провідний канал для електронів між стоком і витоком. Товщина збідненої області у наближенні її різкої границі залежить від різниці потенціалів затвора V_{gs} (з врахуванням контактної різниці на бар'єрі Шоттки V_b) і потенціалом каналу $V(x)$, що збільшується від витоку до стоку, і визначається відомим співвідношенням (див. розд. 7.1)

$$d(x) = \sqrt{\frac{2\epsilon\epsilon_0[V(x) + V_b - V_{gs}]}{eN_d}}. \quad (7.79)$$

Зміни потенціалу затвора приводять до модуляції товщини каналу і відповідно до управління струмом стоку. У певному наближенні канал можна розглядати як нелінійний резистор, керований потенціалом затвора. Стокові характеристики польового транзистора (рис. 7.31), тобто залежності струму стоку I_d від напруги стоку V_{ds} , за малої напруги стоку близькі до лінійних і добре описуються такою «резистивною» моделлю управління.

Зі збільшенням напруги стоку відповідно до (7.79) зростає і поточний потенціал каналу, а товщина збідненого шару збільшується. Тому за великої напруги стоку при звуженні каналу його опір зростає, що викликає зменшення нахилу стокових характеристик.

Ця тенденція справедлива до значення напруги стоку, при якому товщина збідненої області у стоковій частині стає рівною товщині епітаксіального шару a і канал «перекривається». Це поняття, типове для моделі різкої межі збідненої області, не означає відсутності провідності, залишається провідним тонкий шар порядку довжини екранування, в якому напруженість поля досягає значень, вищих за критичне, і швидкість електронів насичується. Після цього значення струму залишається незмінним (рис. 7.31) при збільшенні напруги стоку і рівним струму насичення:

Для більш від'ємних значень потенціалу затвора V_{gs} «початкова» товщина провідного каналу стає меншою і перехід до режиму насичення відбуватиметься раніше, для менших стокових потенціалів. Струм через транзистор буде відсутнім, якщо потенціал затвора дорівнюватиме V_p – так званому потенціалу перекриття, для якого товщина збідненої області біля витоку стає рівною товщині епітаксіального шару a :

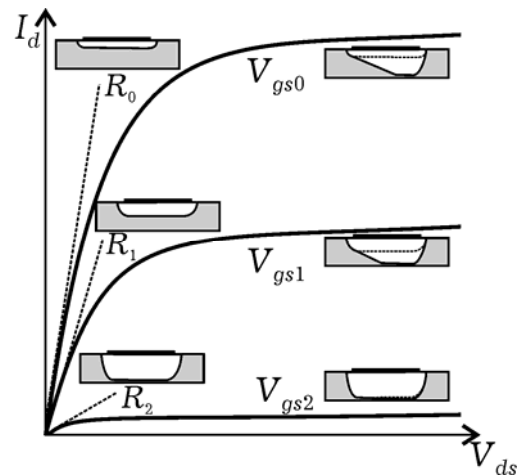


Рис. 7.31. Зміна стокового струму і форми каналу зі зміною напруги затвора і стоку

$$V_p = -\frac{eN_d a^2}{2\epsilon\epsilon_0} - V_b. \quad (7.80)$$

Проста аналітична теорія польових транзисторів [НШП], створена Шоклі, заснована на так званому наближенні «плавного каналу», що передбачає повільну зміну товщини збідненої області. Вона придатна для довгоканальних транзисторів, у яких довжина затвора складає мікрометри, що значно більше товщини епітаксіального шару. Але зменшення довжини затвора, необхідне для підвищення швидкодії, робить припущення цієї моделі неспроможними. Це можна проілюструвати на прикладі стокових характеристик транзистора – укорочення затвора до субмікронних розмірів якісно змінює їх вигляд (рис. 7.32).

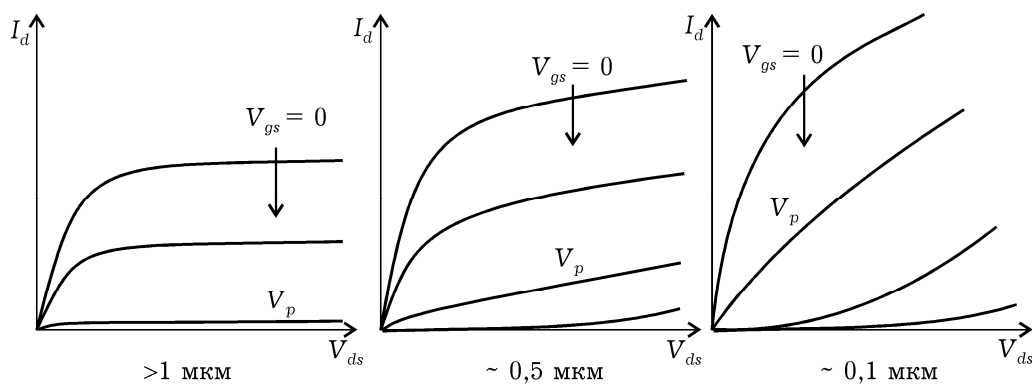


Рис. 7.32. Трансформація форми стокових характеристик при зменшенні довжини затвора

Для субмікронних довжин затвора на вольтамперних характеристиках не спостерігається насичення – вихідна провідність стає значною і тим вищою, чим менша довжина затвора. Примітно також, що для потенціалу затвора, близького до потенціалу перекриття, стокова характеристика близька до лінійної, а її нахил, що визначає вихідну провідність, приблизно дорівнює нахилу інших ВАХ у режимі «насичення». Крім того, її вигляд близький до омичного. Таку ситуацію можна умовно представити як паралельне з'єднання провідності каналу звичайного транзистора і деякої шунтуючої провідності. Кількісні оцінки показують, що таку провідність можуть забезпечити електрони збагаченого шару, що утворюється біля границі епітаксіального шару і підкладки. Підтвердженням такої моделі можуть служити результати моделювання транзистора з довжиною затвора 0,5 мкм методом

Монте-Карло. На рис. 7.33 представлені потенціальний рельєф і просторовий розподіл електронів у каналі відкритого транзистора. Видно, що значна частка електронів знаходиться поза межами епітаксіального шару, створюючи струм, паралельний струму каналу.

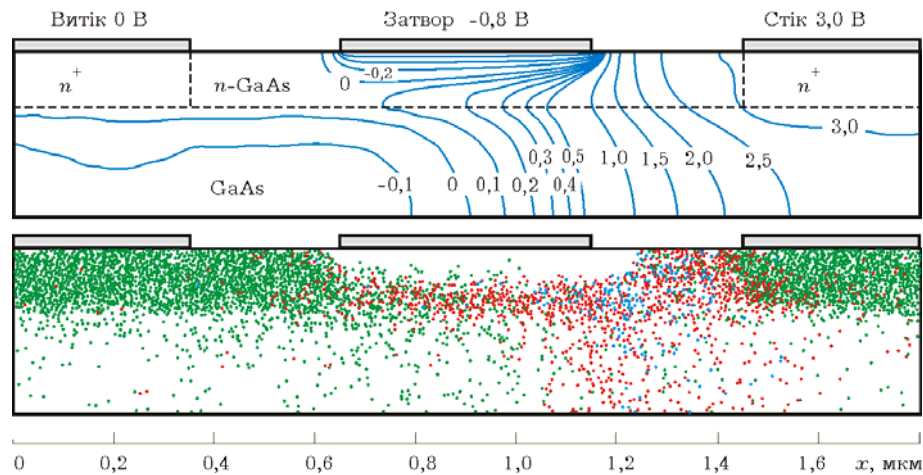


Рис. 7.33. Просторовий розподіл і електронів у каналі відкритого транзистора з довжиною затвора 0,5 мкм. Штрихові лінії позначають межі епітаксіального шару та омичних контактів

Тут треба звернути увагу на те, що умови дрейфового руху електронів в каналі і підкладці істотно різні. Переходячи в підкладку біля витоків за рахунок хаотичної швидкості, вони утворюють контактний потенціальний бар'єр (рис. 7.34, а), висота якого (трохи більша 0,1 еВ) визначається температурою і відношенням між концентраціями домішки в епітаксіальному шарі і підкладці. З розподілу електронів на рис. 7.33 видно також, що в стоковій частині під затвором концентрація електронів у підкладці значно вища. Це відбувається за рахунок розігрівання електронів у каналі і підвищення їх електронної температури. В результаті контактний потенціальний бар'єр істотно підвищується і може досягати значень більш 1 еВ. У результаті з розподілу потенціального рельєфу наочно видно, що подовжнє електричне поле в підкладці слабкіше, ніж в каналі, що створює різні умови при протіканні струму через області каналу і підкладки.

Саме наявність струму через підкладку означає, що для потенціалу перекриття за формулою (7.80) запирання транзистора не відбувається, струм стоку створюють електрони, що проходять через підкладку (рис. 7.34, б). Цей струм також регулюється

затвором, але вплив його потенціалу буде значно меншим, і він поступово спадає, експоненціально зменшуючись.

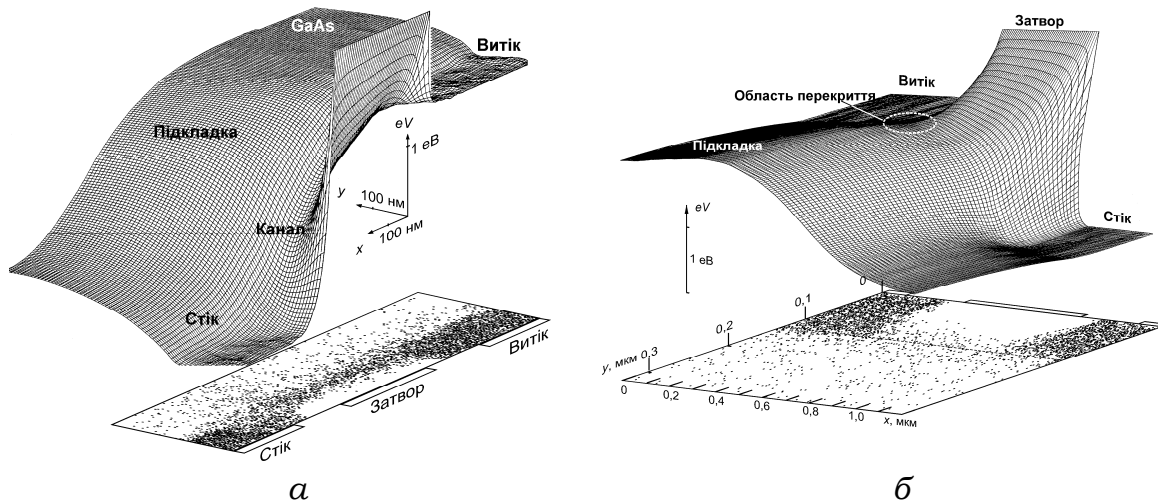


Рис. 7.34. Потенціальний рельєф і розподіл електронів у каналі відкритого ПТШ (а) та поблизу режиму перекриття (б)

Описана фізична модель добре підтверджується ходом експериментальних статичних ВАХ: сток-затворної і стокової характеристик, показаних на рис. 7.35 для транзистора з довжиною затвора 0,5 мкм, шириною – 300 мкм, товщиною епітаксiального шару – 0,12 мкм, концентрацією домішки в епітаксiальному шарі – $1,9 \cdot 10^{17} \text{ см}^{-3}$.

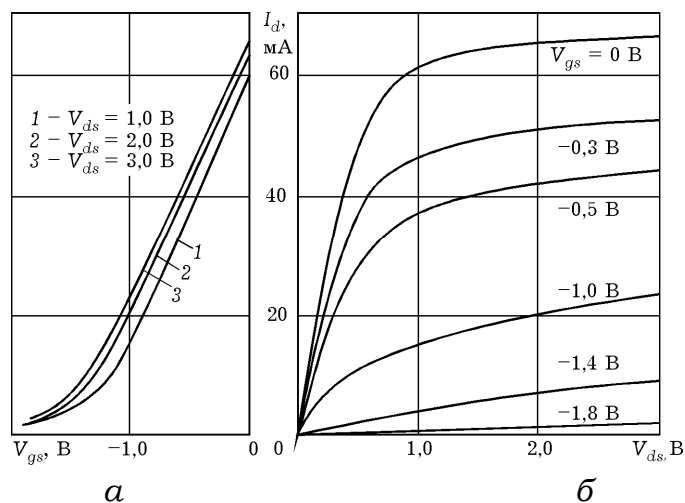


Рис. 7.35. Експериментальні сток-затворні (а) та стокові (б) характеристики ПТШ з довжиною затвора 0,5 мкм.

Експоненціальний характер спадання струму стоку при запиранні транзистора може бути пояснений за допомогою кінетичної моделі Монте-Карло. При пороговій напрузі канал

перекривається у витоковій частині каналу, де утворюється потенціальний бар'єр, добре помітний на рис. 7.34, б. Існування струму стоку при цьому пояснюється надбар'єрним проходженням електронів за рахунок початкових теплових швидкостей від витоку в область прискорюючого поля каналу.

Слід звернути увагу на вигляд стокових характеристик для довжини затвору біля 0,1 мкм (рис. 7.32) – він нагадує анодні характеристики вакуумних тріодів. Це стає зрозумілим, якщо прийняти до уваги, що довжина вільного пробігу електронів має такий самий порядок, тому під затвором електрони пролітають балістично, практично без зіткнень.

7.8.3. Гетероструктурні польові транзистори

Поліпшення параметрів ПТШ може бути досягнуте збільшенням концентрації носіїв в каналі за рахунок підвищення рівня легування. Але при цьому через домішкове розсіяння падає рухливість і відповідно швидкість носіїв. Вихід з цієї ситуації може дати використання гетеропереходів. Основна ідея при цьому полягає в тому, щоб просторово розділити області легування з іонізованими донорними атомами і породжених ними електронів, що повинно істотно понизити домішкове розсіяння і підвищити рухливість. Одним із способів реалізації цієї ідеї є використання гетероконтакта широкозонного високолегованого напівпровідника з нелегованим вузькозонним. Відмінність енергій електронної спорідненості приводить до появи розриву дна зони провідності і, як наслідок, до переходу електронів з широкозонного матеріалу в нелегований вузькозонний. Наявність просторового заряду електронів викликає вигин зони провідності і підвищення концентрації електронів в створюваній потенціальній ямі. Найбільш вивченими і технологічно відпрацьованими є гетероструктури GaAs/AlGaAs. Рухливість електронів, що перейшли в шар вузькозонного матеріалу і створюють там двовимірний електронний газ, досягає для кімнатної температури значень, характерних для чистого матеріалу.

Велика концентрація електронів в тонкому шарі, висока рухливість, а також можливість управління концентрацією носіїв в каналі сприяли швидкій розробці нового типу польових транзисторів на двовимірному електронному газі, так званих ГСПТ (гетероструктурний польовий транзистор), або *HEMT* (*High Electron*

Mobility Transistor), або HFET (Heterostructure Field Effect Transistor).

Схематичне зображення такого транзистора дано на рис. 7.36

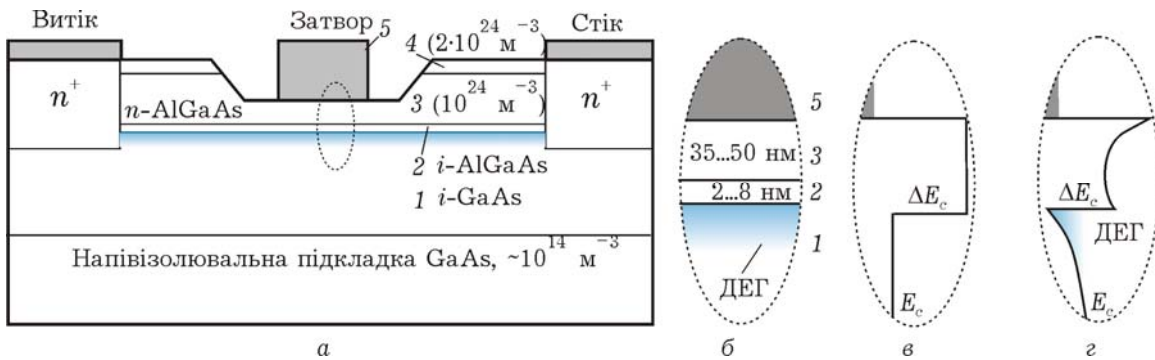


Рис. 7.36. Поперечний перетин ГСПТ (а) з характерною товщиною шарів (б): 1 – нелегований GaAs, 2 – спейсер – нелегований шар AlGaAs, 3 – бар'єрний шар – n-AlGaAs, 4 – контактний шар – n⁺-GaAs. Потенціальний профіль дна зони провідності вихідних матеріалів (в) і в рівноважному стані (г).

Основним структурним «елементом» в ГСПТ на основі GaAs є область двовимірного електронного газу (ДЕГ). Вона формується в зоні гетеропереходу між бар'єрним Al_xGa_{1-x}As-шаром і нелегованим каналним GaAs-шаром. Ширина забороненої зони для Al_xGa_{1-x}As вища, ніж для GaAs, і в області гетеропереходу границя зони провідності GaAs може опинитися нижче за рівень Фермі – утворюється потенціальна яма (рис. 7.36, г).

У ГСПТ на основі GaAs, щоб забезпечити джерело вільних електронів, бар'єрний шар Al_xGa_{1-x}As легують кремнієм. Перехід електронів до потенціальної ями з області бар'єрного шару викликаний тим, що через різні енергії електронної спорідненості GaAs і Al_xGa_{1-x}As в області їх контакту утворюється досить сильне електричне поле, що моделюється зазвичай розривом забороненої зони ΔE_c , що ініціює перехід вільних електронів з Al_xGa_{1-x}As до потенціальної ями. Часто обирають частку Al $x \approx 0,3$, за якої $\Delta E_c \approx 0,3$ еВ. Таким чином, в каналному GaAs-шарі безпосередньо під гетеропереходом утворюється тонкий шар просторового заряду з високою густиною і рухливістю вільних електронів, який і називають *двовимірним електронним газом*.

Між бар'єрним і каналним шаром формують спейсер (spacer) – шар високоомного нелегованого AlAs або Al_xGa_{1-x}As. Основне призначення цього тонкого (1–3 нм) шару – запобігти розсіянню носіїв в каналі на границі з гетеропереходом. Верхня частина

бар'єрного шару утворює з металом затвора бар'єр Шоттки, котрий, як і в звичайному ПТШ, керує провідністю каналу. Часто в бар'єрному шарі формують область з підвищеною концентрацією Si (так зване δ -легування), n^+ - $\text{Al}_x\text{Ga}_{1-x}\text{As}$ шар для поліпшення омичних контактів стоку і витоку і збільшення концентрації носіїв у каналі.

Приклади розподілу потенціальної енергії і електронного заряду в модельованій області зображені на рис. 7.37 для структури з довжиною затвора 0,35 мкм. Шар $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ товщиною 50 нм рівномірно легований донорами з концентрацією 10^{14} см^{-3} . Уздовж межі гетеропереходу розміщено нелегований спейсер товщиною 5 нм.

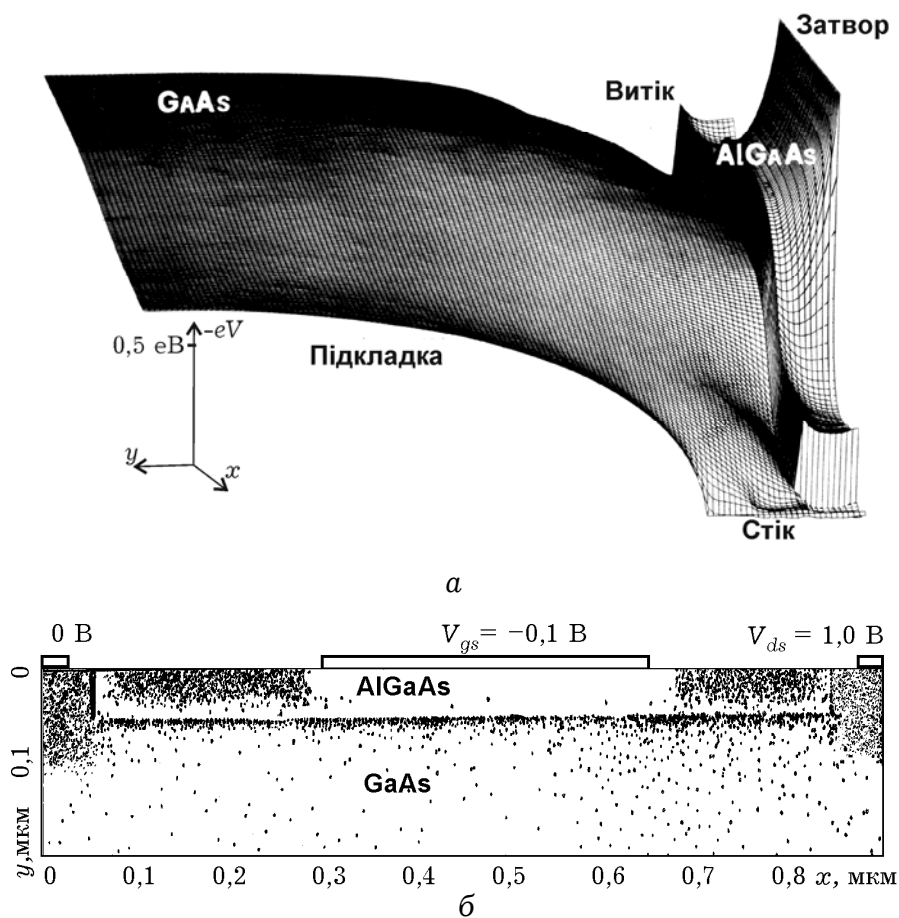


Рис. 7.37. Потенціальний рельєф дна зони провідності (а) і розподіл електронів (б) в поперечному перетині ГСПТ.

На верхньому малюнку слід звернути увагу на зміну уздовж каналу форми потенціальної ями, створеною «стілкою» розриву дна зони провідності і просторовим зарядом електронів, що перейшли в шар GaAs з легованого AlGaAs (рис. 7.36, б). Розподіл електронів у витоковій частині каналу локалізовано біля границі гетеропереходу і, відповідно, в цій частині потенціальна яма більш крута. З рухом

електронів до витoku під дією поздовжнього електричного поля їх температура підвищується. Тому, починаючи з середини затвора, вони глибше проникають у шар GaAs, і форма потенціальної ями стає більш пологою.

Подібна інформація міститься в розподілі електронів по поперечному перетину (рис. 7.37, б). Тут добре видно область збіднення під затвором, практично повністю позбавлена рухливих носіїв, а також збіднені області біля границі гетеропереходу з боку AlGaAs. На цьому малюнку також можна помітити початок утворення паралельного паразитного транзистора в шарі AlGaAs, проте струм, що протікає через нього, малий, оскільки в моделі немає омичного контакту до шару AlGaAs.

На рис. 7.38 представлені результати моделювання і експерименту стокових характеристик транзистора з довжиною затвора 0,35 мкм і шириною 300 мкм. Порівняно з подібними характеристиками арсенідгалійового ПТШ (рис. 7.35) вони відзначаються більшим струмом стоку і відповідно більшими значеннями крутизни.

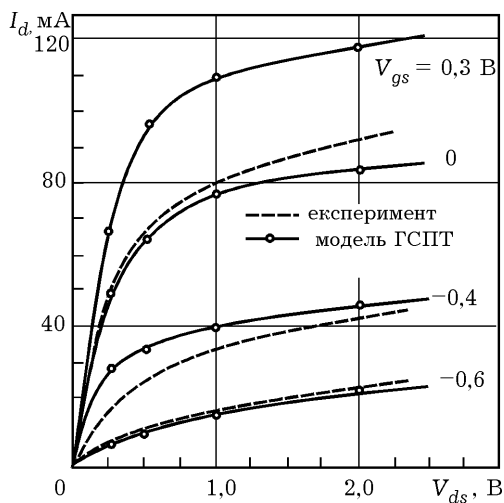


Рис. 7.38. Стокові характеристики транзистора на гетероструктурі $Al_{0,3}Ga_{0,7}As/GaAs$

Структури сучасних ГСПТ досить різноманітні, вони можуть істотно різнитися кількістю шарів, типом їх матеріалу, ступенем легування та іншими параметрам. Перспективними для використання вважаються нітриди галію та алюмінію. Технологічно освоєні структури на гетеропереходах GaN/Al_xGa_{1-x}N. У нелегованому GaN рухливість електронів провідності виявляється досить високою – близько 0,2 м²/(В·с) за кімнатної температури.

Розрив дна зони провідності пари GaN і Al_xGa_{1-x}N істотно більший, ніж в гетеропереході GaAs/Al_xGa_{1-x}As, і складає близько 0,8 еВ. Тому електростатичне поле в області контакту ще сильніше, а концентрація вільних електронів, що перейшли з Al_xGa_{1-x}N до потенціальної ями GaN, може бути істотно вищою. Крім того, в області контакту GaN і Al_xGa_{1-x}N через незбіжність періоду їх

кристалічних ґраток утворюється механічна напруга. А оскільки кристали GaN мають сильний ступінь поляризації, виникає п'єзоефект, стимулюючий додатковий перехід носіїв в канал. У результаті концентрація електронів провідності в області ДЕГ може перевищувати 10^{13} см⁻². Відзначимо, що п'єзоефект настільки великий, що його одного часто вистачає для забезпечення необхідної концентрації носіїв в каналі. Так, якщо в ГСПТ на основі GaAs бар'єрний шар обов'язково легують, щоб забезпечити джерело вільних електронів, то в разі нітриду галію можливо використовувати і нелегований бар'єрний шар Al_xGa_{1-x}N. Проте зазвичай бар'єрний шар в більшості GaN-структур легований. Поєднання високої рухливості і концентрації носіїв в каналі при широкій забороненій зоні (близько 3,4 еВ за кімнатної температури) обумовлює унікальні частотні та потужнісні властивості GaN ГСПТ, вони відзначаються більш високою робочою напругою і густиною струму стоку.

Контрольні запитання та вправи

7.1. Дайте визначення ефектові поля. Опишіть типовий вигляд залежності поверхневої провідності від поверхневого потенціалу для домішкових (донорних та акцепторних) напівпровідників і поясніть характерні розподіли поля та концентрації носіїв для різних режимів на цій кривій.

7.2. На прикладі контакту двох металів покажіть від чого залежить контактна різниця потенціалів. Чому вольт-амперна характеристика такого контакту є нелінійною?

7.3. Від яких фізичних процесів та відповідних параметрів залежить швидкість спадання надлишкової концентрації неосновних носіїв у разі їх інжекції у домішковий напівпровідник?

7.4. Опишіть рівноважний стан електронно-діркового переходу. Від чого залежить контактна різниця потенціалів та яка роль контактного поля у встановленні рівноважного стану?

7.5. Використовуючи співвідношення для товщин збіднених шарів, порівняйте їх значення для сильно несиметричного переходу та прослідкуйте розміщення переходу щодо межі напівпровідників.

7.6. Які процеси відбуваються при електричному зміщенні електронно-діркового переходу? Як впливає зовнішня напруга на розподіл потенціалу та концентрації в переході?

7.7. Опишіть вольт-амперну характеристику електронно-діркового переходу як результат порушення рівноважного стану між складовими струму.

7.8. Опишіть поведінку носіїв за межами збідненої області у разі прямого та зворотного зміщення. уточнить объект?

7.9. Опишіть причини, які визначають ємнісні властивості електронно-діркового переходу.

7.10. Опишіть типову структуру та властивості бар'єру Шоттки.

7.11. У яких контактах можливе утворення збагаченого шару, придатних для використання як омічних?

7.12. Від яких параметрів зонної структури напівпровідників залежать розриви дна зони провідності та стелі валентної зони. Які типи гетеропереходів можливі в залежності від положення цих розривів?

7.13. Опишіть характерну особливість структури та функціональні переваги гетеробіполярного транзистора.

7.14. Які особливості вольт-амперних характеристик субмікронних польового транзистора із затвором Шоттки?

7.15. Опишіть особливості утворення провідного каналу в гетероструктурному польовому транзисторі. Які функції виконує буферний шар?

Додаток до гл. 1

Д1. Модель Кроніга-Пенні

Модель Кроніга-Пенні є однією з повчальних моделей, що дозволяє провести деякі кількісні оцінки особливостей електронних станів у періодичному полі. Як модель потенціального рельєфу оберемо періодичну послідовність прямокутних бар'єрів, що чергуються, і ям (рис. Д1.1, а). За цих умов рівняння Шредінгера з періодичною функцією $U(x)$ може бути представлене у вигляді рівнянь з постійними коефіцієнтами для окремих областей:

I) в області ями $U = 0$

$$\frac{d^2\psi}{dx^2} + \alpha^2\psi = 0, \quad \alpha^2 \equiv 2mE / \hbar^2, \quad (\text{Д1.1})$$

II) в області бар'єру $U = U_0 > E$

$$\frac{d^2\psi}{dx^2} - \beta^2\psi = 0, \quad \beta^2 \equiv 2m(U_0 - E) / \hbar^2 \quad (\text{Д1.2})$$

Розв'язок цих рівнянь можна представити у вигляді

$$\begin{aligned} \psi_1(x) &= A\cos\alpha x + B\sin\alpha x, \\ \psi_2(x) &= C\text{ch}\beta x + D\text{sh}\beta x. \end{aligned} \quad (\text{Д1.3})$$

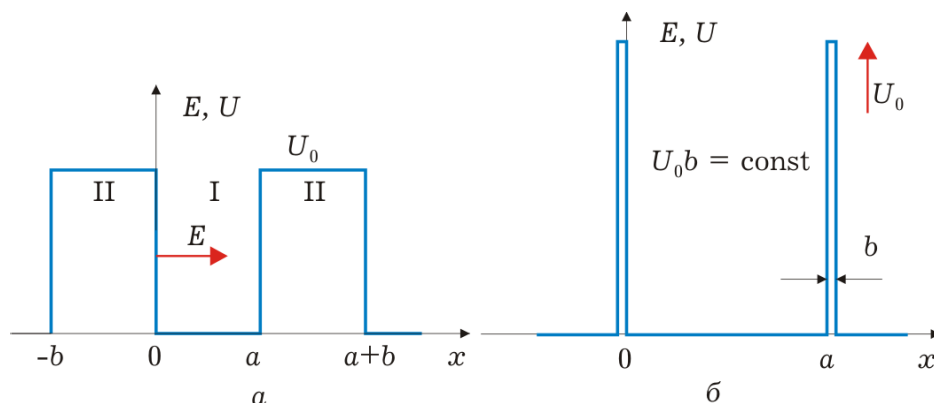


Рис. Д1.1. Періодичний потенціальний рельєф у моделі Кроніга-Пенні (а) та його граничне подання (б)

Але з іншого боку, у періодичному полі функція $\psi(x)$ повинна мати вигляд хвилі Блоха $\psi(x) = \varphi(x)e^{ikx}$, де амплітуда $\varphi(x)$ має бути періодичною $\varphi(x) = \varphi(x + a + \varepsilon)$. Тому

$$\varphi_1(x) = (A \cos \alpha x + B \sin \alpha x) e^{-ikx}, \quad (\text{Д1.4, а})$$

$$\varphi_2(x) = (C \operatorname{ch} \beta x + D \operatorname{sh} \beta x) e^{-ikx}. \quad (\text{Д1.4, б})$$

На межах областей ці функції і їх похідні мають бути неперервні. З умови $\varphi_1(0) = \varphi_2(0)$ отримуємо, що $A = C$, а з умови $\varphi_1'(0) = \varphi_2'(0)$ слідує $\alpha B = \beta D$. «Зшивання» на іншій межі $\varphi_1(a) = \varphi_2(-b)$, $\varphi_1'(a) = \varphi_2'(-b)$ відповідає умові періодичності амплітуд і призводить до співвідношень:

$$\beta(A \cos \alpha a + B \sin \alpha a) = (\beta A \operatorname{ch} \beta b - \alpha B \operatorname{sh} \beta b) e^{ik(a+b)}, \quad (\text{Д1.5, а})$$

$$\alpha(A \sin \alpha a - B \cos \alpha a) = (\beta A \operatorname{sh} \beta b - \alpha B \operatorname{ch} \beta b) e^{ik(a+b)}, \quad (\text{Д1.5, б})$$

які є системою рівнянь для знаходження констант A і B . Вона має розв'язок, якщо її визначник дорівнює 0, або

$$\frac{\beta^2 - \alpha^2}{2\alpha\beta} \sin \alpha a \cdot \operatorname{sh} \beta b + \cos \alpha a \cdot \operatorname{ch} \beta b = \cos k(a+b). \quad (\text{Д1.6})$$

Цей вираз є *дисперсійним рівнянням*, розв'язок якого дозволяє знайти *закон дисперсії*, тобто зв'язок між енергією електрона, що входить в параметри α і β , і хвильовим числом k .

Спростити рівняння (Д1.6) можна без втрати загальності результатів, збільшуючи висоту бар'єрів ($U_0 \rightarrow \infty$) і одночасно їх звужуючи ($b \rightarrow 0$), але так, щоб добуток $U_0 b$, який характеризує їх прозорість, залишався постійним (рис. Д1.1, б). За таких умов $\beta b \rightarrow 0$ і дисперсійне рівняння можна представити у формі

$$\cos ka = f(\alpha(E)) \equiv P(\sin \alpha a / \alpha a) + \cos \alpha a, \quad (\text{Д1.7})$$

де $P \equiv tabU_0 / \hbar^2$. Уявлення про особливості розв'язку цього рівняння можна отримати з рис. Д1.2, де зображена функція $f(\alpha(E))$ для $P = 4$. Оскільки $1 \geq \cos ka \geq -1$, то дійсні рішення можливі лише у відмічених областях значень ka .

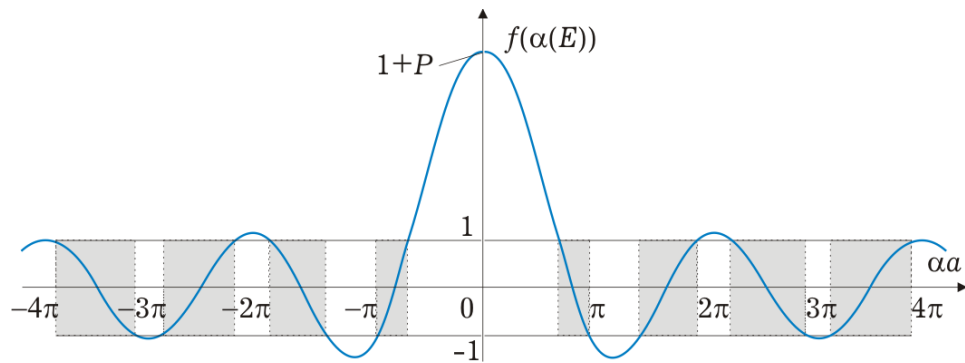


Рис. Д1.2. Графічне визначення областей розв'язку трансцендентного рівняння (Д1.7)

Головним висновком з цієї моделі є твердження про те, що енергетичний спектр електронів є областями (зонами) дозволених і заборонених станів, що чергуються. В інтервалах, де $|f(\alpha(E))| \leq 1$, існують дійсні значення k , яким відповідають розв'язки рівняння Шредінгера у вигляді блохівських хвиль. У проміжках між ними розв'язки (Д1.7) будуть уявними, а хвильові функції стають експоненціально зростаючими і не можуть задовольнити умові нормування для нескінченно довгого кристала. Тому такі стани утворюють «заборонені» зони.

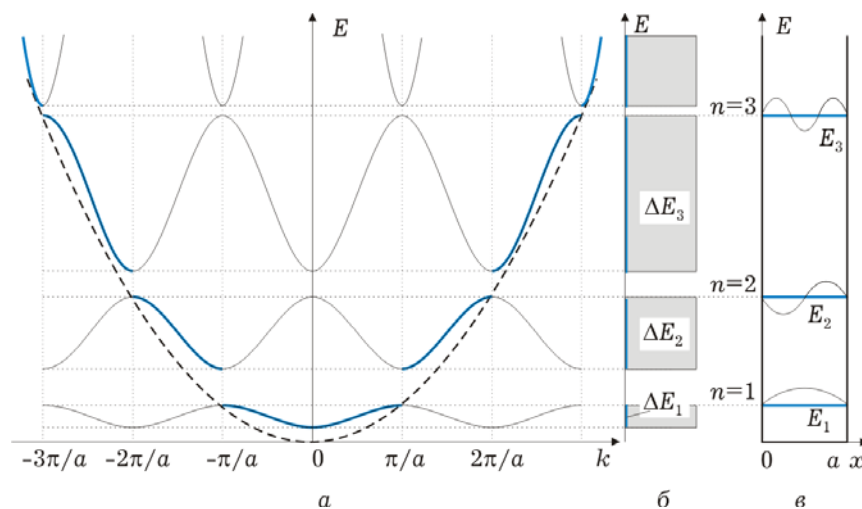


Рис. Д1.3. Дисперсійні криві у моделі Кроніґа-Пенні (а). Верхні краї дозволених зон (б) збігаються з рівнями квантування у нескінченно глибокій ямі (в)

Більш повно результати моделі Кроніґа-Пенні представлені на рис. Д1.3, а у вигляді дисперсійних кривих $E(k)$ для кожної дозволеної енергетичної зони (б). Наявність в (Д1.7) функції $\cos ka$

призводить до того, що, змінюючи значення k на ціле число $2\pi/a$, набудемо тих самих значень енергії. Це означає, що дисперсійні криві для кожної дозволеної зони описуються періодичними і парними функціями. Верхні краї зон відповідають значенням $\alpha a = n\pi$ ($n = 1, 2, \dots$). За такої умови $\cos ka = 1$, $k = n\pi/a$, а значення енергії $E_n = n^2 \hbar^2 \pi^2 / 2ma^2$ будуть такими ж, як і у нескінченно глибокій потенціальній ямі шириною a (рис. Д1.3, в). У цих станах електрони описуються хвильовими функціями $\sin(n\pi x/a)$, характерними для стоячих хвиль. Значенням $\cos ka = -1$ відповідають нижні краї дозволених зон і хвильові функції типу $\cos(n\pi x/a)$, що також описують стоячі хвилі. Це пояснюється тим, що за такої умови на періоді ґратки a вкладається ціле число півхвиль $a = n\lambda/2$ і відбиті від кожного бар'єру хвилі будуть синфазно складатися. Серединам дозволених зон відповідають випадки, коли $f(\alpha(E))$ проходить через 0, а $ka = (2n - 1)\pi/2$, тобто $a = (2n - 1)\lambda/4$ і на довжині a вкладається непарне число чвертей довжини хвиль. У результаті відбиті від бар'єрів хвилі, будучи протифазами, взаємно гасяться і хвильові функції матимуть вигляд $\exp(ikx)$, характерний для хвиль, що біжать.

На рис. Д1.3 штриховою лінією зображена також параболічна дисперсійна крива для вільного електрона. Якщо прозорість бар'єрів нескінченно велика ($P = 0$), то (П1.7) набирає вигляду $\cos ka = \cos \alpha a$. У такому разі дозволені будь-які значення $k = \alpha$ і $E = \hbar^2 k^2 / 2m$, як і для електрона у вільному просторі.

Д2. Інтегралі Фермі-Дірака

Найбільш зручна і така, що найчастіше зустрічається, форма представлення інтегралів Фермі-Дірака має вигляд

$$\Phi_n(\xi) = \frac{1}{\Gamma(n+1)} \int_0^\infty \frac{\varepsilon^n d\varepsilon}{\exp(\varepsilon - \xi) + 1}. \quad (\text{ПЗ.1})$$

Тут n - порядок інтеграла, а $\Gamma(n+1)$ - так звана гамма-функція, яка є узагальненням поняття факторіалу $\Pi(n) = n! = \Gamma(n+1)$ для нецілих значень n . Для цілих додатніх значень n , як завжди, $\Gamma(n+1) = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$; для $n = 3/2$ - $\Gamma(5/2) = 3\sqrt{\pi}/4$; для $n = 1/2$ - $\Gamma(3/2) = \sqrt{\pi}/2$; для $n = -1/2$ - $\Gamma(1/2) = \sqrt{\pi}$. Корисною є властивість гамма-функції: $\Gamma(n+1) = n \cdot \Gamma(n)$, а також властивість інтеграла Фермі-Дірака: $d\Phi_{n+1}(\xi)/d\xi = \Phi_n(\xi)$.

Графіки інтегралів Фермі-Дірака для деяких значень n , що розглядаються у гл. 3, представлені на рис. Д3.1. З графіків видно, що в області досить великих від'ємних значень ξ інтегралі Фермі-Дірака для усіх порядків достатньо добре апроксимуються функцією e^ξ .

Іноді зручно використовувати асимптотичні вирази для цього інтеграла:

$$\Phi_{1/2}(\xi) = \begin{cases} e^\xi, & -\infty < \xi < -1, \quad \text{а)} \\ e^\xi / (0,27 e^\xi + 1), & -1 < \xi < 5, \quad \text{б)} \\ 4\xi^{3/2} / 3\sqrt{\pi}, & 5 < \xi < \infty. \quad \text{в)} \end{cases}$$

Формула (а) придатна для невідроджених напівпровідників, а (в) для сильно вроджених напівпровідників і металів. Формулу (б) можна використовувати для розрахунків як невідроджених колективів, де вона дає вищу точність, ніж (а), так і за слабкого вродження.

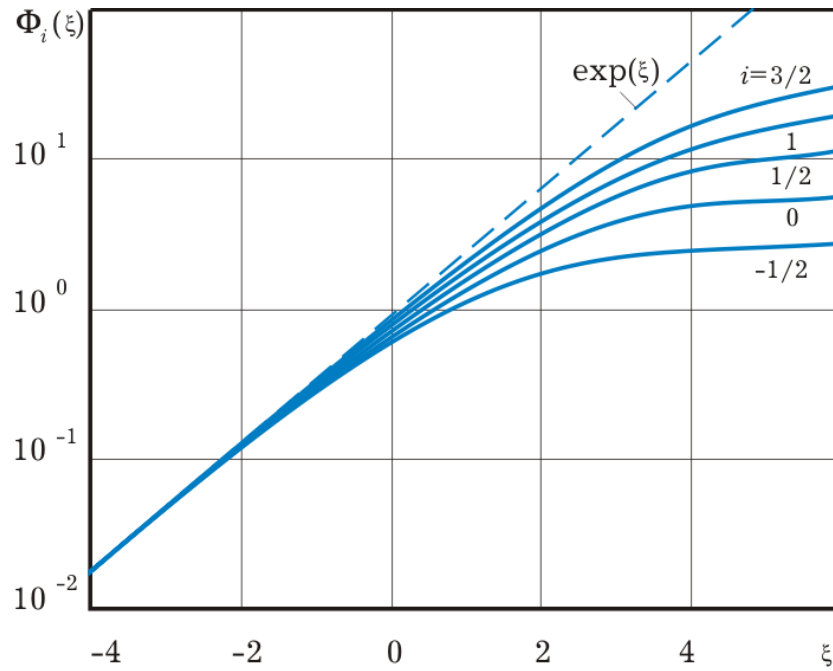


Рис. ДЗ.1. Залежність інтегралів Фермі-Дірака n -порядку від приведенного рівня Фермі. Для невироджених колективів зручна їх апроксимація функцією $\exp(\xi)$

Детальніше з властивостями інтегралів Фермі-Дірака можна ознайомитися, наприклад, в книзі: Дж. Блекмор, «Статистика електронів в напівпровідниках», пер. з англ., видавництво «Мир», Москва, 1964 р., 392 с. Там же наведено таблиці значень функцій $\Phi(\xi)$.

Д3. Надбар'єрне проходження електронів

Найпростіший прояв надбар'єрного проходження має місце під час термоелектронної емісії. Сутність цього явища полягає в тому, що підвищення температури викликає збільшення теплових коливань кристалічної решітки і, відповідно, теплової енергії електронів. Попри те, що середнє значення теплової енергії електронів складає десятки долі електрон-вольта за температури у тисячі кельвінів, певна, хоч і дуже мала, кількість електронів матиме енергію, яка у багато разів вища за середню і може стати достатньою для подолання сил, що заважають виходу електронів із речовини. Хоча таких електронів дуже мало, але саме вони обумовлюють процес емісії. За своєю фізичною суттю цей процес схожий на процес випаровування молекул з рідини.

На рис. Д3.1 з енергетичною діаграмою зони провідності поблизу поверхні поєднана енергетична функція розподілу електронів. З неї видно, що за температури абсолютного нуля всі електрони мають енергію, нижчу за рівень Фермі.

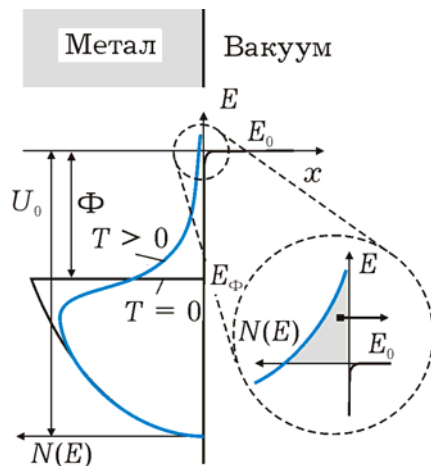


Рис. Д3.1. Енергетична діаграма та функція розподілу електронів біля поверхні металу. З підвищенням температури частина електронів може отримати енергію, достатню для подолання потенціального бар'єру, і матиме можливість виходу

Електрони з енергією, більшою за повну роботу виходу U_0 , тобто в станах, вищих за рівень вакууму E_0 , з'являються з підвищенням температури. Їх кількість пропорційна затушованій області і зростає з нагріванням. Однак не всі електрони, які мають енергію $E_0 > U_0$ приймають участь в емісії. Необхідно, щоб кінетична енергія їх руху в

напрямку поверхні перевищувала повну роботу виходу: $mv_x^2/2 \geq U_0$, або

$$v_x \geq \sqrt{2U_0 / m} \equiv v_{x \min}. \quad (\text{Д3.1})$$

Значення паралельних до поверхні складових швидкості v_y та v_z можуть бути довільними:

$$-\infty < v_y, v_z < \infty. \quad (\text{Д3.2})$$

Оскільки умови виходу (Д3.1) та (Д3.2) формулюють для значень швидкості, то кількість електронів, що можуть приймати участь в емісії, не може бути визначена через енергетичну функцію розподілу. Для цього треба скористатись розподілом по складовим швидкості. Розглянемо групу з dn електронів, які знаходяться в елементарному об'ємі $dv_x dv_y dv_z$ простору швидкостей біля вектора \vec{v} , тобто які мають швидкості в межах від v_x до $v_x + dv_x$, від v_y до $v_y + dv_y$ та від v_z до $v_z + dv_z$ (рис. Д3.2, а).

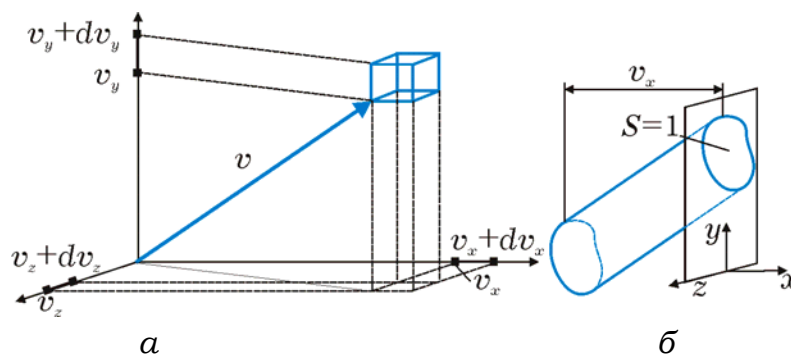


Рис. Д3.2. До виведення закону надбар'єрного проходження електронів
Для розрахунку густини струму емісії біля поверхні емітера обирається
циліндр з висотою, що чисельно дорівнює швидкості v_x , і твірними,
направленими так само, як і вектор швидкості обраної групи електронів

Для розрахунку елементарного струму dj_s , що створюється цією групою електронів, оберемо на поверхні (рис. Д3.2, б) циліндричний об'єм з одиничною основою, висотою v_x , а напрям твірних циліндра має співпадати з напрямом вектора швидкості на рис. Д3.2, а. За такого вибору всі електрони, що знаходяться у циліндрі, покинуть цей об'єм за одиницю часу (секунду), проходячи через обрану на поверхні площадку. Вони і утворять струм dj_s , рівний добутку заряду

електрона на об'єм, чисельно рівний v_x , і на концентрацію цієї групи електронів

$$dj_s = ev_x dn. \quad (\text{Д3.3})$$

Кількість електронів в об'ємі $dv_x dv_y dv_z$ визначається концентрацією станів dn_c і ймовірністю їх заповнення $f(v_x, v_y, v_z)$:

$$dn = dn_c f(v_x, v_y, v_z) = g(\vec{v}) f(v_x, v_y, v_z) dv_x dv_y dv_z. \quad (\text{Д3.4})$$

Густина електронних станів $g(\vec{v})$ у просторі швидкостей постійна і може бути знайдена з густини $g(\vec{k})$ у \vec{k} -просторі (1.10, а) з урахуванням співвідношень $\vec{p} = \hbar\vec{k} = m\vec{v}$. Тому

$$dn_c = g(\vec{v}) dv_x dv_y dv_z = (2m^3/h^3) dv_x dv_y dv_z. \quad (\text{Д3.5})$$

Ймовірність заповнення станів $f(v_x, v_y, v_z)$ задається функцією Фермі – Дірака з заміною $E = m(v_x^2 + v_y^2 + v_z^2)/2$. Відповідно

$$dj_s = \frac{2e(m/h)^3 v_x dv_x dv_y dv_z}{1 + \exp \left[\left(\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T} \right) - \frac{E_\Phi}{k_B T} \right]}. \quad (\text{Д3.6})$$

Повний струм, що створюється всіма електронами, знаходиться інтегруванням (Д3.6) по швидкостях у межах, які задані (Д3.1) і (Д3.2):

$$j_s = \frac{2em^3}{h^3} \int_{-\infty}^{\infty} dv_z \int_{-\infty}^{\infty} dv_y \int_{v_{x \min}}^{\infty} dv_x \frac{v_x dv_x}{1 + \exp \left[\left(\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T} \right) - \frac{E_\Phi}{k_B T} \right]}. \quad (\text{Д3.7})$$

Переходячи до полярних координат у просторі швидкостей v_x і v_z , що досягається заміною $v_y = v_\perp \cos \varphi$ і $v_z = v_\perp \sin \varphi$, та використовуючи табличний інтеграл типу

$$\int \frac{dx}{1 + a e^x} = \ln \frac{e^x}{1 + a e^x},$$

можна отримати

$$j_s = \frac{4\pi m e k_B T}{h^3} \int_{U_0}^{\infty} \ln \left[1 + \exp \left(\frac{E_{\Phi} - E_x}{k_B T} \right) \right] dE_x, \quad (\text{ДЗ.8})$$

де $E_x = m v_x^2 / 2$.

Подальше спрощення можливе для достатньо високих бар'єрів, які мають місце у разі більшості термоелектронних катодів. Навіть для найменших значень швидкості $v_{x \min}$ з (ДЗ.1) аргументом експоненти в (ДЗ.8) буде величина $(E_{\Phi} - U_0) / k_B T = -\Phi / k_B T$, яка для реальних ситуацій ($\Phi > 1$ еВ, $T \approx 1000$ К) за абсолютним значенням на порядок більша за 1. Це дає можливість вважати експоненту малою величиною і використати наближення $\ln(1 + x) \approx x$. Таке припущення означає нехтування одиницею у знаменнику (ДЗ.7), а це є формальною ознакою того, що та частина електронів, яка приймає участь у термоелектронній емісії, достатньо добре описується статистикою Максвелла – Больцмана і є невиродженим електронним газом. Використане припущення зводить останній інтеграл до елементарного і приводить до остаточного результату

$$j_s = AT^2 \exp(-\Phi / k_B T), \quad (\text{ДЗ.9})$$

де $A \equiv 4\pi m e k_B^2 / h^3 = 1,2 \times 10^6$ А/(м²·К²). Вираз (ДЗ.9) звать формулою Річардсона – Дешмана, вона описує основний закон термоелектронної емісії.

В електронній техніці зустрічаються також ситуації, коли висота бар'єрів, які мають подолати електрони, складає десятки частки електрон-вольта. У такому разі останнє припущення може бути неприйнятним і для обчислення інтегралу в (ДЗ.8) доводиться використовувати чисельні методи.

Рекомендована література

1. Физикотехнический институт им. А.Ф.Иоффе Российской Академии Наук, New Semiconductor Materials. Characteristics and Properties URL: <http://www.ioffe.ru/SVA/NSM/>
2. Бонч-Бруевич В.Л., Калашников С.Г. Физика полупроводников.- М.: Наука, 1977. – 367с.
3. Конуэлл Э. Кинетические свойства полупроводников в сильных электрических полях: Пер. с англ. – М.: Мир, 1970. – 384 с.
4. Шур М. Современные приборы на основе арсенида галлия: Пер. с англ. – М.: Мир, 1991. – 632 с.
5. Прохоров Э. Д., Белецкий Н. И. Полупроводниковые материалы для приборов с междолинным переносом. – Х.: Вища шк.; Изд. при Харьк. ун-те, 1982. – 144 с.
6. Иващенко В. М., Митин В. В. Моделирование кинетических явлений в полупроводниках. Метод Монте-Карло. – К.: Наук. думка, 1990. – 192 с.]
7. Зи С. Физика полупроводниковых приборов: В 2 кн.: Пер. с англ. – М.: Мир, 1984. – Кн. 1. – 456 с.
8. Москалюк В.О., Тимофеев В.І., Федяй А.В. «Надшвидкодiючі прилади електронiки», навч. пос. з грифом МОНУ, вид. «Полiтехнiка», Киiв, 2014, С.528.
9. Дж. Блекмор, «Статистика электронов в полупроводниках», пер. з англ., видавництво «Мир», Москва, 1964 р., 392 с.
10. Герасимов С.М., Белоус М.В., Москалюк В.А. «Физические основы электронной техники», учебное пособие с грифом Минвуза СССР, Киiв, "Вища школа", 1981, 368с.
11. Москалюк В.О., Фiзика електронних процесiв. Динамiчнi процеси., навч. посiбник з грифом МОН України Киiв, вид. „Полiтехнiка”, 2004, 180 с.