

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Прикарпатський національний університет імені Василя Стефаника
Національний університет “Львівська політехніка”
Івано-Франківський національний технічний університет нафти і газу

Фізика процесів у напівпровідниках та елементах електроніки

КУРС ЛЕКЦІЙ

Івано-Франківськ
2010

УДК 621.382.8 (075.8)
ББК 22.373.1
Ф50

*Рекомендовано до друку вченою радою
Прикарпатського національного університету імені Василя Стефаника
(протокол № 5 від 25 травня 2010 р.)*

Автори:

Д.М. Фреїк, В.М. Чобанюк, З.Ю. Готра, Б.С. Дзундза, Г.Д. Матеїк, А.І. Ткачук

Рецензенти:

Раренко І.М., доктор фізико-математичних наук, професор,
Чернівецький національний університет ім. Ю. Федьковича;

Дружинін А.О., доктор технічних наук, професор
Національний університет “Львівська політехніка”;

Стасюк З.В., доктор фізико-математичних наук, професор,
Львівський національний університет імені Івана Франка

Фізика процесів у напівпровідниках та елементах електроніки : курс лекцій : [навчальний посібник] / [Д. М. Фреїк, В. М. Чобанюк, З. Ю. Готра та ін. ; за заг. ред. заслуженого діяча науки і техніки України, доктора хімічних наук, проф. Д. М. Фреїка]. – Івано-Франківськ : Видавництво Прикарпатського національного університету імені Василя Стефаника, 2010. – 263 с.

ISBN 978-966-640-275-5

Розглянуто основні фізичні процеси у напівпровідниках, які є важливими для конструювання і функціонування активних елементів електроніки. Значну увагу приділено фізичним процесам роботи діодів, транзисторів, елементів інтегральних схем магніторезисторів, магнітодіодів, сенсорів вологості, тензосенсорів, світлодіодів, фоторезисторів. Крім того, окремо розглянуто нові досягнення у нанотехнологіях наноматеріалів та наноелектроніки і створення сонячних елементів.

Для студентів, аспірантів інженерно-технічних працівників, науковців, що спеціалізуються з проблем фізики напівпровідників та напівпровідникової електроніки.

УДК 621.382.8 (075.8)
ББК 22.373.1

ISBN 978-966-640-275-5

© Фізико-хімічний інститут Прикарпатського національного університету імені Василя Стефаника, 2010

© Видавництво Прикарпатського національного університету імені Василя Стефаника, 2010

ЗМІСТ

| | |
|----------------------------------------------------------------------------------|----|
| ПЕРЕДМОВА | 7 |
| ВСТУП | 8 |
| Лекція 1. СТРУКТУРА ЕНЕРГЕТИЧНИХ ЗОН | 10 |
| 1.1. Утворення енергетичних зон..... | 10 |
| 1.2. Енергетичні зони в твердих тілах..... | 11 |
| 1.3. Електрони і дірки..... | 13 |
| 1.4. Генерація і рекомбінація носіїв заряду..... | 14 |
| 1.5. Домішкові енергетичні рівні..... | 15 |
| Лекція 2.СТАТИСТИКА НОСІЇВ ЗАРЯДУ | 18 |
| 2.1. Густина енергетичних рівнів. Функція розподілу Фермі-Дірака..... | 18 |
| 2.2. Час життя нерівноважних носіїв заряду..... | 23 |
| 2.3. Температурні залежності концентрації носіїв заряду..... | 25 |
| 2.4. Температурні залежності рухливості носіїв заряду і питомої провідності..... | 26 |
| Лекція 3. ЯВИЩА ПЕРЕНОСУ | 28 |
| 3.1. Рухливість носіїв заряду..... | 28 |
| 3.2. Електропровідність..... | 31 |
| 3.3. Квантові ефекти..... | 34 |
| Лекція 4. ТЕРМОЕЛЕКТРИЧНІ ЯВИЩА | 38 |
| 4.1. Ефект Зеебека..... | 38 |
| 4.2. Контактна складова термо-ЕРС..... | 40 |
| 4.3. Об'ємна складова термо-ЕРС..... | 41 |
| 4.4. Ефект Пельт'є..... | 45 |
| Лекція 5. ГАЛЬВАНОМАГНІТНІ ЕФЕКТИ | 48 |
| 5.1. Ефект Холла..... | 48 |
| 5.2. Магніторезистивний ефект..... | 51 |
| 5.3. Магнітоконцентраційний ефект..... | 54 |
| 5.4. Ефект Фарадея..... | 55 |
| 5.5. Ефект Фогта..... | 58 |
| 5.6. Термомагнітні ефекти в напівпровідниках..... | 59 |
| Лекція 6. ПРОЦЕСИ ДИФУЗІЇ | 61 |
| 6.1. Дифузійні рівняння..... | 61 |
| 6.2. Рівняння Айнштайна..... | 63 |
| Лекція 7. ПРОЦЕСИ ПОГЛИНАННЯ СВІТЛА | 65 |
| 7.1. Власне поглинання світла..... | 65 |

| | |
|----------------------------------------------------------------------|------------|
| 7.2. Домішкове поглинання світла..... | 68 |
| 7.3. Екситонне поглинання світла..... | 69 |
| 7.4. Поглинання світла вільними носіями заряду..... | 70 |
| 7.5. Поглинання світла кристалічною ґраткою..... | 70 |
| Лекція 8. ЯВИЩЕ ФОТОПРОВІДНОСТІ..... | 72 |
| 8.1. Квантовий вихід фотоэффекту..... | 72 |
| 8.2. Релаксація фотопровідності і стаціонарна фотопровідність..... | 73 |
| 8.3. Ефект Дембера..... | 76 |
| 8.4. Вентильна фото-ЕРС..... | 78 |
| 8.5. Фотомагнітоелектричний ефект (ефект Кікоїна-Носкова)..... | 79 |
| Лекція 9. ФІЗИЧНІ ПРОЦЕСИ В ПОВЕРХНЕВИХ ШАРАХ..... | 83 |
| 9.1. Поверхневі стани..... | 83 |
| 9.2. Приповерхневий просторовий заряд..... | 85 |
| 9.3. Поверхнева провідність і ефект поля..... | 90 |
| 9.4. Поверхнева рекомбінація..... | 92 |
| Лекція 10 КОНТАКТНІ ЯВИЩА В НАПІВПРОВІДНИКАХ..... | 94 |
| 10.1 Контакт напівпровідник-напівпровідник..... | 94 |
| 10.2 Нерівноважний стан переходу..... | 99 |
| 10.3. Вольт-амперна характеристика і ємність переходу..... | 101 |
| 10.4. Неінжектуючі контакти, тунельні переходи і гетеропереходи..... | 104 |
| 10.4.1. Неінжектуючі контакти..... | 104 |
| 10.4.2. Тунельні переходи..... | 105 |
| 10.4.3. Гетеропереходи..... | 106 |
| 10.5. Контакти..... | 107 |
| 10.5.1. Контакт метал-напівпровідник..... | 107 |
| 10.5.2. Випрямний контакт..... | 108 |
| 10.5.3. Омичний контакт..... | 110 |
| 10.5.4. Контакт метал-метал..... | 110 |
| Лекція 11. ДІОДИ І ТРАНЗИСТОРИ..... | 112 |
| 11.1. Напівпровідникові діоди..... | 112 |
| 11.2. Біполярні транзистори..... | 115 |
| 11.3. Модель Еберса-Молла..... | 120 |
| Лекція 12. ЕЛЕМЕНТИ ІНТЕГРАЛЬНИХ СХЕМ..... | 126 |
| 12.1. Активні елементи інтегральних схем..... | 126 |
| 12.2. Транзистор системи метал-діелектрик-напівпровідник..... | 127 |
| 12.3. Польові транзистори та прилади з зарядовим зв'язком..... | 131 |
| Лекція 13. МАГНІТОРЕЗИСТОРИ..... | 135 |
| 13.1. Загальна характеристика магніtoresисторів..... | 135 |
| 13.2. Матеріали для магніtoresисторів..... | 136 |

| | |
|----------------------------------------------------------------------------------|------------|
| 13.3. Магніторезистори на основі магнітоконцентраційного ефекту..... | 137 |
| 13.4. GMR-магніторезистори..... | 138 |
| 13.5. Параметри магніторезисторів..... | 139 |
| Лекція 14. МАГНІТОДІОДИ..... | 142 |
| 14.1. Загальна характеристика магнітодіодів..... | 142 |
| 14.2. Матеріали для магнітодіодів | 144 |
| 14.3. Торцеві магнітодіоди | 145 |
| 14.4. Планарні магнітодіоди | 146 |
| 14.5. Планарні магнітодіоди з металевим контактом | 148 |
| 14.6. Полярні магнітодіоди..... | 149 |
| 14.7. Магнітодіоди з ефектами порогового перемикання і "пам'яті"..... | 151 |
| 14.8. Інтегральні магнітодіоди..... | 152 |
| Лекція 15. ПЛІВКОВІ НАПІВПРОВІДНИКОВІ СЕНСОРИ ВОЛОГОСТІ..... | 155 |
| 15.1. Принцип роботи сенсорів вологості..... | 155 |
| 15.2. Тонкоплівкові алюмінієво-оксидні сенсори вологості..... | 156 |
| 15.3. Плівкові сенсори для вимірювання точки роси..... | 159 |
| Лекція 16. НАПІВПРОВІДНИКОВІ ТЕНЗОСЕНСОРИ..... | 163 |
| 16.1. Напівпровідникові тензосенсори..... | 163 |
| 16.1.1. Конструкція інтегральних тензосенсорів..... | 163 |
| 16.1.2. Інтегральний сенсор тиску на основі біполярних тензотранзисторів..... | 164 |
| 16.1.3. Тензорезистивні плоскі сенсори відносного і абсолютного тиску.. | 165 |
| 16.1.4. Інтегральний балочний акселерометр..... | 166 |
| 16.2. Напівпровідникові сенсори газу резистивного типу..... | 166 |
| 16.2.1. Матеріали чутливих елементів сенсорів..... | 166 |
| 16.2.2. Конструкції сенсорів..... | 170 |
| 16.2.3. Технологія виготовлення сенсорів..... | 174 |
| 16.2.4. Тонкоплівкові сенсори..... | 176 |
| 16.2.5. Товстоплівкові сенсори..... | 181 |
| Лекція 17. АКТИВНІ НАПІВПРОВІДНИКОВІ ФОТОЕЛЕКТРИЧНІ ЕЛЕМЕНТИ..... | 192 |
| 17.1. Світло діоди..... | 192 |
| 17.1.1. Основні параметри та характеристики світлодіодів..... | 193 |
| 17.1.2. Принцип роботи інжекційних світлодіодів з р-п-переходами..... | 194 |
| 17.1.3. Світлодіоди з антистоксовими люмінофорами..... | 197 |
| 17.2. Напівпровідникові лазери..... | 198 |
| 17.2.1. Лазери на гомоструктурах..... | 201 |
| 17.3. Фоторезистори..... | 202 |
| 17.3.1. Принцип дії фоторезисторів..... | 202 |
| 17.3.2. Основні параметри фоторезисторів..... | 203 |

| | |
|-------------------------------------------------------|-----|
| 17.3.3. Конструкція фоторезисторів..... | 206 |
| 17.3.4. Области використання фоторезисторів..... | 207 |
| 17.4. Фотодіоди..... | 209 |
| 17.4.1. Класифікація фотодіодів..... | 209 |
| 17.4.2. Р-п-фотодіоди..... | 211 |
| 17.4.3. Р-і-п-фотодіоди..... | 214 |
| 17.4.4. Матеріали і конструкція р-і-п-фотодіодів..... | 215 |
| 17.4.5. Лавинні фотодіоди..... | 215 |
| 17.4.6. Високочастотні фотодіоди..... | 219 |
| 17.4.7. Інжекційні фотодіоди..... | 222 |
| 17.4.8. S-фотодіоди..... | 223 |
| 17.4.9. Використання фотодіодів..... | 223 |

**Лекція 18. НАПІВПРОВІДНИКОВІ НАНОМАТЕРІАЛИ,
НАНОТЕХНОЛОГІЇ ТА НАНОЕЛЕКТРОНІКА.....** 225

| | |
|----------------------------------------------------------------|-----|
| 18.1. Фізичні основи квантово-розмірних структур..... | 226 |
| 18.2. Методи отримання наноструктурованих матеріалів..... | 229 |
| 18.2.1. Літографія..... | 229 |
| 18.2.2. Молекулярно-променева епітаксія..... | 229 |
| 18.2.3. Осадження наноструктур із металоорганічних сполук..... | 234 |
| 18.2.4. Іонне бомбардування та імплантація..... | 236 |
| 18.2.5. Молекулярне нашарування – метод хімічної зборки..... | 237 |
| 18.2.6. Методи колоїдної хімії..... | 238 |
| 18.2.7. Осадження газодинамічного потоку пари..... | 239 |
| 18.3. Штучні атоми і кристали..... | 244 |
| 18.4. Елементи наноелектроніки..... | 246 |
| 18.4.1. Наноласери..... | 247 |
| 18.4.2. Одноелектронні транзистори..... | 247 |
| 18.4.3. Одноелектронне тунелювання..... | 248 |

**Лекція 19. СОНЯЧНІ ЕЛЕМЕНТИ НА ОСНОВІ
СЕНСИБІЛІЗОВАНИХ БАРВНИКАМИ НАПІВПРОВІДНИКІВ.....** 250

| | |
|------------------------------------------------------------------------|-----|
| 19.1. Принцип дії барвникових-сенсibilізованих сонячних елементів..... | 250 |
| 19.2. Матеріали для барвникових-сенсibilізованих сонячних елементів.. | 253 |
| 19.3. Барвник – сенсibilізатор..... | 255 |

ПЕРЕДМОВА



Створення цього курсу лекцій – давня мрія колективу авторів і насущна потреба для забезпечення навчального процесу із ряду дисциплін як на фізичних, так і технічних спеціальностях університетів, пов'язаних із фізикою напівпровідників та напівпровідниковою електронікою.

Що стосується Прикарпатського національного університету імені Василя Стефаника, то слід зазначити, що початки вивчення фізики напівпровідників та їх прикладних аспектів пов'язані із новими напрямками спеціалізації спочатку на кафедрі фізики, а із створенням університету – на кафедрах фізики твердого тіла, фізики і хімії твердого тіла та радіофізики і електроніки, завдяки ініціативам у першу чергу проф. Кланічки В.М., проф. Фреїка Д.М., проф. Рувінського М.А., проф. Новосядлого С.П. Фактично менш як за 20 років створені не тільки наукові школи із фізико-хімічних основ напівпровідникового матеріалознавства та напівпровідникової електроніки, але й започатковані зовсім нові напрямки для спеціалізації студентів, які стали визнаними на Україні, та за її межами. Останнє і спонукало нас (Д.М. Фреїк, В.М. Чобанюк, Б.С. Дзундза, А.І. Ткачук) ініціювати підготовку цього матеріалу.

Створена нова кафедра фізики новітніх технологій (проф. Галушак О.М. – завідувач кафедри, доценти Матеїк Г.Д., Солонисний Я.В., Рувінський Б.М. – вихованці наукової школи Прикарпатського університету) в Івано-Франківському національному технічному університеті нафти і газу стала осередком впровадження у навчальний процес інженерних спеціальностей прикладних досягнень фізики напівпровідників.

Необхідно особливо відзначити великий досвід підготовки спеціалістів із цих напрямків у Національному університеті “Львівська політехніка”, які реалізуються під загальним керівництвом проф. Готри З.Ю. Якраз їх доробок, за доброю згодою, використаний при підборі матеріалів до курсу лекцій. В його основу покладені наступні видання:

Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.

Мікроелектронні сенсори фізичних величин: Науково-навчальне видання. В трьох томах. / [Вуйцік В., Готра З.Ю., Григор'єв В.В. та ін.]; за ред. З.Ю. Готри. – Львів: Ліга-Прес, 2002.

Є надія, що зусилля авторів, які вкладені при підборі матеріалу та підготовці його до видання, принесуть користь як студентам, так і їх викладачам і будуть позитивно оцінені спеціалістами.

Дмитро Фреїк

*м. Івано-Франківськ
14 травня 2010р.*

ВСТУП

Підвищена увага в останнє десятиліття до вивчення напівпровідників та структур на їх основі викликано об'єктивними причинами: подальший розвиток напівпровідникової електроніки привело до пошуку принципіально нових шляхів розв'язання проблем, пов'язаних із покращенням основних параметрів напівпровідникових приладів (потужності, швидкодії та ін.). Гетероструктури, поряд з традиційними способами керування типом провідності, рівнем легування та кількістю *p-n*-переходів, відкрили можливість впливати всередині самої структури на найважливіші характеристики напівпровідника – шириною забороненої зони, а значить, і на оптичні константи, тим самим, націлено покращувати основні параметри та характеристики майже всіх напівпровідникових приладів, і в ряді випадків створювати принципово нові.

В основі роботи більшості напівпровідникових структур лежить явище інжекції. Тому створення напівпровідникових пристроїв на їх основі стало можливим після отримання ефективно інжектуючих систем GaAs – AlAs. Саме ця гетероструктура сприяла реалізації багатьох цікавих пропозицій щодо їх використання у різних приладах, дала поштовх інтенсивної розробки принципово нових пристроїв, вказала на шляхи пошуку та реалізації нових ідеальних гетеропереходів. Створення гетеропереходів, близьких за своїми властивостями до ідеальних, дозволило приступити до дослідження специфічних особливостей електричних та оптичних явищ, обумовлених особливостями їх зонної енергетичної діаграми, намітити шляхи використання в напівпровідникових приладах. Таким чином, фундаментальні дослідження гетеропереходів призвели до народження нового напрямку в напівпровідниковій техніці, і знайшли широке використання у лазерах і світлодіодах, перетворювачах сонячної енергії і фотодетекторах, біполярних і польових транзисторах, діодах і тиристорах, сенсорах фізичних величин.

Сьогодні напівпровідники широко використовуються для потреб радіоелектроніки і обчислювальної техніки як дискретні елементи, а також як компоненти аналогових і цифрових мікросхем. Крім того, їхні властивості такі, що залежно від вибраної конструкції і режиму роботи, можуть виконувати функції як активних, так і пасивних компонентів схем. Отже, існує можливість побудови складних електронних схем з використанням напівпровідникових структур, що особливо важливо для мікро- та наноелектроніки.

Через наявність високого вхідного опору польові транзистори широко застосовують в операційних і електрометричних підсилювачах. Розширюється їх використання в низькочастотних малощумових пристроях, мікропотужних підсилювачах, тощо. Можливість побудови електронних схем на взаємодоповнювальних польових транзисторах з різним типом провідності каналу, що називаються комплементарними (КМОН), відкриє значні перспективи їх широкого використання в дискретній техніці. Логічні елементи, елементи збереження дискретної інформації (елементи пам'яті) на КМОН-структурах практично не споживають потужності в статичному режимі. Це дає змогу значно зменшити енергетичні затрати на збереження і перетворення інформації і підвищити ступінь інтеграції компонентів напівпровідникових мікросхем.

Пропонований посібник містить 19, лекцій у яких викладено матеріал з питань структури енергетичних зон, статистики носіїв заряду, явищ переносу, термоелектричних явищ, гальваномагнітних ефектів, процесів дифузії, процесів поглинання світла, явищ фотопровідності, фізичних процесів в поверхневих шарах, контактних явищ у напівпровідниках, діодів і транзисторів, елементів інтегральних схем, магніторезисторів, магнітодіодів, плівкових напівпровідникових сенсорів вологості, напівпровідникових тензосенсорів, активних напівпровідникових фотоелектричних елементів, напівпровідникових наноматеріалів, нанотехнологій та наноелектроніки, сонячних елементів на основі сенсibilізованих барвниками напівпровідників.

Навчальний посібник підготовлено відповідно до програми курсів "Фізика напівпровідникових приладів", "Твердотільна електроніка" і призначений для студентів, які навчаються за напрямком "Фізика", "Прикладна фізика", а також може бути корисним для студентів інших спеціальностей в галузі напівпровідникової електроніки, мікро- та наноелектроніки.

**Д.М. Фреїк, В.М. Чобанюк, З.Ю. Готра,
Б.С. Дзундза, Г.Д. Матеїк, А.І. Ткачук**

Лекція 1

СТРУКТУРА ЕНЕРГЕТИЧНИХ ЗОН

План

- 1.1. Утворення енергетичних зон
- 1.2. Енергетичні зони в твердих тілах
- 1.3. Домішкові енергетичні рівні
- 1.4. Електрони і дірки
- 1.5. Генерація і рекомбінація носіїв заряду

Література

- 1.1. Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.
- 1.2. Мікроелектронні сенсори фізичних величин: Науково-навчальне видання. В 3 томах. Том 1. / Вуйцік В., Готра З.Ю., Григор'єв В.В., Каліта В., Мельник О.М., Потенцік Є.; за ред. З.Ю. Готри. – Львів: Ліга-Прес, 2002. – 475 с.
- 1.3. Мікроелектроніка. В 9 кн. / Под ред. Л. А. Коледова. Кн. 1. Физические основы функционирования изделий микроэлектроники / О.В. Митрофанов, Б.М. Симонов, Л.А. Коледов.: Высшая школа, 1987. – 168 с.
- 1.4. Пасынков В. В., Чиркин Л. К. Полупроводниковые приборы. – М.: Высшая школа, 1987. – 479 с.

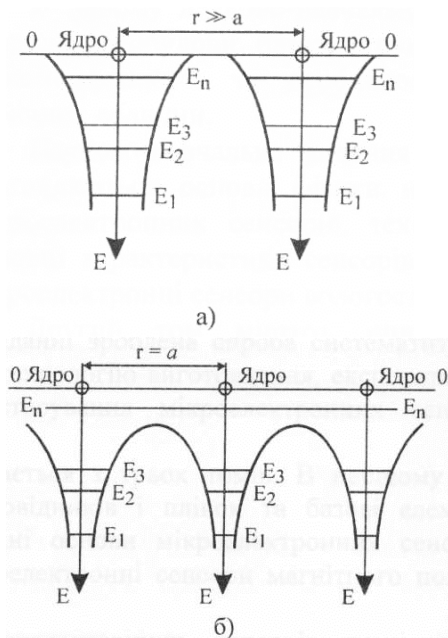


Рис. 1.1. Енергетичні схеми атомів: а – віддалених один відносно одного на відстань $r \gg a$; б – наближених на відстань $r = a$.

1.1. Утворення енергетичних зон

За умови, що атоми у матеріалі розміщені на великих відстанях один від одного взаємодією між ними можна знехтувати (рис. 1.1, а), Потенційна енергія взаємодії електрона з ядром $E(r)$ має вигляд енергетичної ями. Електрон у такій ямі має від'ємну енергію і може знаходитись на одному з рівнів E_1, \dots, E_n . Рівні, які розміщені вище ніж E_n , – вільні (рис. 1.1, а). При цьому потенціальні бар'єри перешкоджають вільному переходу електронів від одного атома до другого.

При наближенні атомів потенціальні криві, які відділяють сусідні атоми, частково перекриваються і в результаті утворюють потенціальну криву, яка проходить нижче від нульового рівня 00. Це означає, що зближення атомів приводить до зменшення не тільки ширини але й висоти бар'єра. Оскільки ця висота є нижче початкового положення енергетичного рівня валентних електронів E_n , то останні отримують можливість практично без перешкод переходити від одного атома до другого.

Квантово-механічний рух електрона в такому кристалі описується рівнянням плоскої хвилі:

$$\Psi(\mathbf{r}) = c \cdot \exp(i \cdot \mathbf{k} \cdot \mathbf{r}) = c \cdot \exp[i(k_x x + k_y y + k_z z)], \quad (1.1)$$

де Ψ - хвильова функція; $c = \text{const}$; \mathbf{k} - хвильовий вектор; \mathbf{r} – радіус-вектор, а $k_x = n_1 \cdot 2\pi/L$; $k_y = n_2 \cdot 2\pi/L$; $k_z = n_3 \cdot 2\pi/L$ – його проєкції на осі x, y, z відповідно; L –

лінійний розмір кристала.

Електрони, внаслідок своєї корпускулярно-хвильової природи не мають одночасно визначених координат (x, y, z) і складових імпульсу (p_x, p_y, p_z) в напрямку цих координат. Відповідно до фундаментального твердження квантової теорії - принципу невизначеності Гейзенберга, їх стан визначається співвідношенням невизначеності

$$\Delta x \cdot \Delta p_x \geq h; \Delta y \cdot \Delta p_y \geq h; \Delta z \cdot \Delta p_z \geq h; \quad (1.2)$$

Тут $\Delta x, \Delta y, \Delta z$ – невизначеності в координатах, а $\Delta p_x, \Delta p_y, \Delta p_z$ – невизначеності в складових імпульсу частинки: $h = 6,6 \cdot 10^{-34}$ Дж·с – стала Планка.

1.2. Енергетичні зони в твердих тілах

При наближенні атомів один до одного, як схематично показано на рис. 1.2, дискретні енергетичні рівні, властиві окремому атому, розщеплюються на окремі підрівні в результаті чого тверде тіло характеризується певною зонною діаграмою, в якій дозволені енергетичні зони чергуються з забороненими. Кожна зона складається з множини дискретних енергетичних підрівнів, відстань між якими настільки мала, що зони можна вважати практично неперервними.

Енергетичні зонні діаграми металу, напівпровідника та діелектрика зображено на рис. 1.3, а-в. Верхня дозволена зона називається вільною або зоною провідності, а розміщена безпосередньо під нею дозволена зона – валентною зоною. При температурі $T = 0$ К валентна зона завжди повністю заповнена електронами, а зона провідності може бути заповнена тільки в нижній частині в металах або бути пустою - в напівпровідниках і діелектриках, різниця між якими полягає в значно більшій ширині забороненої зони діелектриків.

Для лужних металів, характерне часткове заповнення валентної зони (рис. 1.4, а).

Лужноземельні елементи мають повністю заповнену двома електронами зовнішню s-підоболонку і, здавалося б, ці кристали повинні мати повністю заповнену електронами енергетичну зону. Насправді це не так. Над валентним s-рівнем цих елементів розташовується вільний p-рівень, якому відповідають 6 дозволених квантових станів електронів. У разі об'єднання атомів у кристал цей рівень розмивається у вільну дозволена р-зону, в якій можливі $6N$ (квантових) станів електронів. При цьому відбувається перекриття енергетичних s- і р-зон, як схематично показано на прикладі кристала берилію (рис. 1.4, б). За температури абсолютного нуля електрони верхніх рівнів s-зони, прагнучи до стану з мінімальною енергією, займають нижні енергетичні рівні р-зони. Отже, в об'єднаній (гібридній) s-р-зоні зайняті $2N$ нижні квантові стани, а $6N$ квантових станів вільні.

Аналогічно формується частково заповнена енергетична зона в елементів III групи періодичної

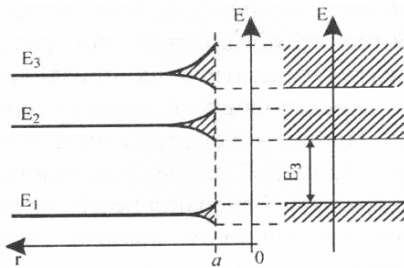


Рис. 1.2. Утворення енергетичних зон з дискретних рівнів при наближенні атомів (а параметр ґратки; r - відстань між атомами).



Рис. 1.3. Енергетичні зонні діаграми при температурі $T = 0$ К: а - металу; б – напівпровідника; в – діелектрика.

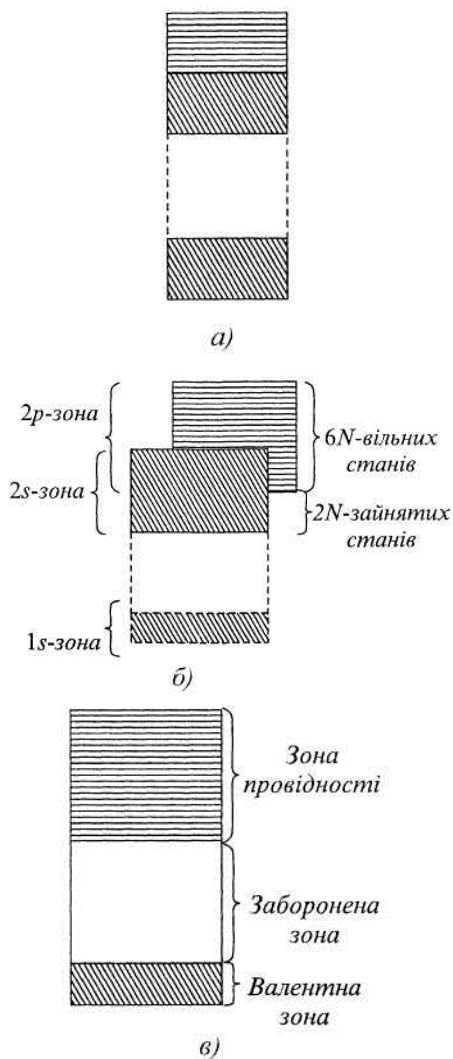


Рис. 1.4. Формування енергетичних діаграм: лужних металів (на прикладі кристала натрію) (а); лужноземельних металів (на прикладі кристала берилію) (б); у власних напівпровідниках і діелектриках (в).

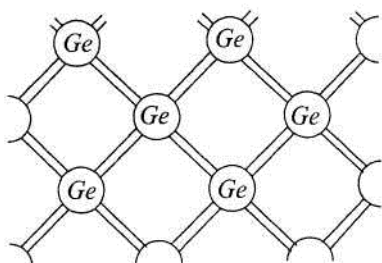


Рис. 1.5. Модель валентних зв'язків германію.

системи елементів.

До типових напівпровідників належать елементи IV групи періодичної системи елементів – кремній і германій, що мають кристалічну ґратку типу алмаза. Як показано на рис. 1.5, чотири валентні електрони цих елементів утворюють ковалентні зв'язки, що мають квантово-механічну природу, з чотирма найближчими атомами, розташованими у вершинах тетраедра. Цією особливістю кристалічної структури визначається сильний зв'язок валентних електронів кремнію і германію з атомами кристалічної ґратки.

Зовнішня електронна оболонка кремнію ($Z = 14$) може містити 8 електронів згідно з принципом Паулі: два електрони в 3s-станах і шість електронів в 3p-станах. У кремнію є тільки чотири електрони: два – в 3s-станах і два – в 3p-станах. У разі об'єднання атомів кремнію в кристал чотири енергетичні рівні розмиваються в чотири дозволені зони. При зменшенні відстані між атомами вони спочатку перекриваються, утворюючи єдину змішану зону, а потім остання розпадається на дві зони, розділені забороненою зоною завширшки ~ 1 еВ. У цих зонах виявляється по $2N$ енергетичні рівні, кожний з яких може бути зайнятий двома електронами. За температури абсолютного нуля електрони, прагнучи зайняти стани, які відповідають мінімуму енергії системи, повністю заповнюють нижню з цих зон – валентну зону, і повністю звільняють верхню зону – зону провідності.

Аналогічно формуються енергетичні діаграми германію і вуглецю в модифікації алмазу. Тільки в алмазі ширина забороненої зони, що відділяє валентну зону від зони провідності, дуже велика, внаслідок чого алмаз є досконалим діелектриком.

Отже, речовини з неповністю заповненими валентними зонами є провідниками, а речовини з повністю заповненими – напівпровідниками або діелектриками.

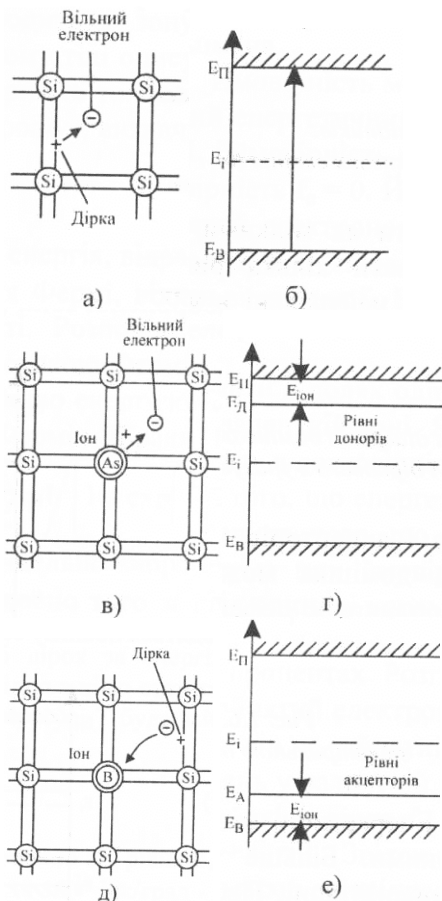


Рис. 1.6. Утворення вільних носіїв заряду в кремнії: власному (а, б), електронному (в, г) і діркового (д, е)

1.3. Домішкові енергетичні рівні

Домішкові атоми у напівпровідниках, можуть суттєво впливати на енергетичну структуру. Процеси утворення вільних носіїв заряду, які здатні приймати участь в явищах переносу, у власному і домішковому кремнії та їх енергетичні діаграми представлені на рис. 1.6. Для кремнію характерні домішки заміщення, тобто атоми домішки заміщають атоми кремнію у вузлах кристалічної ґратки. У власному кремнії одночасно утворюються вільний електрон і дірка (рис. 1.6, а), що відповідає переходу електрону з валентної зони в вільну (рис. 1.6, б).

При введенні в кремній атома елемента V групи Періодичної таблиці (наприклад, миш'яку As) чотири з п'яти його валентних електронів вступають в зв'язок з чотирма валентними електронами сусідніх атомів кремнію і утворюють стійку оболонку з восьми електронів. Дев'ятий електрон стає слабо зв'язаним з ядром п'ятивалентного елемента, він легко відривається і перетворюється на вільний електрон (рис. 1.6, в), дірка при цьому не утворюється. На енергетичній діаграмі цей процес відповідає переходу електрона з рівня донорів E_D у вільну зону (рис. 1.6, г). Домішковий атом перетворюється на нерухомий іон з одиничним позитивним зарядом. Домішка такого типу називається донорною, а напівпровідники, в які введено атоми донорів, електронними або n-типу електропровідності. В таких напівпровідниках вільних електронів більше, ніж дірок, і вони характеризуються переважно електронною електропровідністю.

Якщо в кремній ввести атом тривалентного елемента III групи (наприклад, бору B), то всі три його валентних електрони вступають у зв'язок з чотирма електронами сусідніх атомів кремнію. Для утворення стійкої оболонки з восьми електронів не вистачає одного. Ним стає один з валентних електронів, який буде забрано від найближчого сусіднього атома, у якого в результаті утворюється незаповнений зв'язок – дірка (рис. 1.6, д). На енергетичній діаграмі цей процес відповідає переходу електрона з валентної зони на рівень акцепторів E_A і утворенню в валентній зоні дірки (рис. 1.6, е). Домішковий атом перетворюється на нерухомий іон з одиничним негативним зарядом, вільний електрон при цьому не утворюється. Домішка такого типу називається акцепторною, а напівпровідники, в які введено атоми акцепторів, – дірковими або p-типу провідності. Дірок в них більше ніж вільних електронів, тому ці напівпровідники мають переважно діркову електропровідність.

З енергетичних діаграм електронних і діркових напівпровідників (рис. 1.6, г, е) видно, що рівні донорів E_D і акцепторів E_A розміщені в забороненій зоні: рівні E_D – поблизу дна зони провідності, а рівні E_A – поблизу

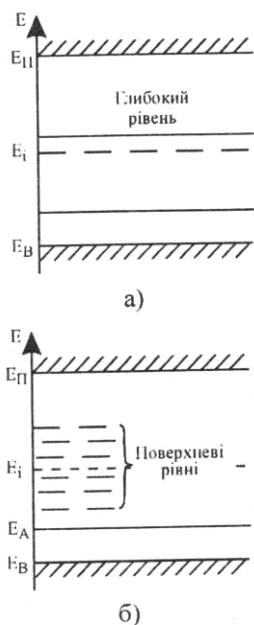


Рис. 1.7. Енергетичні діаграми домішкових напівпровідників: з нейтральною домішкою (а) і діркового з урахуванням поверхневих станів (б).

верхньої границі валентної зони. Відривання лишнього електрона від донора чи додавання додаткового електрона до акцептора вимагає затрати енергії $E_{\text{іон}}$, яка показана на діаграмах. Енергетичні рівні домішкових атомів, розміщені поблизу дозволених зон, називаються мілкими. Деякі домішки утворюють глибокі рівні, які знаходяться поблизу середини забороненої зони (рис. 1.7, а).

Особливу роль в будь-якому реальному напівпровіднику відіграє його поверхня. Структурні порушення кристалічної ґратки і наявність адсорбованих атомів створюють поблизу поверхні додаткові енергетичні рівні, які називаються поверхневими. Ці рівні можуть займати різне положення на енергетичній діаграмі, однак частіше – в межах забороненої зони (рис. 1.7, б).

1.4. Електрони і дірки

З підвищенням температури у напівпровідникових кристалах з'являється ймовірність того, що внаслідок взаємодії з тепловими коливаннями ґратки електрони валентної зони отримають порції енергії, достатні для переходу на рівні вищої вільної зони – зони провідності. Кількість переходів з валентної зони в зону провідності протягом одиниці часу в одиниці об'єму кристала тим більша, чим більша концентрація валентних електронів, чим вища температура і чим менша ширина забороненої зони. Найбільш імовірними будуть переходи в зону провідності електронів, розташованих на верхніх рівнях валентної зони, оскільки для переходу їм потрібна найменша додаткова енергія. З цієї ж причини в зоні провідності заселяються передусім нижні енергетичні рівні, на яких електрони мають додатну, приблизно постійну ефективну масу і поведуться в першому наближенні як вільні. Збуджені електрони "живуть" у зоні провідності дуже нетривалий (~ 1 мкс) час, після чого повертаються на вільні рівні валентної зони. Кількість таких зворотних переходів збільшується із зростанням концентрації електронів зони провідності. Це означає, що за певної температури встановлюється *рівноважна концентрація* електронів зони провідності n . У речовин, ширина забороненої зони яких не перевищує 1 еВ, вже при кімнатних температурах концентрація електронів зони провідності досить велика, що й визначає їх напівпровідникові властивості.

Важливим є факт, що електропровідність власних напівпровідників визначається не тільки появою

електронів зони провідності, але і звільненням при цьому такої ж кількості енергетичних рівнів у валентній зоні.

Ці "частинки" називають *дірками*. Тут йдеться не про реальні частинки, а про деяку умовну модель, яка істотно полегшує аналіз фізичних процесів у напівпровідниках, оскільки задача про рух величезної кількості електронів у майже заповненій валентній зоні замінюється задачею про рух невеликої кількості дірок у майже вільній енергетичній зоні.

За своїми властивостями дірка є неначе дзеркальним відображенням електрона, що знаходився раніше на цьому енергетичному рівні. Через те, що її заряд протилежний за знаком до заряду електрона e , то напрям руху дірки в електричному полі протилежний до напрямку руху електрона. Для того, щоб струм дірок був еквівалентний до струму, створеному насправді рухом електронів валентної зони, кожній дірці треба приписати ефективну масу m^* , яка дорівнює за величиною і протилежна за знаком ефективній масі електрона на цьому енергетичному рівні. Оскільки перехід електронів у зону провідності відбувається, як правило, з верхніх рівнів валентної зони, для яких характерна від'ємна ефективна маса електронів, ефективна маса дірок виявляється додатною. Звідси впливає можливість розгляду руху дірок в електричному полі як вільних частинок з приблизно постійною масою. Нарешті, відліку енергії електронів зони провідності вгору від дна цієї зони відповідає відлік енергії дірок вниз від стелі валентної зони. Внаслідок рівності у власних напівпровідників кількості дірок у валентній зоні кількості електронів у зоні провідності обидва механізми електропровідності – електронний і дірковий – у них рівноправні та однаково впливають на створення електричного струму.

1.5. Генерація і рекомбінація носіїв заряду

Хімічно чистий напівпровідник без будь-яких домішок називають власним. При температурі $T = 0 \text{ K}$ у власному напівпровіднику немає вільних носіїв заряду і його електрична провідність $\sigma = 0$.

Атоми в кристалічній ґратці кремнію та інших напівпровідників зв'язані між собою за рахунок сил обміну, які виникають в результаті попарного з'єднання валентних електронів сусідніх атомів, при цьому кожен з атомів залишається електрично нейтральним. Такий зв'язок називається ковалентним. Підвищення температури викликає коливальний рух атомів кристалічної ґратки. В результаті ковалентні зв'язки між атомами можуть розриватися, що призводить утворення пари носіїв заряду: вільного електрона та незаповненого зв'язку – дірки поблизу того атома, від якого відірвався електрон. Процес утворення електронно-діркових пар називається генерацією носіїв заряду. Якщо цей процес відбувається під впливом теплоти, то його називають

термогенерацією. Генерація носіїв заряду може відбуватися під впливом поглинутих напівпровідником квантів світла (світлова генерація) та інших енергетичних факторів.

Незаповнений електроном зв'язок швидко заповнюється одним з валентних електронів сусіднього атома, на місці якого утворюється дірка, і цей процес повторюється. Наявність дірки в напівпровіднику еквівалентна наявності в ньому частинки з позитивним зарядом, який дорівнює за абсолютним значенням заряду електрона.

Одночасно з генерацією в напівпровіднику неперервно відбувається і зворотний процес – рекомбінація носіїв заряду, тобто повернення електронів з зони провідності до валентної зони, в результаті чого зникає пара носіїв заряду. В стані термодинамічної рівноваги процеси генерації та рекомбінації носіїв заряду взаємно врівноважені. При цьому в напівпровіднику існують рівноважні концентрації електронів n_0 і дірок p_0 .

При дії на напівпровідник не теплового зовнішнього енергетичного фактора (світла, сильного електричного поля та ін.) внаслідок генерації нових носіїв заряду їх концентрація n і p (нерівноважна концентрація) буде перевищувати рівноважну концентрацію на величину Δn (або Δp), яку називають надлишковою концентрацією. Таким чином.

$$\Delta n = n - n_0; \Delta p = p - p_0 \quad (1.3)$$

Надлишкова концентрація носіїв заряду може виникати в окремих областях напівпровідникової структури приладу не тільки в результаті зовнішнього енергетичного впливу, але й за рахунок різних процесів (інжекції, екстракції, акумуляції і т.д.).

Механізми рекомбінації можуть бути різними (рис. 1.8). Міжзонна або безпосередня рекомбінація відбувається при переході вільного електрона з зони провідності у валентну зону на один з вільних енергетичних рівнів, що рівноцінно зникненню пари носіїв заряду – вільного електрона і дірки. Однак такий процес міжзонної рекомбінації мало ймовірний, оскільки вільний електрон і дірка повинні знаходитися одночасно в одному і тому ж місці кристала. Крім того, повинен виконуватися закон збереження імпульсу, тобто рекомбінація електрона і дірки можлива тільки при однакових, але протилежно направлених імпульсах електрона і дірки. Тому, наприклад, в германії на 10 тис. рекомбінацій лише одна відбувається в результаті міжзонної рекомбінації (рис. 1.8, а).

Рекомбінація з участю рекомбінаційних пасток відбувається в два етапи (рис. 1.8, а). На першому етапі рекомбінаційна пастка (або енергетичний рівень рекомбінаційної пастки) захоплює, наприклад, електрон з зони провідності який вибуває з процесу

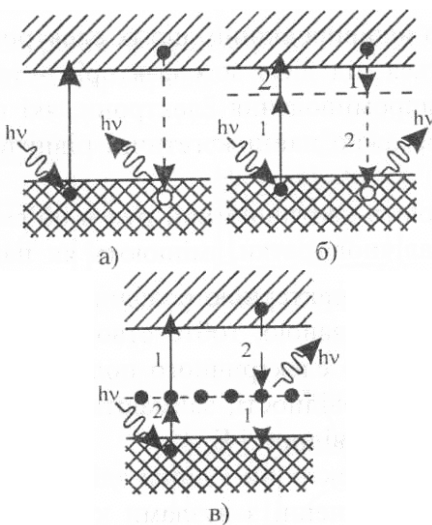


Рис. 1.8. Різні механізми генерації та рекомбінації носіїв заряду: а – міжзонна генерація (—) і рекомбінація (----); б – генерація і рекомбінація з участю пустих рекомбінаційних пасток; в – генерація і рекомбінація з участю заповнених електронами рекомбінаційних пасток.

електропровідності. В цьому стані пастка перебуватиме до того часу, поки до неї не підійде дірка, тобто, поки в даному місці кристала не виникне вільний енергетичний рівень валентної зони. При виконанні цих умов здійснюється другий етап рекомбінації – електрон переходить на вільний рівень валентної зони (що еквівалентно захопленню дірки з валентної зони від'ємно зарядженою пасткою). Двоетапний процес рекомбінації ймовірніший, оскільки він не вимагає одночасної присутності в даному місці кристала вільного електрона і дірки. Рекомбінаційна пастка сприймає кількість руху, необхідну для збереження закону збереження імпульсу, і може забрати частину енергії, яка звільниться в процесі рекомбінації.

В залежності від того, як витрачається енергія, звільнена при рекомбінації електрона і дірки, рекомбінацію можна розділити на два види:

– випромінювальна, при якій енергія, яка звільняється при переході електрона на нижчий енергетичний рівень, випромінюється в вигляді кванта світла (фотона).

– безвипромінювальна (фононна), при якій надлишкова енергія електрона передається кристалічній ґратці напівпровідника та витрачається на утворення фононів – квантів теплової енергії.

В якості рекомбінаційних пасток можуть виступати домішкові атоми або іони, різні включення в кристалі, незаповнені вузли кристалічної ґратки, тріщини та інші недосконалості об'єму або поверхні. На поверхні кристала перерахованих дефектів значно більше, ніж в об'ємі, тому процес рекомбінації на поверхні відбувається значно інтенсивніше. Його розглядають і оцінюють окремо, вважаючи поверхневу рекомбінацію різновидом рекомбінації з участю рекомбінаційних пасток.

Лекція 2

СТАТИСТИКА НОСІЇВ ЗАРЯДУ В НАПІВПРОВІДНИКАХ

План

- 2.1. Густина енергетичних рівнів. Функція розподілу Фермі-Дірака
- 2.2. Час життя нерівноважних носіїв заряду
- 2.3. Температурні залежності концентрації носіїв заряду
- 2.4. Температурні залежності рухливості носіїв заряду і питомої провідності

Література

- 2.1. Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.
- 2.2. Мікроелектронні сенсори фізичних величин: Науково-навчальне видання. В 3 томах. Том 1. / Вуйцік В., Готра З.Ю., Григор'єв В.В., Каліта В., Мельник О.М., Потенцкі Є.; за ред. З.Ю. Готри. – Львів: Ліга-Прес, 2002. – 475 с.
- 2.3. Мікроелектроніка. В 9 кн. / Под ред. Л. А. Коледова. Кн. 1. Физические основы функционирования изделий микроэлектроники / О.В. Митрофанов, Б.М. Симонов, Л.А. Коледов.: Высшая школа, 1987. – 168 с.
- 2.4. Степаненко И. П. Основы микроэлектроники. – М.: Сов. радио, 1980. – 286 с.

2.1. Густина енергетичних рівнів. Функція розподілу Фермі-Дірака

Напівпровідник, при деякій температурі $T = const$, перебуває в стані термодинамічної рівноваги. Цей стан характерний рівністю швидкостей генерації $\nu_{ген}$ і рекомбінації $\nu_{рек}$.

В одиничному об'ємі напівпровідника є певна для даного напівпровідника і температури кількість вільних носіїв заряду, яка називається концентрацією. Кількість електронів у зоні провідності та дірок у валентній зоні звичайно значно менша від кількості енергетичних станів, які є в цих зонах. Тому середня густина заповнення енергетичних станів електронами і дірками f відповідає нерівності

$$f \ll 1. \quad (2.1)$$

Якщо ця умова виконується, напівпровідник називається невідродженим, якщо, навпаки, не виконується – виродженим.

У власному напівпровіднику, при температурі $T = 0$ К, всі енергетичні рівні валентної зони заповнені електронами, а рівні зони провідності – вільні. З підвищенням температури деяка кількість електронів залишає валентну зону і переходить у зону провідності. Розподіл електронів і дірок за енергіями в твердому тілі описується статистикою Фермі-Дірака. Згідно з цією статистикою ймовірність того, що стан з деякою енергією E при температурі T буде зайнятий електроном, визначає функція Фермі-Дірака:

$$f_n(E, T) = 1 / \left[1 + \exp\left(\frac{E - E_F}{kT}\right) \right], \quad (2.2)$$

де E – енергія, відрахована від довільно вибраного рівня (звично від рівня E_B); E_F – енергія Фермі, відрахована відносно того ж рівня; $k = 1,38 \cdot 10^{-23}$ Дж/град – стала Больцмана.

Ймовірність можна виразити в долях одиниці або в процентах. Розглянемо деякий енергетичний рівень у валентній зоні. Якщо він зайнятий електроном, то $f_n = 1$, і ймовірність перетворилась в достовірність. Якщо ж він не зайнятий, то ця ймовірність $f_n = 0$. Ймовірність того, що енергетичний рівень у валентній зоні не зайнятий електроном, є ймовірність того, що на цьому рівні є дірка f_p . Таким чином, $f_p = 1$, якщо $f_n = 0$, а сума цих ймовірностей $f_p + f_n = 1$. Звідси можна знайти вираз для f_p , віднявши від одиниці вираз (2.2), отримаємо співвідношення

$$f_p(E, T) = 1 / \left[1 + \exp\left(\frac{E_F - E}{kT}\right) \right]. \quad (2.3)$$

Воно визначає ймовірність того, що даний енергетичний рівень E при температурі T зайнятий діркою. Енергетичний рівень Фермі E_F – це рівень енергії, ймовірність заповнення якого електроном при будь-яких значеннях температури дорівнює 0,5%.

Розглянемо деякий інтервал значень енергії dE , який лежить у зоні провідності. Якщо кількість енергетичних станів, які відповідають енергії E , які можуть бути зайняті електронами, через $N(E)$, то в інтервалі dE буде $N(E) dE$ таких станів. Виразивши концентрацію електронів dn , які мають енергії в інтервалі dE , як добуток $N(E) dE$ на $f_n(E, T)$ і інтегруючи за всіма значеннями енергії в зоні провідності, знайдемо кількість електронів в зоні провідності в одиниці об'єму

$$n = \int_{E_c}^{\infty} N(E) \cdot f_n(E, T) dE. \quad (2.4)$$

Аналогічно для концентрації дірок у валентній зоні

$$p = \int_0^{E_B} N(E) \cdot f_p(E, T) dE. \quad (2.5)$$

Для обчислення інтегралів в (2.4) і (2.5) використаємо закон розподілу густини енергетичних станів, якій візьмемо з квантової механіки для області енергій $E > E_c$:

$$N(E) = \frac{4p(2m_n)^{3/2}}{h^3} \cdot (E - E_c)^{1/2}, \quad (2.6)$$

де m_n – ефективна маса електрона в зоні провідності. Даний вираз (2.6) справджується для вузької смуги енергій поблизу зони провідності. Після інтегрування (2.4) отримуємо вираз для концентрації вільних електронів у напівповіднику:

$$n = N_c \cdot \exp\left[-(E_c - E_F) / kT\right]. \quad (2.7)$$

Аналогічні розрахунки для дірок дають вираз

$$p = N_v \cdot \exp\left[-(E_F - E_B) / kT\right]. \quad (2.8)$$

Коефіцієнти N_c і N_v називаються еквівалентними густинами станів відповідно електронів і дірок. Припустивши, що всі вільні електрони сконцентровані поблизу дна зони провідності, а дірки – поблизу стелі валентної зони, їх можна розрахувати за формулами:

$$N_c = 2 \left(\frac{2pm_n kT}{h^2} \right)^{3/2}; \quad (2.9)$$

$$N_v = 2 \left(\frac{2pm_p kT}{h^2} \right)^{3/2}, \quad (2.10)$$

де m_n , m_p – ефективні маси електрона і дірки в напівпровіднику; m – маса вільного електрона; k – стала Больцмана; h – стала Планка.

У власному напівпровіднику енергетичний рівень Фермі E_F лежить приблизно посередині забороненої зони:

$$E_F = \frac{E_g}{2} + \frac{kT}{2} \cdot \ln \frac{N_v}{N_c} = \frac{E_v}{2} + \frac{3}{4} \cdot kT \cdot \ln \frac{m_p}{m_n}. \quad (2.11)$$

З ростом температури рівень E_F звичайно піднімається (як правило, $m_p > m_n$). Для більшості власних напівпровідників (кремній, германій та ін.) в робочих інтервалах температур цим зміщенням можна знехтувати. При розрахунках часто припускають, що $N_c = N_v$. Тоді можна записати

$$E_F = E_g / 2 = E_i,$$

де через E_i позначена середина забороненої зони.

Знайдемо добуток концентрацій носіїв заряду в напівпровіднику:

$$n \cdot p = N_c N_v \exp\left(\frac{-E_g}{kT}\right), \quad (2.12)$$

з якого випливає, що цей добуток постійний для даної температури і не залежить від характеру та кількості домішок, які містить напівпровідник. Концентрація носіїв заряду в власному напівпровіднику

$$n_i = n = p = \sqrt{N_c N_v} \exp\left(\frac{-E_g}{kT}\right). \quad (2.13)$$

З (2.13) видно, що власна концентрація залежить від ширини забороненої зони E_g напівпровідника і температури. Вплив температури на величину n_i тим сильніший, чим вище E_g . Рівень Фермі E_F однаковий в усіх частинах рівноважної системи, якою б різнорідною вона не була, тобто $E_F = const$.

Функція Фермі-Дірака (2.2), (2.3) справедлива не тільки для власних, але й для домішкових напівпровідників. У напівпровідниках n-типу велика кількість електронів переходить у зону провідності з

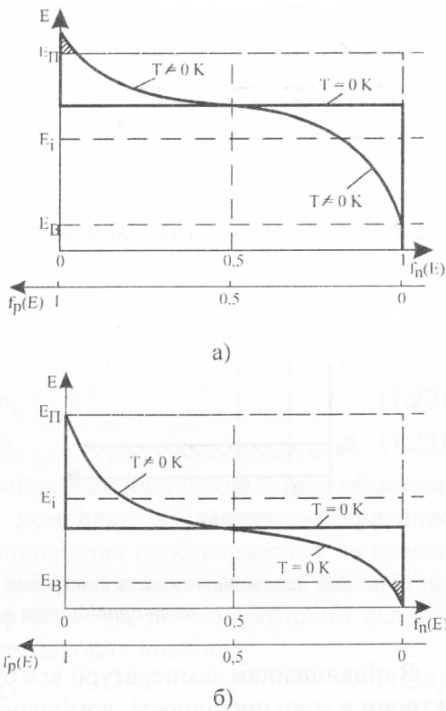


Рис. 2.1. Функції Фермі-Дірака $f_n(E)$ і $f_p(E)$ домішкових напівпровідників *n*- (а) і *p*-типу (б).

рівнів донорів, при цьому дірки у валентній зоні не з'являються. Тому ймовірність появи електрона в зоні провідності вища від ймовірності появи дірки у валентній зоні. Це можливо у випадку зміщення рівня Фермі E_F в бік дна зони провідності. Чим вища концентрація атомів донора в напівпровіднику N_D , тим більше це зміщення. В напівпровіднику *p*-типу, навпаки, рівень E_F розміщується нижче рівня E_V , і з підвищенням концентрації атомів акцепторів N_A в ньому рівень Фермі буде лежати ближче до стелі валентної зони. На рис. 2.1 представлені функції Фермі-Дірака для домішкових напівпровідників *n*-типу (а) і *p*-типу (б). Заштриховані ділянки пропорційні концентраціям носіїв заряду в зонах.

Залежності положення рівня Фермі від температури в домішкових напівпровідниках *n*- і *p*-типів показані на рис. 2.2, а, б. Розглядаючи характер цих залежностей на прикладі напівпровідника *n*-типу (рис. 2.2, а) в області низьких температур можна знехтувати переходами електронів з валентної зони в зону провідності і вважати, що всі електрони в зоні провідності з'являються в результаті іонізації донорів. Тоді вираз для рівня Фермі E_F при низьких температурах можна записати у вигляді:

$$E_F = \frac{E_c - E_D}{2} + \frac{kT}{2} \cdot \ln\left(\frac{N_D}{N_c}\right), \quad (2.14)$$

де E_D – енергетичний рівень донорів; N_D – їх концентрація. З цього виразу при $T = 0$ К отримуємо $E_F = (E_c - E_D)/2$.

З підвищенням температури все більша кількість атомів донорів віддає свої електрони в зону провідності, домішкові рівні поступово виснажуються. Повне їх виснаження настає при $n = N_D$. Положення рівня Фермі E_{FS} , яке відповідає цій умові, можна визначити з рівності $E_{FS} = E_D$. Виходячи з визначення рівня Фермі, для концентрації електронів у зоні провідності можна записати $n_n = N_D/2$ і обчислити за допомогою (2.14) температуру виснаження домішки:

$$T_S = \frac{E_c - E_D}{k \cdot \ln(2N_c / N_D)}. \quad (2.15)$$

При всіх температурах, відмінних від нуля, електрони в зоні провідності з'являються не тільки за рахунок донорних атомів, але й за рахунок переходів у зону провідності з валентної зони, тому можна записати

$$n_n = n_{np} + n_i, \quad (2.16)$$

де n_{np} – концентрація електронів, зумовлена іонізацією домішки, n_i – концентрація електронів, зумовлена генерацією власних носіїв заряду.

Рис. 2.2. Залежність положення рівня Фермі E_F від температури в домішкових напівпровідниках *n*-типу (а) і *p*-типу (б).

З підвищенням температури власна концентрація n_i зростає, досягає домішкової концентрації $n_{np} = N_D$ і перевищує її, що відповідає переходу до власної електропровідності, який відбувається при деякій температурі T_i , названій температурою іонізації. При

температурі $T > T_i$ першим членом у виразі (2.17) можна знехтувати, тоді температуру T_i можна виразити як:

$$T_i = \frac{E_g}{k \cdot \ln(N_c N_v / N_D^2)}. \quad (2.17)$$

Положення рівня Фермі та концентрації носіїв заряду в напівпровідниках р-типу визначають таким же чином, як і для напівпровідників n-типу.

Якщо відомі концентрації основних носіїв заряду, то можна обчислити концентрації неосновних носіїв у домішкових напівпровідниках, використавши співвідношення

$$np = n_n p_n = n_p p_p = n_i^2, \quad (2.18)$$

де n , p , n_n , p_n , n_p , p_p – концентрації електронів і дірок відповідно в напівпровідниках: власному, n-типу і р-типу. Співвідношення (2.18) часто називають законом діючих мас. Звідси можна знайти концентрації неосновних носіїв заряду:

$$p_n = n_i^2 / n_n; \quad (2.19)$$

$$n_p = n_i^2 / p_p. \quad (2.20)$$

Робочий температурний діапазон домішкових напівпровідників обмежений знизу температурою повної іонізації домішок, а зверху – критичною температурою, при якій домішковий напівпровідник перетворюється на власний. У робочому діапазоні можна вважати всі домішкові атоми повністю іонізованими і знехтувати власною концентрацією n_i , приймаючи, що концентрації основних носіїв заряду дорівнюють концентраціям домішкових атомів:

$$n_n = N_D; \quad (2.21)$$

$$p_p = N_A. \quad (2.22)$$

Отже, за температури абсолютного нуля в металі немає електронів, з енергією, вищою, ніж E_F .

У діелектриків і власних напівпровідників рівень Фермі розташовується приблизно посередині забороненої зони і практично не залежить від температури.

В електронного напівпровідника рівень Фермі при $T = 0 \text{ K}$ розташовується посередині між: донорним рівнем і дном зони провідності, а з підвищенням температури він зміщується вниз, причому тим сильніше, чим менша концентрація донорної домішки n_d .

У діркового напівпровідника рівень Фермі при $T = 0 \text{ K}$ розташовується посередині між акцепторним рівнем і стелею валентної зони, а за підвищення температури він зсувається догори, причому тим сильніше, чим менша концентрація акцепторів n_a .

Рівноважна концентрація основних і неосновних носіїв заряду в напівпровіднику дорівнює квадрату концентрації власних носіїв заряду. Цей висновок зумовлений термодинамічною рівновагою в

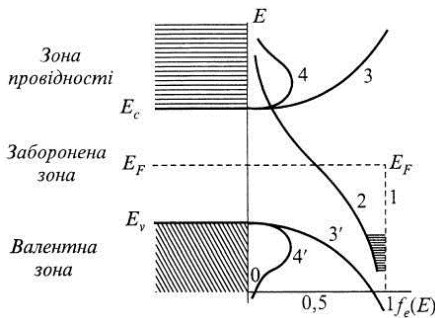


Рис. 2.3. Положення рівня Фермі і заповнення енергетичних рівнів електронами та дірками у власних напівпровідників: 1-2 – ймовірність заповнення електронами рівнів, відповідно, при $T = 0 \text{ K}$ і $T > 0 \text{ K}$; 3-3' – густина рівнів у зоні провідності та валентній зоні; 4-4' – розподіл за енергіями електронів у зоні провідності та дірок у валентній зоні.

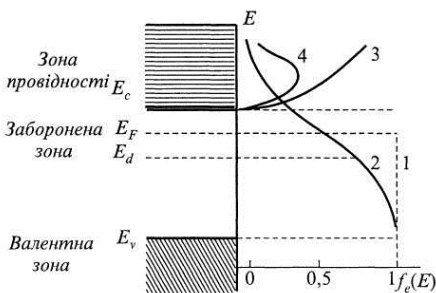


Рис. 2.4. Положення рівня Фермі і заповнення енергетичних рівнів електронами в донорного напівпровідника: 1,2 – ймовірність заповнення електронами рівнів відповідно при $T=0 \text{ K}$ і $T>0 \text{ K}$; 3 – густина рівнів у зоні провідності; 4 – розподіл за енергіями електронів у зоні провідності.

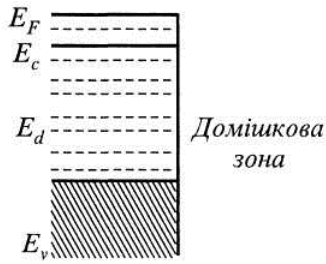


Рис. 2.5. Енергетична діаграма виродженого напівпровідника.

напівпровіднику і є окремим випадком відомого з хімії закону діючої маси.

Внаслідок високої концентрації домішкових атомів, а отже, і електронів у зоні провідності сильно легованого напівпровідника його електронний газ виявляється виродженим. Тому самі напівпровідники з подібною енергетичною діаграмою називають виродженими.

2.2. Час життя нерівноважних носіїв заряду

Після припинення енергетичного впливу на напівпровідник надлишкова концентрація носіїв заряду в ньому внаслідок процесу рекомбінації через деякий час зменшується до нуля. Кількість носіїв, які рекомбінують в одиниці об'єму (швидкість зміни концентрації), пропорційна до надлишкової концентрації і обернено пропорційна до деякого параметра τ , який називають часом життя:

$$\frac{dn}{dt} = -\frac{d(\Delta n)}{dt} = \frac{\Delta n}{t_n}, \quad \frac{dp}{dt} = -\frac{d(\Delta p)}{dt} = \frac{\Delta p}{t_p}. \quad (2.23)$$

Таким чином, часом життя нерівноважних носіїв заряду є співвідношення надлишкової концентрації (Δn або Δp) нерівноважних носіїв заряду до швидкості зміни цієї концентрації внаслідок рекомбінації:

$$t_n = \frac{|\Delta n|}{|d(\Delta n)/dt|} \quad \text{і} \quad t_p = \frac{|\Delta p|}{|d(\Delta p)/dt|}. \quad (2.24)$$

Процес перенесення спостерігають у напівпровідниках при наявності електронів у зоні провідності і при неповному заповненні електронами валентної зони. При виконанні цих умов і за відсутності градієнту температури перенесення носіїв заряду може відбуватися або під дією електричного поля, або під дією градієнта концентрації носіїв заряду.

Направлений рух носіїв заряду під дією електричного поля називають дрейфом. Електрони отримують в електричному полі на довжині вільного пробігу додаткову енергію і при цьому переходять на вищі енергетичні рівні. При наступній взаємодії електрона з атомом кристалічної ґратки електрон віддає кристалічній ґратці нагромаджену протягом вільного пробігу енергію і повертається на один з нижчих енергетичних рівнів у дозволений зоні. Так можна описати процес електропровідності в результаті руху електронів у зоні провідності та валентній зоні. Однак, враховуючи майже повне заповнення електронами енергетичних рівнів валентної зони, зручніше розглядати у валентній зоні рух дірок: дірки, рухаючись у напрямку вектора електричного поля і отримуючи в цьому полі додаткову енергію, переходять за енергетичною діаграмою на нижчі енергетичні рівні.

В результаті дрейфу електронів у напівпровіднику

з'являється електронна складова густини дрейфового струму, яку на основі закону Ома запишемо як

$$J_{n\text{др}} = g_n E = qn\mu_n E, \quad (2.25)$$

де γ_n – питома електропровідність напівпровідника при наявності тільки електронів; μ_n – рухливість електронів, тобто величина, яка чисельно дорівнює середній швидкості їх направленої руху в електричному полі з напруженістю, яка дорівнює одиниці; q – заряд електрона.

Аналогічно, діркова складова густини дрейфового струму

$$J_{p\text{др}} = g_p E = qp\mu_p E, \quad (2.26)$$

Знаки в правих частинах співвідношень (2.25) і (2.26) однакові, оскільки заряди електронів і дірок різні за знаком і рухаються в електричному полі в різні сторони.

Повна густина дрейфового струму при наявності вільних електронів і дірок дорівнює сумі електронної та діркової складових:

$$J = J_n + J_p = qgE = q(n\mu_n + p\mu_p)E, \quad (2.27)$$

де γ – питома провідність напівпровідника з урахуванням вільних електронів і дірок.

В результаті нерівномірного розподілу концентрації носіїв заряду в об'ємі напівпровідника за відсутності градієнта температури відбувається дифузія – рух носіїв заряду внаслідок градієнта концентрації, вирівнювання концентрації носіїв по напівпровіднику.

Відомо, що потік частинок при дифузії (кількість частинок, які перетинають за одиницю часу одиничну площу, перпендикулярну напрямку градієнта концентрації) пропорційний градієнту концентрацій цих частинок:

$$\Phi_m = -D_m \cdot \text{grad } m, \quad (2.28)$$

де D_m – коефіцієнт дифузії, який дорівнює абсолютному значенню відношення густини потоку частинок до градієнта їх концентрації.

Внаслідок того, що вектор градієнта концентрації направлений у бік зростання аргументу, а частинки дифундують туди, де їх менше, тобто проти градієнта концентрації, знаки перед лівою і правою частинами співвідношення (2.28) різні.

Густину електронної складової дифузійного струму можна одержати шляхом множення правої частини (2.28) на елементарний заряд електрона. Електрони дифундують проти вектора градієнта концентрації і мають негативний заряд. Тому напрям вектора густини дифузійного струму електронів повинен збігатися з напрямком вектора градієнта концентрації електронів, тобто

$$J_{n\text{диф}} = qD_n \cdot \text{grad } n. \quad (2.29)$$

Аналогічно для густини діркової складової дифузійного струму маємо вираз

$$J_{p\text{диф}} = -qD_p \cdot \text{grad } p. \quad (2.30)$$

Оскільки заряд дірок позитивний, напрям вектора густини дифузійного струму дірок повинен збігатися з напрямком їх дифузії, внаслідок чого в співвідношенні (2.30) зберігається знак мінус.

Одночасно з процесом дифузії нерівноважних носіїв заряду відбувається процес їх рекомбінації. Тому надлишкова концентрація зменшується в напрямку від джерела надлишкової концентрації. Відстань, на якій при одномірній дифузії в напівпровіднику без електричного поля в ньому концентрація носіїв заряду зменшується внаслідок рекомбінації в $e = 2,718\dots$ рази, називають дифузійною довжиною L . Не можна плутати дифузійну довжину з довжиною вільного пробігу носіїв заряду, яка визначається як середня відстань, яку проходить носій заряду між двома послідовними актами розсіювання. Іншими словами, це відстань, на яку носій заряду дифундує за час життя. Таким чином, дифузійна довжина зв'язана з часом життя носіїв співвідношеннями

$$L_n = \sqrt{D_n t_n}; L_p = \sqrt{D_p t_p}, \quad (2.31)$$

де D_n і D_p – коефіцієнти дифузії електронів і дірок.

Коефіцієнти дифузії зв'язані з рухливістю носіїв заряду співвідношеннями Ейнштейна

$$D_n / m_n = kT / q; D_p / m_p = kT / q. \quad (2.32)$$

Необхідно зауважити, що дифузія носіїв заряду може відбуватися в напівпровіднику, який має рівномірний розподіл концентрації носіїв, тобто рівний нулю градієнт концентрації, але при наявності в напівпровіднику градієнта температури. В цьому випадку виникає дифузія електронів з нагрітих областей у холодні, внаслідок чого виникає градієнт концентрації.

При наявності електричного поля і градієнта концентрації носіїв заряду в напівпровіднику будуть існувати дрейфові та дифузійні струми. Тому, враховуючи (2.25) і (2.29), густина електронного струму

$$J_n = J_{n \text{ др}} + J_{n \text{ диф}} = qn m_n E + qD_n \cdot \text{grad } n. \quad (2.33)$$

Густина діркового струму з урахуванням (2.26) і (2.30)

$$J_p = J_{p \text{ др}} + J_{p \text{ диф}} = qp m_p E + qD_p \cdot \text{grad } p. \quad (2.34)$$

Густина повного струму в напівпровіднику дорівнює сумі струмів (2.33), (2.34) і густини струму зміщення, який дорівнює похідній за часом від вектора електричної індукції:

$$J = J_n + J_p + \frac{\partial}{\partial t}(\epsilon \epsilon_0 E). \quad (2.35)$$

2.3. Температурні залежності концентрації носіїв заряду

Температурні залежності концентрації носіїв заряду і положення рівня Фермі в широкому діапазоні температур і при різних концентраціях домішок представлені на рис. 2.6.

При відносно малій концентрації домішки (донорів)

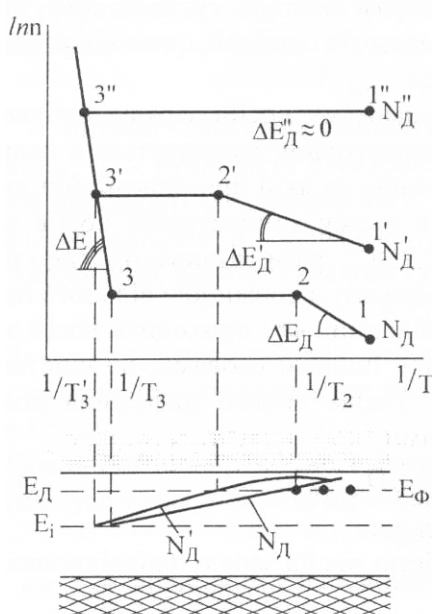


Рис. 2.6. Температурні залежності концентрації вільних електронів в напівпровіднику при різних концентраціях донорів і відповідні залежності положення рівня Фермі.

N_D в області низьких температур зі збільшенням температури росте концентрація вільних електронів за рахунок іонізації донорів (ділянка кривої між точками 1 і 2). Кут нахилу цієї ділянки кривої характеризує енергію іонізації домішки. В даному діапазоні температур рівень Фермі перебуває між дном зони провідності та енергетичними рівнями донорів. При деякій температурі (точка 2) ймовірність заповнення донорних рівнів стає рівною 50 %. Тому рівень Фермі при температурі T_2 збігається з енергетичним рівнем донора.

При подальшому збільшенні температури концентрація вільних електронів практично не збільшується (ділянка між точками 2 і 3), оскільки всі домішки вже іонізовані, а ймовірність іонізації власних атомів напівпровідника ще дуже мала. Ця ділянка називається ділянкою виснаження домішок. Ділянки 1-2 і 2-3 відповідають домішковій електропровідності напівпровідника.

При відносно великих температурах (ділянка кривої після точки 3) концентрація вільних електронів (носіїв заряду) росте за рахунок переходу електронів через заборонену зону. Нахил цієї ділянки кривої характеризує ширину забороненої зони напівпровідника. Рівень Фермі при цих температурах розміщений поблизу середини забороненої зони, а напівпровідник можна вважати власним, оскільки концентрацію носіїв заряду визначає іонізація власних атомів напівпровідника. Принцип дії більшості напівпровідникових приладів базується на використанні домішкової електропровідності, тому виникнення власної електропровідності негативно впливає на задовільну роботу приладу. З цієї причини температура в точці 3 є максимальною робочою температурою напівпровідникового приладу з концентрацією домішки N_D .

Розглянемо зміщення кривих і зміну їх характеру при збільшенні концентрації домішки ($N_D'' > N_D' > N_D$). При збільшенні концентрації домішки кут нахилу ділянки іонізації домішки зменшується внаслідок розщеплення домішкових рівнів і зменшення енергії іонізації домішок. При достатньо високій концентрації домішки енергія іонізації домішки прямує до нуля, оскільки енергетична зона домішки перекривається з зоною провідності. Такий напівпровідник називається виродженим.

Температура переходу від домішкової до власної електропровідності з ростом концентрації домішки зростає ($T_3' > T_3$).

2.4. Температурні залежності концентрації носіїв заряду і питомої провідності

На рухливість носіїв заряду в основному впливають два фізичних фактори: хаотичні теплові коливання атомів кристалічної ґратки (розсіювання носіїв заряду на теплових коливаннях атомів кристалічної ґратки) та

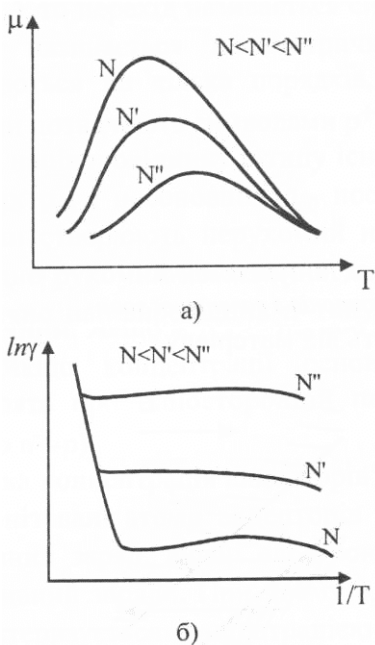


Рис. 2.7. Температурні залежності рухливості носіїв заряду (а) та питомої провідності напівпровідника (б) при різних концентраціях домішки.

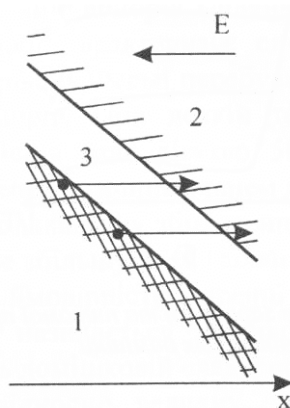


Рис. 2.8. Тунелювання електронів з валентної зони в зону провідності при сильному електричному полі в напівпровіднику: 1 – валентна зона; 2 – зона провідності; 3 – заборонена зона.

електричні поля іонізованих домішок. При високих температурах переважає розсіювання носіїв заряду на теплових коливаннях атомів кристалічної ґратки. Тому зі збільшенням температури в цьому діапазоні температур рухливість носіїв зменшується (рис. 2.7, а).

В діапазоні низьких температур при зменшенні температури зменшуються теплові швидкості хаотичного руху носіїв заряду, що приводить до збільшення часу перебування носія поблизу іона домішки, тобто збільшення тривалості дії електричного поля іона домішки на носій заряду. Тому в діапазоні малих температур зі зменшенням температури рухливість електронів також зменшується.

При зростанні концентрації домішок збільшується і розсіювання на іонах домішок, тобто зменшується рухливість носіїв заряду. Однак у діапазоні високих температур переважає механізм розсіювання носіїв на теплових коливаннях атомів кристалічної ґратки і відповідно криві температурної залежності рухливості носіїв заряду в діапазоні високих температур практично не зміщуються зі збільшенням концентрації домішки.

Питома провідність пропорційна концентрації носіїв заряду та їх рухливості. Температурна залежність питомої провідності схожа на температурну залежність концентрації носіїв при дуже низьких і високих температурах. У діапазоні температур, які відповідають виснаженню домішки, температурні зміни питомої провідності зумовлені температурною залежністю рухливості (рис. 2.7, б).

Напівпровідники в сильних електричних полях, в сильних електричних полях у напівпровіднику можуть відбуватися процеси, які приводять до зміни питомої провідності напівпровідника. Тому розглянемо фізичні процеси, які впливають на концентрацію носіїв заряду.

Під дією високої напруженості електричного поля вільний електрон (або дірка) набуває енергії достатньої для іонізації домішки і власного атома напівпровідника. Процес іонізації атома таким електроном називається ударною іонізацією. Кількісно процес ударної іонізації характеризують коефіцієнтом ударної іонізації, які сильно залежать від напруженості електричного поля

$$a = A \cdot |E|^m,$$

де m залежить від матеріалу.

Для сильного електричного поля в напівпровіднику характерний сильний нахил енергетичних зон (рис. 2.8). При цьому електрони можуть проходити через вузький потенційний бар'єр без зміни своєї енергії – тунелювати завдяки своїм квантово-механічним властивостям. Напруженість електричного поля, при якій відбувається ефект тунелювання, різна для різних матеріалів, тому що товщина потенційного бар'єру залежить від ширини забороненої зони напівпровідника.

Лекція 3

ЯВИЩА ПЕРЕНОСУ В НАПІВПРОВІДНИКАХ

План

- 3.1. Рухливість носіїв заряду
- 3.2. Електропровідність
- 3.3. Квантові ефекти

Література

- 3.1. Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.
- 3.2. Мікроелектронні сенсори фізичних величин: Науково-навчальне видання. В 3 томах. Том 1. / Вуйцік В., Готра З.Ю., Григор'єв В.В., Каліта В., Мельник О.М., Потенцкі Є.; за ред. З.Ю. Готри. – Львів: Ліга-Прес, 2002. – 475 с.
- 3.3. Микроэлектроника. В 9 кн. / Под ред. Л. А. Коледова. Кн. 1. Физические основы функционирования изделий микроэлектроники / О.В. Митрофанов, Б.М. Симонов, Л.А. Коледов.: Высшая школа, 1987. – 168 с.
- 3.4. Степаненко И. П. Основы микроэлектроники. – М.: Сов. радио, 1980. – 286 с.

3.1. Рухливість носіїв заряду

При розгляді поведінки носіїв заряду в твердому тілі найсуттєвішими є кінетичні явища, зв'язані з рухом носіїв заряду під дією електричних, магнітних полів або різниці температур. До них належить електропровідність, термоелектричні, гальваномагнітні та термомагнітні явища.

За наявності одного типу вільних носіїв заряду (електронів) *питома електропровідність* становить:

$$S_n = en\mu_n, \quad (3.1)$$

де e – заряд електрона; n – концентрація вільних електронів; μ_n – дрейфова рухливість електронів, тобто величина, яка чисельно дорівнює середній швидкості їх дрейфу \bar{u} , в електричному полі одиничної напруженості \mathcal{E} , тобто:

$$\mu_n = \bar{u} / \mathcal{E}.$$

Якщо носії заряду - дірки, то міркування аналогічні.

Згідно з елементарною електронною теорією матеріалів “кінцева” дрейфова швидкість електронів - швидкість у момент часу, що передує зіткненню,

$$u_k = at,$$

де a - прискорення електрона під дією електричного поля; τ – час релаксації, тобто середній час між двома зіткненнями електрона з ґраткою (середній час вільного пробігу). Оскільки вважається, що після кожного наступного зіткнення напрямлений рух припиняється, то початкова дрейфова швидкість $u_0 = 0$.

Другий закон Ньютона для електрона у електричному

полі E_0 має такий вигляд:

$$eE_0 = ma.$$

де

$$a = \frac{c}{m} E_0$$

Оскільки $\bar{u} = \frac{u_k + u_0}{2} = \frac{et}{2m} E_0$, то остаточно отримуємо

$$m_n = et / 2m. \quad (3.2)$$

При більш строгому підході, із врахуванням розподілу електронів за швидкостями, отримують:

$$m_n = et / m. \quad (3.3)$$

тобто вдвічі більший результат.

Якщо використовувати формулу (3.3) для одного з кристалографічних напрямів у напівпровіднику, то:

$$m_{ni} = \frac{e}{m_i} t, \quad (3.4)$$

де μ_{ni} – рухливість електрона в одному з напрямків (наприклад, вздовж осі еліпсоїда – ізоенергетичної поверхні для Si, Ge); m_i – відповідний компонент тензора ефективної маси.

У загальному випадку рухливість (і електропровідність) у напівпровідникових кристалах є тензором. Але оскільки в кристалах типу кремнію та германію є симетричні еквівалентні долини, то рухливість (і електропровідність) таких кристалів усереднюється і стає ізотропною:

$$m_{s_n} = \frac{1}{3} (m_{n1} + m_{n2} + m_{n3}).$$

Крім того, за аналогією з (3.4) можна записати:

$$m_{s_n} = \frac{e}{m_{s_n}} t.$$

Отже,

$$\frac{1}{m_{s_n}} = \frac{1}{3} \left(\frac{1}{m_1} + \frac{1}{m_2} + \frac{1}{m_3} \right),$$

тобто отримуємо співвідношення, з якого можна визначити ефективну масу провідності $m_{\sigma n}$. Зокрема можна знайти величину $m_{\sigma n}$ для кремнію та германію. Підрахунок дає значення: $(m_{\sigma n})_{Si} = 0,22 m$; $(m_{\sigma n})_{Ge} = 0,12 m$ (m – маса електрона).

На відміну від елементарної електронної теорії в цьому випадку τ визначається як час релаксації, що характеризує відновлення рівноважного стану, порушеного електричним полем, після зняття останнього.

Можна вважати, що цей час приблизно дорівнює середньому часу вільного пробігу. Дійсно, якщо швидкість напрямленого руху носієм заряду втрачається внаслідок одиничного зіткнення, то це припущення допустиме. Однак і більш строгий підхід призводить до висновку, що час релаксації і середній час вільного

пробігу є величинами одного порядку.

При взаємодії електрона (або дірки) з коливаннями ґратки – фононами може спостерігатися передавання енергії ґратці і народження нового фонона або зворотний процес. При цьому справедливі закони збереження енергії та квазіімпульсу, тобто при народженні фонона маємо:

$$E(k) = E_1(k_1) + \hbar\omega_q$$

$$k = k_1 + q,$$

де $E(k)$ – енергія електрона до зіткнення (до взаємодії з коливанням ґратки); k – відповідний хвильовий вектор; $E_1(k_1)$ – енергія електрона після народження нового фонона з хвильовим вектором q і енергією $\hbar\omega_q$. При поглинанні електронем фонона маємо:

$$E_1(k_1) = E(k) + \hbar\omega_q$$

$$k_1 = k + q,$$

де $E_1(k_1)$ – енергія електрона після поглинання фонона, k_1 – хвильовий вектор цього електрона.

Описаний механізм розсіювання називають однофононим. Оскільки кількість фононів визначається температурою, то і розсіювання електронів на теплових коливаннях ґратки залежить від температури.

Менш вірогідним, однак можливим, є багатофононний механізм розсіювання, коли спостерігається народження або поглинання більше ніж одного фонона.

У реальних кристалах центрами розсіювання служать дефекти структури, відстань між якими сумірна з довжиною хвилі електрона.

Розрахунок часу релаксації для кожного типу дефектів є досить трудомістким. Не вдаючись до деталей, наведемо лише залежності рухливості носіїв заряду в кристалі від температури, розраховані для різних механізмів розсіювання.

При розсіянні тепловими коливаннями ґратки $\mu \sim T^{-3/2}$ для невиродженого газу і $\mu \sim T^{-1}$ для виродженого. При розсіянні на іонізованих домішках $\mu \sim T^{-3/2}$. Розсіювання на нейтральних атомах домішки в кристалі практично не залежить від температури.

За наявності декількох різних типів дефектів у твердому тілі час релаксації визначається з виразу:

$$\frac{1}{t} = \sum_i \frac{1}{t_i},$$

де t_i – час релаксації в кристалі з i -м типом дефектів. При цьому рухливість

$$m = \frac{e}{m^*} \left\langle \sum_i \frac{1}{t_i} \right\rangle,$$

де m^* – ефективна маса носія заряду. Однак з достатнім наближенням можна вважати, що

$$m = \left(\sum_i \frac{1}{m_i} \right)^{-1}, \quad (3.5)$$

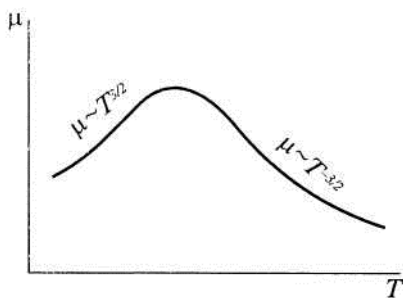


Рис. 3.1. Залежність рухливості від температури в невироджених домішкових напівпровідниках.

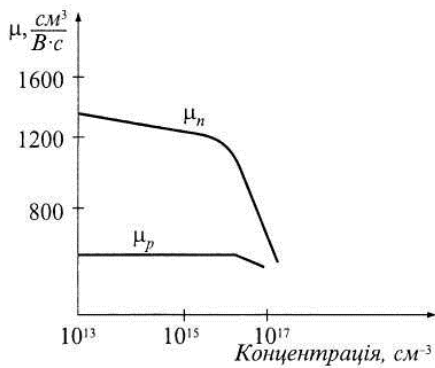


Рис. 3.2. Залежність рухливості електронів μ_n від концентрації донорів N_d і дірок μ_p від концентрації акцепторів N_a в кремнії.

де μ_i - рухливість носіїв заряду в кристалі з i - м типом дефекту.

У невідроджених домішкових напівпровідниках залежність рухливості від температури може бути якісно описана графіком рис. 3.1. За високих температур переважає розсіяння носіїв заряду тепловими коливаннями ґратки і $\mu \sim T^{-3/2}$ за низьких температур визначальним є розсіяння на іонах домішки і $\mu \sim T^{3/2}$.

Результати розрахунків вказують також на залежність рухливості від ефективної маси носіїв заряду. Збільшення ефективної маси призводить до зменшення рухливості і, оскільки ефективна маса дірок більша від ефективної маси електронів, рухливість дірок звичайно менша від рухливості електронів.

Рухливість залежить також від концентрації домішок. Із їх зростанням, як і у газі виникнення іншої недосконалості в ґратках, розсіювання зростає, а рухливість падає. На рис. 3.2 наведено залежність рухливості електронів і дірок від концентрації домішок в кремнії при $T = 300$ К.

Отже, зазвичай зі збільшенням опору цього матеріалу (наприклад, кремнію) в домішковій області рухливість збільшується або більш високоомний зразок має меншу концентрацію домішок.

3.2. Електропровідність

Розглянемо залежність електропровідності напівпровідників від температури і концентрації домішок. У загальному випадку питома електропровідність є сумою, де кожен доданок визначається формулою типу (3.1). Зокрема, якщо носіями заряду є електрони і дірки, то отримуємо:

$$S = enm_n + em_p. \quad (3.6)$$

В області власної електропровідності рівноважні концентрації електронів і дірок однакові тому:

$$S_i = e(m_n + m_p)n_i = e(m_n + m_p)\sqrt{N_c N_v} \exp\left(-\frac{E_g}{2kT}\right), \quad (3.7)$$

де E_g – ширина забороненої зони.

Хоча рухливості μ_n і μ_p , а також ефективні густина станів N_c і N_v залежать від температури, все ж основний вклад у температурну залежність σ_i дає експоненціальний член. Тому коефіцієнт перед експонентою вважатимемо константою, що спостерігається в широких інтервалах температур.

Звідси,

$$\ln S_i = A - \frac{0,43E_g}{2k} \frac{10^3}{T},$$

де $A : \ln\left[(m_n + m_p)\sqrt{N_c N_v}\right]$, тобто в області власної електропровідності $\lg \sigma_i$, є практично лінійною функцією $(1/T)$; причому кутовий коефіцієнт прямої визначається шириною забороненої зони E_g , а саме:

$$tga = \frac{0,43E_g}{2k \cdot 10^3} \text{ і } E_g = 0,4tga, \text{ eВ.}$$

У домішковій області один із доданків (3.7) малий, порівняно з іншими, оскільки концентрація неосновних носіїв заряду набагато менша від концентрації основних носіїв заряду. Звідси для рівноважних концентрацій (3.7) дає одночленну експоненціальну залежність від температури аж до області виснаження домішки або, якщо така відсутня, до області власної електропровідності. В електронному напівпровіднику маємо:

$$s_n = em_n \sqrt{\frac{N_c N_d}{2}} e^{-\frac{E_d}{2kT}},$$

відповідно у дірковому напівпровіднику –

$$s_p = em_p \sqrt{\frac{N_v N_a}{2}} e^{-\frac{E_a}{2kT}},$$

де N_c, N_d, N_v, N_a – густини станів в зоні провідності, донорних рівнів, валентній зоні, акцепторних рівнів відповідно. Тоді

$$\ln s_n = B - \frac{0,43E_d}{2k \cdot 10^3} \cdot \frac{10^3}{T},$$

де $B: \ln m_n \sqrt{\frac{N_c N_v}{2}}$.

Звідки

$$tga_d = \frac{0,43E_d}{2k \cdot 10^3}$$

і

$$E_d = 0,4tga_d, \text{ eВ.} \quad (3.8)$$

Аналогічно

$$\ln s_p = C - \frac{0,43E_a}{2k \cdot 10^3} \cdot \frac{10^3}{T},$$

де

$$C \sim \ln m_p \sqrt{\frac{N_v N_a}{2}}.$$

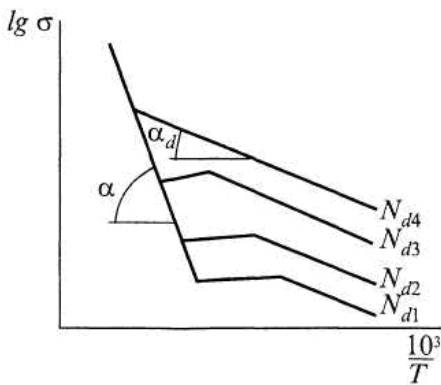
Звідки

$$tga_a = \frac{0,43E_a}{2k \cdot 10^3};$$

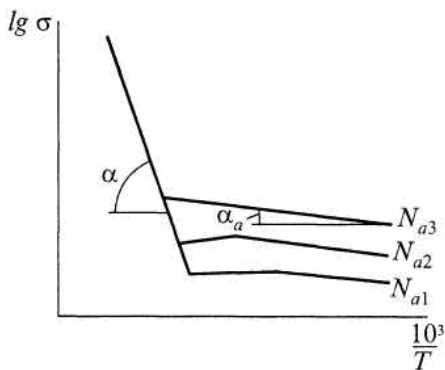
$$E_a = 0,4tga_a, \text{ eВ.} \quad (3.9)$$

Для області виснаження домішок хід функції $\sigma(T)$ визначається залежністю $\mu(T)$. Зокрема, якщо в області виснаження μ зменшується зі збільшенням T , то $\sigma(T)$ також буде зменшуватися.

На рис. 3.3 наведено криві $\lg \sigma = f(1/T)$ для електронного напівпровідника при різній концентрації донорів N_d і для діркового напівпровідника при різній концентрації акцепторів N_a . Якщо в обох випадках розглядається один і той самий напівпровідник (наприклад, кремій), то в області власної



а)



б)

Рис. 3.3. Залежність питомої електропровідності σ від абсолютної температури T в напівпровіднику з різною концентрацією донорів N_d (а); з різною концентрацією акцепторів N_a (б) (концентрація зростає зі збільшенням порядкового номера)

електропровідності крива має однаковий нахил. Кут нахилу кривих у домішковій області визначається енергією іонізації донорів і акцепторів, згідно з (3.8), (3.9).

Однак у реальних напівпровідниках хід кривих може дещо відрізнятись від ідеалізованих кривих, зображених на рис. 3.3.

При концентрації домішки $\sim 10^{15} \text{ см}^{-3}$ наприклад, в германії хвильові функції сусідніх домішкових атомів помітно перекриваються. Оскільки в реальних кристалах практично завжди є частково скомпенсовані домішки при $N_d > N_a$ то при перекритті хвильових функцій сусідніх донорних центрів виникає можливість тунелювання електронів по донорних центрах. За дуже низької температури (при $T \approx 0 \text{ К}$) концентрація зайнятих центрів становить $N'_d = N_d - N_a$, а концентрація вільних центрів – N_a . У цьому випадку навіть при $T = 0 \text{ К}$ можлива електропровідність по домішках (по локалізованих станах). Але домішкові центри розташовані далеко один від одного, то рухливість електронів по домішкових рівнях значно менша від рухливості електронів зони провідності. Тому з підвищенням температури переважає звичайна зонна електропровідність. Але при підвищенні концентрації домішки перекриття хвильових функцій електронів сусідніх домішкових центрів виявляється настільки значним, що одиночний домішковий рівень розширюється, утворюючи в забороненій зоні напівпровідника або діелектрика домішкову зону, так що густина станів виявляється відмінною від нуля в тому місці забороненої зони, де виникає домішкова зона. При подальшому підвищенні концентрації донорів домішкова зона може зімкнутися із зоною провідності (або з валентною зоною для напівпровідника р-типу при підвищенні концентрації акцепторів). Оскільки різні дефекти структури, як і домішкові атоми (як окремих випадок дефектів структури), створюють рівні в забороненій зоні, домішкові зони можливі в неоднорідних напівпровідниках та діелектриках. Зазначимо, що електропровідність по домішковій зоні (по локалізованих станах) може впливати на загальну електропровідність, і масштаби цього впливу залежать від температури, а також від частоти, якщо електричне поле змінне. Хоч звичайно електропровідність по валентній зоні або по зоні провідності значно перевищує електропровідність по домішковій зоні, в ряді випадків останню необхідно враховувати.

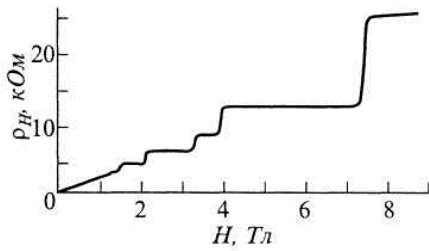


Рис. 3.4. Квантові ефекти Холла – залежність опору ρ_H від H в гетероструктурі $GaAs/Al_{0,3}Ga_{0,7}As$ при $T=8$ мК.

3.3. Квантові ефекти

Одним з яскравих експериментальних відкриттів минулого століття було відкриття фон Клітцингом, Дорда та Пеппером 1980 р. квантування холлівського опору, яке відоме нині як *квантовий ефект Холла* (КЕХ). За це відкриття Клаусу фон Клітцингу була присвоєна Нобелівська премія за 1985 рік.

Суть явища полягає в тому, що в низькотемпературному, при $T \sim 1$ К, холлівському опорі $\rho_H 2D$ (тобто двовимірного) електронного газу в сильному магнітному полі H , вимірюваному залежно від H чи N – концентрації носіїв, спостерігається плато (рис. 3.4).

На плато

$$r_H = \frac{h}{ve^2}, \quad (3.10)$$

тобто холлівський опір виражається через світлові константи h і e та v -фактор заповнення

$$v = \frac{N_S}{N_L},$$

де

$$N_L = \frac{1}{2} p l^2 = \frac{eH}{ch} = \frac{H}{\Phi_0}.$$

Тут N_L – кількість станів Ландау в розрахунку на одиницю поверхні, $l = \sqrt{\frac{ch}{eH}}$ – магнітна довжина,

$\Phi_0 = \frac{ch}{e}$ – квант магнітного потоку.

Унікальною особливістю наведеного є те, що v , визначена на формулою (3.10), дорівнює

$$v = \frac{p}{q}, \quad (3.11)$$

де p і q – цілі числа.

Рівність (3.11) витримується з надзвичайною точністю яка досягала 10^{-5} . Константа з величиною $\frac{h}{e^2} \approx 25,812$ кОм, що фігурує в (3.10), увійшла в теорію твердого тіла ще до відкриття КЕХ. Вона була введена як граничний опір для дротів, при якому ще зберігається металева провідність. Тобто таку константу можна розцінювати як *квант опору* (або $\frac{h}{e^2}$ – як *квант провідності* з урахуванням факту, що провідність $S_H = r_H^{-1}$). Саме в такому розумінні прийнято говорити про квантування ρ_H чи σ_H .

Через два роки після відкриття фон Клітцингом КЕХ ефекту Тсуї, Штермер і Госсард виявили так званий *дробовий квантовий ефект Холла*. Формальною відмінністю між ними є те, що у факторі v (3.11) $p = 1$ для КЕХ і $p > 1$ для дробового КЕХ. Насправді, фізичні

механізми, що відповідають за ці ефекти, є різними. Якщо в першому визначальним є вплив випадкового потенціалу, то в другому – міжелектронна взаємодія.

В обох випадках умовою їх реалізації є двовимірність носіїв. Виникає питання, як практично досягти двовимірності в нашому тривимірному світі. Існує декілька типів систем, в яких електронний (чи дірковий) газ поводить себе як двовимірний:

а) інверсійні і акумульовані шари в кремнієвих МДН (метал-діелектрик-напівпровідник)-структурах. У таких структурах кристал Si має плоску межу розділу з тонкою ($\sim 10^{-5}$ см) плівкою діелектрика SiO₂, на протилежну поверхню якої наноситься металічний електрод-затвор. Прикладене постійне електричне поле V_G між затвором і границею розділу створює в SiO₂ однорідне електричне поле, а в Si біля межі розділу Si/SiO₂ – об'ємний заряд і потенціальну яму, що утримує його. З боку діелектрика яма обмежена потенціальним бар'єром. Тоді тонкий приповерхневий шар ($\approx 30-50 \text{ \AA}$), обмежений стінками, заповнюється електронами – носіями протилежного (інверсного) за знаком відносно до носіїв в об'ємі (дірок). Отже, біля поверхні *p-Si* при $V_G > 0$ виникає інверсійний шар: *n*-канал (біля поверхні *n-Si* при $V_G < 0$ *p*-канал);

б) двовимірні шари носіїв заряду в гетероструктурах – на межі розділу двох напівпровідників з близькими періодами кристалічних ґраток, але різними ширинами заборонених зон. Прикладом може служити гетероструктура GaAs/AlGaAs. Ширина забороненої зони в AlGaAs більша, ніж в GaAs. З легованого сплаву AlGaAs носії тунелюють у порівняно чистий GaAs, в результаті чого хімічний потенціал по обидва боки бар'єра вирівнюється. Через переміщення заряду на межі розділу виникає сильне електричне поле ($E \sim 10^5$ В/см), що породжує потенціальну яму в GaAs – виникає шар двохмірних носіїв заряду *n*- чи *p*-типу (на контакті з *n*-AlGaAs:Si чи *p*-AlGaAs:Be, відповідно).

Рівняння, що пов'язують густину струму \mathbf{j} з електричним полем \mathbf{E} для 2D-систем у випадку гальваномагнітних досліджень у компонентах мають вигляд:

$$\begin{aligned} j_x &= s_{xx} E_x + s_{xy} E_y & E_x &= r_{xx} j_x + r_{xy} j_y \\ j_y &= s_{yx} E_x + s_{yy} E_y & E_y &= r_{yx} j_x + r_{yy} j_y \end{aligned} \quad \text{чи}$$

де $\sigma_{\alpha\beta}$, $\rho_{\alpha\beta}$ – тензори питомих провідності та опору, відповідно: σ_{xy} , ρ_{xy} – холлівські компоненти, які для кристалів кубічної симетрії задовольняють умову:

$$s_{xx} = s_{yy}; \quad r_{xx} = r_{yy};$$

а згідно зі співвідношеннями Онзагера

$$s_{xy} = -s_{yx}; \quad r_{xy} = -r_{yx}.$$

З умови $\hat{\sigma}\hat{\rho} = \hat{1}$ ($\hat{1}$ – одиничний діагональний тензор) випливає:

$$r_{xx} = \frac{S_{xx}}{S_{xx}^2 + S_{xy}^2}; r_{xy} = \frac{S_{yx}}{S_{xx}^2 + S_{xy}^2}. \quad (3.12)$$

За умов квантового ефекту Холла $S_{xx} \ll S_H$ (σ_H, ρ_H – абсолютне значення (σ_{xx}, ρ_{xx} відповідно), і тоді (3.12) набирає вигляд:

$$S_{xx} = r_{xx} = 0, S_H = r_H^{-1}.$$

Відомо, що при накладанні на кристал зовнішнього поля паралельно до Oz електронний спектр носіїв у площині xOy стає дискретним – зона розбивається на *рівні Ландау*. У випадку 2D-шару, що лежить у такій площині, густина станів $D(\varepsilon)$ є системою δ -функцій (рис. 3.5). Присутність в реальному кристалі випадкових потенціалів, спричинених неоднорідностями, домішками тощо, призводить до розмиття δ -функцій у вузькі піки. Експерименти показали, що між піками густина станів $\rho(\varepsilon)$ має майже постійне значення.

Саме по собі знання густини станів $D(\varepsilon)$ не вирішує питання про поведінку хвильових функцій, що є визначальною у проблемі існування електропровідності та її механізмів.

Квантові стани можуть бути локалізованими і колективізованими: хвильові функції локалізованих станів загасають з відстанню, тоді як колективізовані не загасають до нескінченності. З цього погляду останні нагадують собою блохівські функції ідеального кристала, але відрізняються від них своїм аналітичним виглядом, що визначається як магнітним полем, так і сильним впливом потенціалу дефекту. Відомо, що навіть короткодійчий потенціал ізольованої домішки відщепляє від рівня Ландау нескінченна кількість рівнів, що відповідають локалізованим станам. Існування колективізованих станів, що підтверджується експериментально, забезпечує холлівський струм.

Області локалізованих станів називаються *щільностями рухливості*, а їх межі з областями колективізованих (чи протяжних) станів – *краями чи порогами рухливості*. При $T \neq 0$ К стани, першопочатково локалізовані, вносять внесок у провідність за рахунок термічно активованих процесів. У результаті локалізація поступово руйнується.

Якісне пояснення цілочисельного квантового ефекту Холла можна отримати на наведеній вище структурі електронних станів. Рівень Фермі зсувається зі зміною N_s і H . При цьому складові провідності σ_{xx} і σ_{xy} демонструють принципово різну поведінку. Відомо, що в нескінченному зразку σ_{xx} визначається тільки електронами з околу рівня Фермі, тоді як σ_{xy} (холлівська провідність) – електронами з усіх заповнених станів. Тому, коли рівень Фермі знаходиться в щільності рухливості, $\sigma_{xx} = 0$ при $T=0$ К. Згідно з (3.12) при цьому $\rho_{xx} = 0$ і струм протікає без розсіювання.

При $T=0$ К σ_{xx} відмінна від нуля, але експоненціально

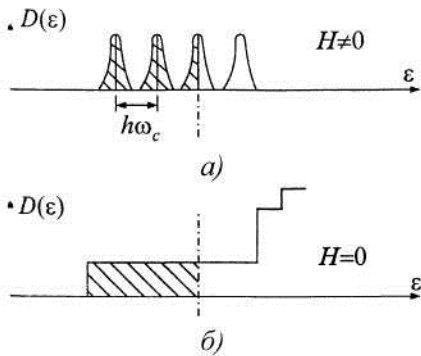


Рис. 3.5. Густина станів електрона у зовнішньому магнітному полі: а- з полем., б - без поля.

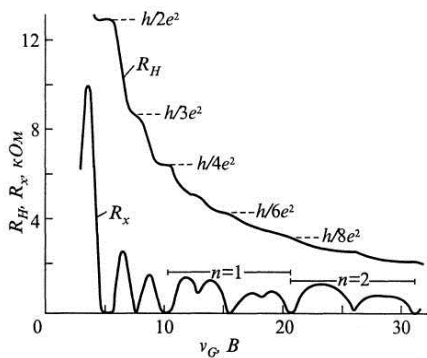


Рис. 3.6. Квантовий ефект Холла в Si – МДН-структурі залежність холлівського опору R_H і опору R_x від напруги на затворі v_G (v_G пропорційна N_s).

мала, така провідність визначається або активаційним збудженням електронів і дірок в колективізовані (протяжні) стани, що належать сусіднім рівням Ландау або стрибками (рис. 3.6). Цим пояснюються глибокі провали в σ_{xx} і ρ_{xy} при $T=0$ К. Навпаки, $\sigma_H \neq 0$ навіть при $T=0$ К, оскільки за рахунок дрейфу носіїв у схрещених електричному (\mathbf{E}) і магнітному (\mathbf{H}) полях у цю провідність дають внесок протяжні стани на всіх рівнях Ландау, що є нижче від рівня Фермі. Так, при зміні рівня Фермі в межах щілини рухливості заповнення протяжних станів залишається практично постійним ($T \rightarrow 0$ К). Тому при цьому і σ_H залишається постійною, що й пояснює появу плато на залежності $S_H = (N_s)$ і $S_H = (N)$.

Коли рівень Фермі проходить через протяжні стани, ситуація міняється корінним чином, оскільки імовірність переходів з розсіюванням носіїв втрачає експоненціальну малість і σ_{xx} вже досягає величини, співвимірної з σ_H . В цій же області швидко міняється і σ_H , оскільки міняється кількість протяжних станів під рівнем Фермі.

Підсумовуючи, можна сказати, що основні властивості цілочисельного квантового ефекту Холла можна пояснити без залучення розгляду взаємодії між носіями. Природно припустити, що ефекти взаємодії електронів, що належать одному рівневі Ландау $\sim \frac{e^2}{\epsilon l}$, ϵ – діелектрична проникливість) будуть у мінімальній мірі ослаблені поляризацією інших станів. Це спостерігається за низьких концентрацій і/або сильних магнітних полях. Реалізація таких вимог можлива на досконалих зразках за температур, нижчих від тих, при яких спостерігаємо цілочисельний квантовий ефект Холла.

Лекція 4

ТЕРМОЕЛЕКТРИЧНІ ЯВИЩА

План

- 4.1. Ефект Зеєбека
- 4.2. Контактна складова термо-ЕРС
- 4.3. Об'ємна складова термо-ЕРС
- 4.4. Ефект Пельтьє

Література

- 4.1. Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.
- 4.2. Мікроелектронні сенсори фізичних величин: Науково-навчальне видання. В 3 томах. Том 1. / Вуйцік В., Готра З.Ю., Григор'єв В.В., Каліта В., Мельник О.М., Потенцкі Є.; за ред. З.Ю. Готри. – Львів: Ліга-Прес, 2002. – 475 с.
- 4.3. Мікроелектроніка. В 9 кн. / Под ред. Л. А. Коледова. Кн. 1. Физические основы функционирования изделий микроэлектроники / О.В. Митрофанов, Б.М. Симонов, Л.А. Коледов.: Высшая школа, 1987. – 168 с.
- 4.4. Степаненко И. П. Основы микроэлектроники. – М.: Сов. радио, 1980. – 286 с.

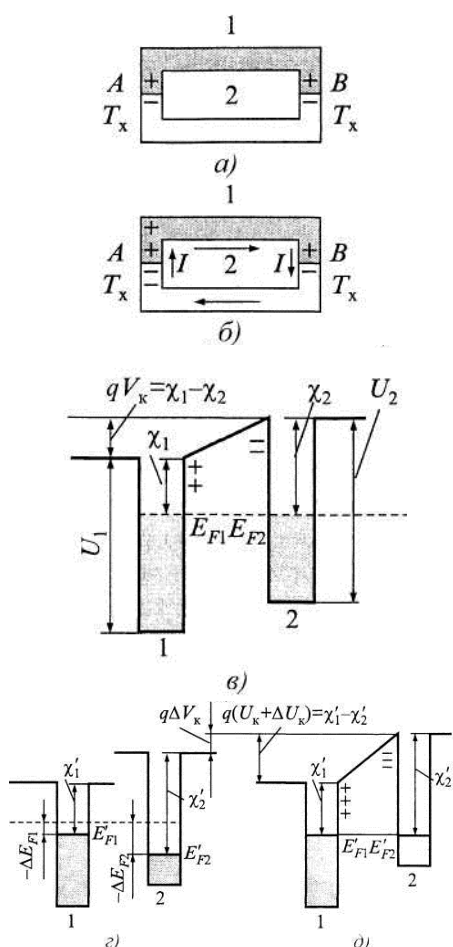


Рис. 4.1. Виникнення термоелектрорушійної сили.

4.1. Ефект Зеєбека

В 1823 р. Т. Зеєбек встановив, що в електричному колі, яке складається з двох різнорідних провідників 1 і 2, виникає електрорушійна сила U_m , якщо контакти цих провідників А і В підтримуються при різних температурах T_r і T_x (рис. 4.1, б). Ця ЕРС називається *термоелектрорушійною силою*. Як показує експеримент, у порівняно незначному діапазоні температур вона пропорційна різниці температур контактів А і В:

$$U_T = a(T_r - T_x). \quad (4.1)$$

Коефіцієнт пропорційності

$$a = \frac{dU_T}{dT} \quad (4.2)$$

називають *диференційною* або *питомою термо-ЕРС*. Він залежить від природи провідників, які дотикаються, і від температури.

За наявності різниць температур спаїв у колі з термоелементом з'являється термоелектрорушійна сила (термо-ЕРС), яка складається з трьох складових.

Перша складова термо-ЕРС зумовлена дифузією носіїв заряду від нагрітого спаю, температура якого внаслідок підведеної до нього теплової потужності від будь-якого джерела вища від температури тепловидільного спаю. Дифузія носіїв заряду в гілках термоелемента може виникати з двох причин. По-перше, у гілках термоелемента біля нагрітого спаю виявляється більша кількість іонізованих атомів. При додатковій іонізації збільшується концентрація основних носіїв

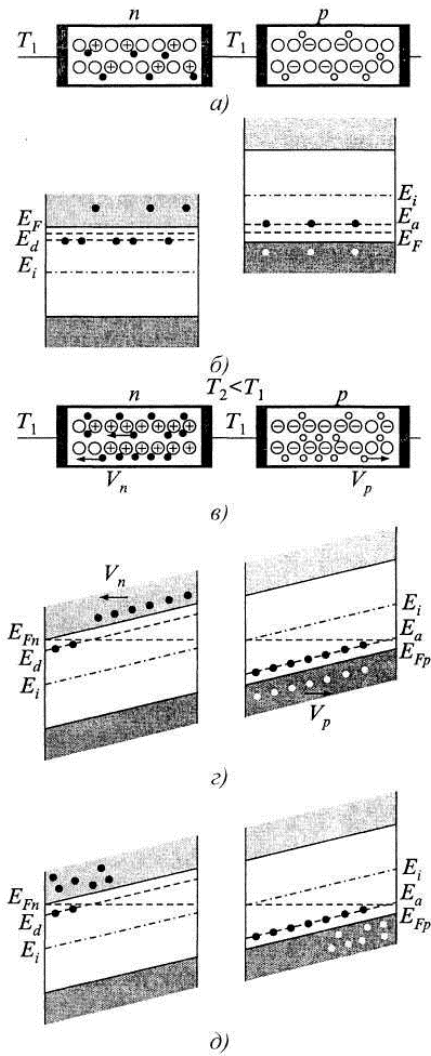


Рис. 4.2. Діаграма термо-ЕРС: а – спаї термоелемента при однаковій температурі T (не всі домішки іонізовані); б – в стані термодинамічної рівноваги (при T_1); в – спаї при різних температурах (поблизу спаїв всі домішки іонізовані); г – при опорі навантаження $R_H = 0$ (струм дорівнює струму короткого замикання термоелемента); д – при опорі навантаження R_H (виникла термо-ЕРС).

заряду на нагрітих кінцях гілок термоелемента. У цьому випадку виникає дифузія основних носіїв у кожній гілці термоелемента внаслідок градієнта концентрації (рис. 4.2). По-друге, якщо в гілках термоелемента всі атоми були іонізовані вже при низькій температурі (ця температура вища від температури виснаження атомів), то концентрація основних носіїв заряду під час нагрівання практично не збільшиться. Однак на нагрітих кінцях гілок термоелемента носії заряду набувають більшої енергії. Тому знову відбувається дифузія основних носіїв заряду від нагрітого кінця у кожній гілці термоелемента, зв'язана з вирівнюванням середньої енергії, що припадає на один носій визначеного знака.

Дифузія, наприклад, електронів у від'ємній гілці може відбуватися тільки від нагрітого кінця цієї гілки вздовж неї і не може переходити у додатну гілку, тому що перешкоджає потенціальний бар'єр нагрітого спаю термоелемента. Аналогічно, дірки у додатній гілці можуть дифундувати вздовж неї від нагрітого кінця.

Переміщення носіїв заряду, пов'язане з їх дифузією, порушує електричну нейтральність у гілках термоелемента – на нагрітих кінцях гілок термоелемента залишаються некомпенсовані іонізовані домішкові атоми, а на протилежних кінцях утворюється надлишок основних носіїв заряду. У результаті виникає перша складова термо-ЕРС, яку можна назвати дифузійною, оскільки вона зумовлена процесами дифузії.

Друга складова термо-ЕРС – це наслідок температурної залежності контактної різниці потенціалів. Якщо обидва спаї термоелемента мають однакову температуру, то контактні різниці потенціалів на цих спаях рівнонаправлені в протилежні боки при обході контура з термоелементом і не дають результуючої ЕРС. Якщо ж температура спаїв термоелемента різна, то значення контактної різниці потенціалів на спаях буде також різне (рис. 4.3). Тоді у колі термоелемента виникає друга складова термо-ЕРС з тією ж полярністю, що і перша складова.

Третя складова термо-ЕРС виникає у термоелементі внаслідок захоплення носіїв заряду квантами теплової енергії – фононами. Якщо у гілках термоелемента є градієнт температури, виникає направлений рух фононів від нагрітих кінців гілок. У результаті зіткнень фононів з носіями заряду фонони тягнуть за собою електрони у від'ємній гілці і дірки у додатній гілці. Цей ефект може виявитися переважаючим за низьких температур.

Результуюча термо-ЕРС, що складається з трьох розглянутих складових, залежить від різниці температур спаїв термоелемента і від електрофізичних властивостей напівпровідникових матеріалів, що утворюють гілки термоелемента. У невеликому діапазоні температур термо-ЕРС є можна вважати (з достатньою для

практичної мети точністю) пропорційною різниці температур спаїв термоелемента і деякому коефіцієнту $\alpha_{1,2}$, який називається коефіцієнтом термо-ЕРС:

$$a_r = a_{1,2} \Delta T.$$

Існують два основні джерела виникнення термо-ЕРС: зміна контактної різниці потенціалів при зміні температури (контактна складова U_K) і утворення направленого руху носіїв в провіднику за наявності градієнта температури (об'ємна складова $U_{об}$).

4.2. Контактна складова термо-ЕРС

На рис. 4.1, а показано електричне коло, яке складається з двох різнорідних провідників, які перебувають в контакті один з одним при температурі T_x ; на рис. 4.1, в – зони провідності цих провідників в рівноважному стані. Рівні Фермі E_{F1} і E_{F2} встановлюються на одній висоті, і між провідниками виникає контактна різниця потенціалів,

$$U_k = \frac{c_2 - c_1}{q},$$

де χ_1 і χ_2 – термодинамічні роботи виходу електронів з провідників. При однаковій температурі контактів А і В різниці потенціалів однакові за величиною і направлені назустріч одна одній. Тому вони зрівноважують одна одну і результуюча напруга дорівнює нулю.

Нагріємо тепер контакт А до температури $T_2 > T_x$, залишаючи контакт В при температурі T_x . В нормальних металах і електронних напівпровідниках підвищення температури викликає пониження рівня Фермі. Позначимо це пониження в провіднику 1 через $-\Delta E_{F1}$, в провіднику 2 через $-\Delta E_{F2}$. Зміна положення рівня Фермі призводить до зміни термодинамічної роботи виходу, яка в провіднику 1 дорівнює $c'_1 = c_1 - \Delta E_{F1}$, в провіднику 2 $c'_2 = c_2 - \Delta E_{F2}$. Якщо $\Delta E_{F1} \neq \Delta E_{F2}$, то при цьому рівні Фермі в провідниках 1 і 2 у гарячому контакті виявляються різні, що свідчить про порушення рівноваги в контакті, яка була встановлена при температурі T_x (рис. 4.1, з). Нова рівновага досягається за рахунок перетікання електронів з першого провідника в другий до встановлення рівнів Фермі на одній висоті (рис. 4.1, д). Різниця потенціалів, яка тепер виникла, дорівнюватиме

$$U_k + \Delta U_k = \frac{c'_2 - c'_1}{q} = \frac{c_2 - c_1}{q} - \frac{\Delta E_{F2} - \Delta E_{F1}}{q}.$$

Вона є більшою, ніж в контакті В, який перебуває при температурі $T_x < T_r$, на величину

$$\Delta U_k = \frac{\Delta E_{F2} - \Delta E_{F1}}{q}. \quad (4.3)$$

Отже, підвищення температури одного з контактів призводить до зміни контактної різниці потенціалів в ньому і виникнення між гарячим і холодним контактом різниці потенціалів ΔU_k , яка і являє собою контактну

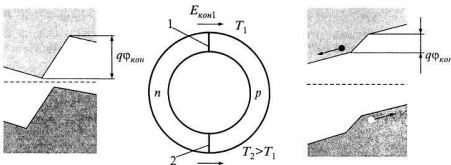


Рис. 4.3. Виникнення термо-ЕРС внаслідок температурної залежності контактної різниці потенціалів під час короткого замикання гілок термоелемента і енергетичні діаграми спаїв, які перебувають при різних температурах.

складову термо-ЕРС.

Позначимо різницю температур між контактами через ΔT , тоді ΔE_{F1} і ΔE_{F2} можна записати наступним чином:

$$\Delta E_{F1} = (\partial E_{F1} / \partial T) \Delta T ;$$

$$\Delta E_{F2} = (\partial E_{F2} / \partial T) \Delta T .$$

Підставляючи в (4.3), отримаємо:

$$\Delta U_k = \frac{1}{q} \left(\frac{\partial E_{F2}}{\partial T} - \frac{\partial E_{F1}}{\partial T} \right) \Delta T . \quad (4.4)$$

Звідси легко визначити контактну складову питомої термо-ЕРС електричного кола, яка складається з провідників 1 і 2 (рис. 4.1, б):

$$a_k^{1,2} = \frac{dU_k}{dT} = \frac{1}{q} \left(\frac{\partial E_{F2}}{\partial T} - \frac{\partial E_{F1}}{\partial T} \right) . \quad (4.5)$$

Вона дорівнює різниці контактних складових

$$a_k = \frac{1}{q} \frac{\partial E_F}{\partial T} \quad (4.6)$$

окремих провідників, які утворюють це коло. Для металів енергія E_F дорівнює

$$E_F(T) = E_F(0) \cdot \left[1 - \frac{p^2}{12} \left(\frac{kT}{E_F(0)} \right)^2 \right],$$

де $E_F(0)$ – енергія Фермі при 0 К. Диференціюючи E_F по T і підставляючи $\partial E_F / \partial T$ в (4.6), знаходимо:

$$a_k = \frac{p^2 k^2}{6q E_F(0)} T . \quad (4.7)$$

Для невироджених напівпровідників n-типу енергія Фермі дорівнює

$$E_{Fn} = kT \ln \frac{n}{N_c} = kT \ln \frac{nh^3}{2(2pm_n kT)^{3/2}} .$$

Диференціюючи E_F по T і підставляючи в (4.6), отримуємо

$$a_k = \frac{k}{q} \left(\frac{E_{Fn}}{kT} + T \frac{d(\ln n)}{dT} - \frac{3}{2} \right) . \quad (4.8)$$

4.3. Об'ємна складова термо-ЕРС

Концентрація носіїв заряду в провіднику залежить від температури. У металів ця залежність дуже слабка і зумовлена термічним розширенням, яке викликає зміну об'єму провідника. У напівпровідників навпаки, при збільшенні температури концентрація носіїв може істотно зростати. Тому на гарячому кінці напівпровідника концентрація носіїв може виявитися вищою, ніж на холодному, внаслідок чого від гарячого кінця до холодного виникає дифузійний потік, який в напівпровіднику n-типу призводить до утворення на холодному кінці від'ємного об'ємного заряду, а на гарячому – додатного. Ці заряди створюють різницю

потенціалів $U_{об}$ – об’ємну складову термо-ЕРС.

У стаціонарному стані в кожній точці провідника дифузійний потік носіїв зрівноважується дрейфовим потоком, викликаний різницею потенціалів $U_{об}$ і який направлений від гарячого кінця до холодного. Для n -напівпровідника дифузійний потік носіїв дорівнює $-D_n(dn/dx)$ (D_n – коефіцієнт дифузії електронів); а dn/dx – градієнт їх концентрації; дрейфовий потік дорівнює $n\mathbf{v}_d = -n\mathbf{m}_n\mathbf{E}$, де v_d – дрейфова швидкість; μ_n – рухливість електронів; ε – напруженість поля, яке виникло в напівпровіднику. У стаціонарних умовах:

$$n\mathbf{m}_n\mathbf{E} + D_n(dn/dx) = 0. \quad (4.9)$$

Розв’язування цього рівняння призводить до наступного результату для об’ємної складової термо-ЕРС, яка зумовлена дифузією носіїв заряду і яка виникає внаслідок зміни концентрації носіїв заряду з температурою:

$$a_{об} = \frac{dU_{об}}{dT} = -\frac{k}{q}T \frac{d(\ln n)}{dT}. \quad (4.10)$$

Підкреслимо ще раз, що у металів, в яких концентрація носіїв надзвичайно слабо залежить від температури, ця складова практично дорівнює нулю; вона також буде малою в напівпровідниках в температурному інтервалі виснаження домішок.

Існує ще одна причина виникнення об’ємної складової термо-ЕРС – інтенсифікація теплового руху носіїв заряду у разі збільшення температури. Зростання температури викликає підвищення середньої енергії носіїв, а, відповідно, і зміни довжини їх вільного пробігу λ , яка таким чином пов’язана з енергією носіїв E :

$$I \sim E'. \quad (4.11)$$

Показник r залежить від механізму розсіювання: при розсіюванні носіїв на теплових коливаннях ґратки $r = 0$, при розсіюванні на іонізованих домішках $r = 2$ і т. д.

З (4.11) видно, що збільшення температури провідника, яке викликає збільшення енергії носіїв, повинно призводити в загальному випадку до збільшення довжини їх вільного пробігу, а, відповідно, і до зміни коефіцієнта дифузії, пропорційного λ . Тому від гарячого кінця до холодного встановиться термодифузійний потік носіїв, який призводить до формування додатної різниці потенціалів $V'_{об}$. Розрахунок показує, що для невироджених напівпровідників n -типу термодифузійна складова термо-ЕРС:

$$a'_{ia} = \frac{k}{q} \left(r + \frac{1}{2} \right). \quad (4.12)$$

Підсумовуючи (4.8), (4.10) і (4.12), отримуємо результуючу питому термо-ЕРС напівпровідника n -типу:

$$a_n = -\frac{k}{q} \left(r + 2 - \frac{E_{Fc}}{kT} \right), \quad (4.13)$$

де E_{Fc} - відстань від рівня Фермі до зони провідності. Аналогічно для напівпровідника р-типу:

$$a_p = -\frac{k}{q} \left(r + 2 - \frac{E_{Fp}}{kT} \right), \quad (4.14)$$

де через E_{Fp} відстань від рівня Фермі до валентної зони.

У металів термодифузійна складова термо-ЕРС також не дорівнює нулю. Враховування її і контактної складової (4.7) призводить до наступного виразу для питомої термо-ЕРС металу з параболічною залежністю енергії від хвильового вектора:

$$a_i = -\frac{p^2 k}{3q} (r+1) \frac{kT}{E_F}. \quad (4.15)$$

Для електричного кола, яке складається з напівпровідника і металу, внеском в термо-ЕРС металом, можна знехтувати і вважати, що вся термо-ЕРС зумовлена напівпровідником.

Необхідно відзначити, що γ , яке входить в (4.15), може бути і від'ємною величиною. У цьому випадку зі збільшенням температури довжина вільного пробігу не зростає, а, навпаки, спадає, внаслідок чого термодифузійний потік направлений не від гарячого до холодного, а від холодного до гарячого кінця провідника і зумовлена ним різниця потенціалів направлена протилежно до контактної складової термо-ЕРС. У цьому випадку результуюча термо-ЕРС може виявитися такою, яка дорівнює нулю або навіть змінити свій напрямок. Така картина спостерігається, зокрема, в ряді перехідних металів і сплавів (ніхром, хромель та ін.).

Для напівпровідників зі змішаною провідністю, в яких електричний струм переноситься одночасно електронами і дірками, термо-ЕРС визначається таким співвідношенням:

$$a_{i\delta} = \frac{a_n m_n n + a_p m_p p}{n m_n + p m_p}. \quad (4.16)$$

Для власних напівпровідників $n=p=n_i$

$$a_i = \frac{a_n m_n + a_p m_p}{m_n + m_p}. \quad (4.17)$$

Нехтуючи у виразах (4.13) і (4.14) доданком $(r+2)$ в порівнянні з E_F/kT і пам'ятаючи, що у власному напівпровіднику $E_{Fi} \approx E_g/2$, отримуємо:

$$a_i \approx \frac{E_g}{2qT} \frac{m_n - m_p}{m_n + m_p}. \quad (4.18)$$

Оскільки переважно $\mu_n > \mu_p$, то a_i у власному напівпровіднику є величиною від'ємною. В акцепторному напівпровіднику при переході до власної провідності відбувається зміна додатного знака термо-ЕРС на від'ємний.

На рис. 4.4 наведено залежність термо-ЕРС акцепторного напівпровідника від температури. В області

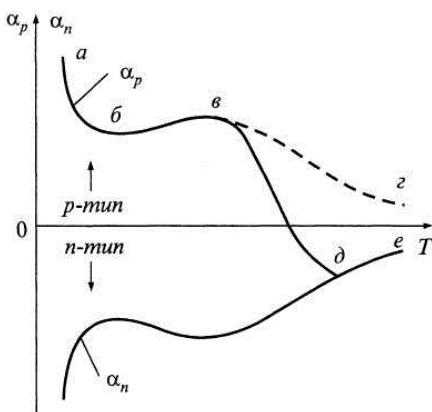


Рис. 4.4. Залежність термо-ЕРС акцепторного і донорного напівпровідників від температури.

низьких температур, в якій рівень Фермі порівняно слабо мінється з температурою, величина $-E_{Fp}/kT$ буде приблизно обернено пропорційна температурі T . Тому і питома термо-ЕРС $\alpha_p \sim 1/T$ (ділянка $a-b$ кривої на рис. 4.4). В області виснаження домішки $-E_{Fp}/kT = \ln(N_v/N_a) = \ln N_v - \ln N_a$. Оскільки $N_v \sim T^{3/2}$, то в цій області α_p незначно росте при підвищенні температури (ділянка $b-c$). В області переходу до власної провідності μ стає величиною практично постійною і такою, яка дорівнює μ_i тому α_p знову зменшується при зростанні температури і крива повинна би поводити себе так, як показано на відріжку $c-d$ на рис. 4.4. Насправді ж внаслідок того, що в цій області концентрації електронів і дірок практично зрівнюються, термо-ЕРС різко спадаючи, переходить через нуль і стає від'ємною (відрізок $d-e$).

При подальшому зростанні температури абсолютна величина термо-ЕРС, згідно з (4.18), зменшується (крива $d-e$ для власного напівпровідника).

У нижній частині рис. 4.4 наведено криву зміни термо-ЕРС донорного напівпровідника з температурою. В області власної провідності вона зливається з кривою для напівпровідника p -типу.

Термоелектричний ефект отримав широке практичне застосування, в тому числі і в радіоелектроніці. Він дає змогу безпосередньо перетворювати теплову енергію в електричну, що використовується в термогенераторах. Теорія таких генераторів була розроблена А.Ф. Йоффе. Згідно з цією теорією ККД перетворення теплової енергії в електричну визначається величиною $\alpha^2 \sigma / K$, де K – коефіцієнт теплопровідності напівпровідника; σ – питома електропровідність.

Якісно це стає зрозумілим з таких міркувань. У термогенераторах прагнуть отримати найбільший перепад температур між гарячим і холодним кінцями провідника при якомога меншій затраті теплової енергії. Чим нижча теплопровідність провідника, тим більша, відповідно, термо-ЕРС. При цьому зменшувати теплопередачу від гарячого кінця до холодного за рахунок видовження провідника не можна, оскільки при цьому буде зростати внутрішній опір термогенератора і ККД буде спадати. З цієї ж причини вигідно мати максимальну електропровідність σ напівпровідника. Оскільки при зростанні ступеня легування напівпровідника σ спадає, а K і α зростають, то для кожного напівпровідника існує оптимальний ступінь легування, який забезпечує максимальну величину $\alpha^2 \sigma / K$, а відповідно і ККД.

Термоелектричний ефект використовується також для вимірювання температур (термопари) і при інших вимірюваннях, які можуть бути зведені до вимірювання температури. У теплових фотоприймачах (термоелементах) світло поглинається зачорненою приймальною площадкою, до якої приєднано спай

термопари, і світло нагріває їх. За значенням термо-ЕРС, яка виникає, визначається сила струму. У вакууметрах через металевий провідник, до середини якого приєднано спай термопари, пропускається фіксований струм. Температура спаю буде різною залежно від теплопровідності газу, який оточує провідник. Теплопровідність, в свою чергу, визначається тиском газу. Тому, вимірюючи термо-ЕРС, що виникає, можна визначити тиск газу. Цим методом зручно вимірювати тиски в діапазоні 10^{-1} -10 Па.

У техніці вимірювань електрофізичних параметрів напівпровідників термоелектричний ефект використовується для визначення переважного типу провідності (за знаком термо-ЕРС) і ширини забороненої зони.

4.4. Ефект Пельтьє

Ефект Пельтьє полягає в тому, що у разі пропускання струму через контакт двох різнорідних матеріалів в додаток до джоулевого тепла в контакті виділяється або поглинається тепло, кількість якого Q_{Π} пропорційна заряду $q = It$, який пройшов через контакт:

$$Q_{\Pi} = \pm \Pi It. \quad (4.19)$$

Коефіцієнт пропорційності Π називається коефіцієнтом Пельтьє.

Пояснимо природу цього ефекту на прикладі нейтрального контакту напівпровідника з металом. Нагадаємо, що *нейтральним* називають контакт двох матеріалів, які володіють однаковими роботами виходу. У такому контакті відсутні шари збагачення і збіднення, немає вигину зон. На рис. 4.5, а наведено енергетичну діаграму нейтрального контакту металу з напівпровідником n-типу. У рівноважному стані рівні Фермі металу (E_{F_M}) і напівпровідника (E_{F_n}) розташовуються на одній висоті, а дно зони провідності напівпровідника знаходиться вище рівня Фермі металу на $-E_{F_n}$, так що для електронів, які переходять з металу в напівпровідник існує потенціальний бар'єр заввишки $-E_{F_n}$.

Прикладемо до такого контакту зовнішню різницю потенціалів, як показано на рис. 4.5, б. Під дією цієї різниці потенціалів в напівпровіднику, на якому спадає практично вся прикладена напруга, рівень Фермі, дно зони провідності й інші енергетичні рівні будуть зазнавати поступового піднесення на висоту qV при переміщенні від лівого контакту до правого. У системі виникає струм, направлений зліва направо; потік електронів, показаний на рис. 4.5, б стрілками, направлений справа наліво.

З рис. 4.5, б видно, що електрони, які переходять в лівому контакті з напівпровідника в метал, мають більшу енергію, ніж в металі, на величину

$$-E_{F_u} + \bar{E}_u,$$

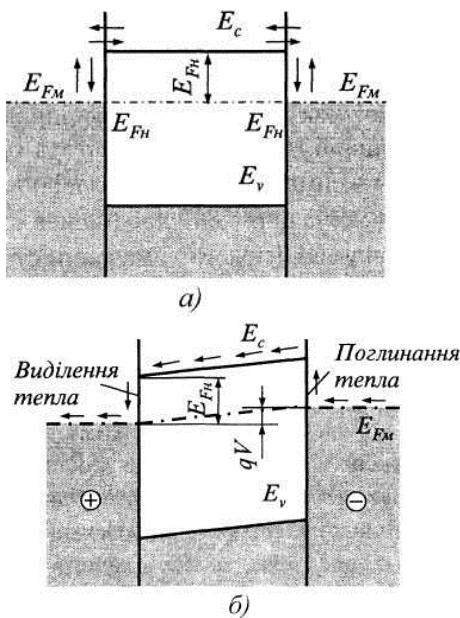


Рис. 4.5. До пояснення ефекту Пельтьє: а – енергетична діаграма кола метал-напівпровідник-метал в стані теплової рівноваги; б – те саме при пропусканні струму.

де $-E_{Fn}$ – висота бар'єра, з якого скочуються електрони, що надходять з напівпровідника в метал; \bar{E}_n – середня енергія електронів у напівпровіднику, які беруть участь у створенні струму. Вона не дорівнює середній тепловій енергії електронів $3/2 kT$, як може здатися з першого погляду, оскільки відносна роль швидких електронів у формуванні електричного струму є вищою, ніж повільних. Для невиродженого електронного газу розрахунок показує, що

$$\bar{E}_n = (r + 2)kT, \quad (4.20)$$

де r – показник степеня у співвідношенні (4.11).

Отже, кожний електрон, який переходить в лівому контакті з напівпровідника в метал, переносить надлишкову енергію:

$$\Delta E = -E_{Fn} + (r + 2)kT, \quad (4.21)$$

яка виділяється в цьому контакті в формі додаткового (до Джоульовського) тепла. Це тепло являє собою теплоту Пельтьє. Поділивши (4.21) на заряд електрона q , отримаємо коефіцієнт Пельтьє

$$|\Pi| = \frac{k}{q} \left(r + 2 - \frac{E_{Fn}}{kT} \right) T. \quad (4.22)$$

У правому контакті електрони переходять з металу в напівпровідник, переборюючи потенціальний бар'єр $-E_{Fn}$. Крім цього, для встановлення рівноваги цих електронів з електронним газом у провіднику їм необхідно набути ще енергію \bar{E}_{in} ; всю цю енергію вони черпають з ґратки, охолоджуючи її в правому контакті.

При проходженні постійного струму через термоелемент в його спаях поглинається або виділяється, залежно від напрямку струму, деяка кількість тепла, пропорційна часу, струму та коефіцієнту Пельтьє:

$$Q_n = \pm \Pi_{1,2} I t. \quad (4.23)$$

Якщо напрямок струму в колі з термоелементами такий, як показано на рис. 4.6, то на спаї 1 вільні електрони і дірки, що виникли в результаті теплової генерації, рухаються в різні боки під дією сумарного електричного поля (дифузійного поля спаю та зовнішнього поля). При тепловій генерації носіїв заряду в області спаю 1 на перекидання електрона з валентної зони в зону провідності витрачається деяка теплова енергія кристалічної ґратки напівпровідника. Тому при заданому напрямку струму спай 1 буде охолоджуватися, спай 2 при цьому нагрівається, тому що до нього з різних боків підходять електрони і дірки, які віддають при рекомбінації деяку енергію у вигляді тепла.

У результаті при проходженні струму термоелемент працює як своєрідний тепловий насос, забираючи теплову енергію на спаї 1 (теплопоглинальному спаї за даним напрямком струму) і виділяючи її на спаї 2 (тепловидільному спаї).

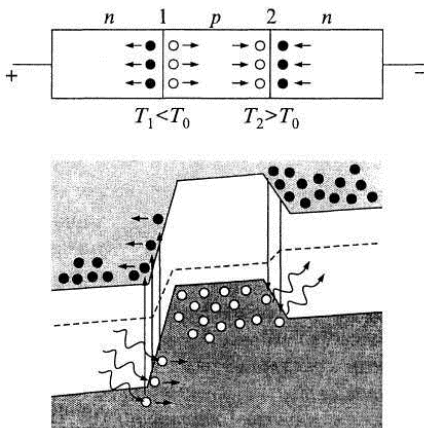


Рис. 4.6. Охолодження спаю (1) і нагрівання спаю (2) напівпровідникових термоелементів при проходженні струму та енергетичні діаграми, які пояснюють ці явища.

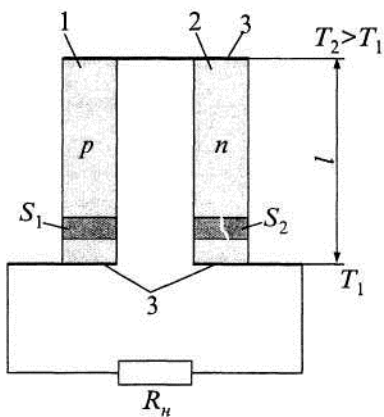


Рис. 4.7. Схема напівпровідникового термоелемента з опором навантаження: 1 — додатна гілка; 2 - від'ємна гілка; 3 - металеві контактні пластини.

Ефект Пельтьє обернений ефекту Зеебека. Тому для одного і того ж термоелемента існує зв'язок між коефіцієнтом Пельтьє та коефіцієнтом термо-ЕРС:

$$P_{1,2} = a_{1,2} T. \quad (4.24)$$

Це співвідношення може бути отримане у разі застосування до термоелектричних явищ першого і другого законів термодинаміки.

У напівпровідникових термоелектричних пристроях використовують напівпровідникові термоелементи, кожен з яких складається з двох гілок з різним типом електропровідності. Гілку термоелемента, матеріал якого має електропровідність р-типу, називають додатною гілкою, гілку термоелемента з електропровідністю n-типу – від'ємною. Додатну та від'ємну гілки напівпровідникового термоелемента з'єднують між собою послідовно контактною пластиною (рис. 4.7). Зону електричного з'єднання гілок напівпровідникового термоелемента називають спаєм. Під час роботи термоелемента його спаї мають різну температуру: один з них є теплопоглинальним, інший – тепловидільним.

Напівпровідниковий термоелектричний пристрій звичайно має велику кількість послідовно з'єднаних між собою термоелементів у єдиному конструктивному виконанні – напівпровідникову термобатарею. Напівпровідниковий термоелектричний пристрій, що містить одну чи кілька електрично з'єднаних напівпровідникових термобатарей у єдиному конструктивному виконанні з системою теплообміну, називають *напівпровідниковим термоелектричним блоком*.

Явище охолодження контакту у разі проходження струму особливо важливе, оскільки дозволяє створювати термоелектричні холодильні пристрої – домашні холодильники, пристрої для охолодження бортової радіоелектронної апаратури, мікрохолодильники для біологічних цілей, різноманітні термоелектричні термостати, мікроскопні столики з охолодженням і нагріванням тощо. Зручним є те, що, змінюючи напрям струму, можна перетворювати холодильник на нагрівач і, навпаки, нагрівач в холодильник.

Лекція 5

ГАЛЬВАНОМАГНІТНІ ЕФЕКТИ В НАПІВПРОВІДНИКАХ

План

- 5.1. Ефект Холла
- 5.2. Магніторезистивний ефект
- 5.3. Магнітоконцентраційний ефект
- 5.4. Ефект Фарадея
- 5.5. Ефект Фогта
- 5.6. Термомагнітні ефекти в напівпровідниках

Література

- 5.1. Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.
- 5.2. Мікроелектронні сенсори фізичних величин: Науково-навчальне видання. В 3 томах. Том 1. / Вуйцік В., Готра З.Ю., Григор'єв В.В., Каліта В., Мельник О.М., Потенцкі Є.; за ред. З.Ю. Готри. – Львів: Ліга-Прес, 2002. – 475 с.
- 5.3. Микроэлектроника. В 9 кн. / Под ред. Л. А. Коледова. Кн. 1. Физические основы функционирования изделий микроэлектроники / О.В. Митрофанов, Б.М. Симонов, Л.А. Коледов.: Высшая школа, 1987. – 168 с.
- 5.4. Степаненко И. П. Основы микроэлектроники. – М.: Сов. радио, 1980. – 286 с.

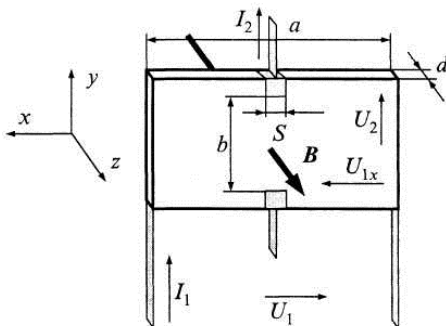


Рис. 5.1. Нормальний ефект Холла

5.1. Ефект Холла

Для пояснення фізичного процесу ефекту Холла розглянемо явища, які відбуваються в плоскій напівпровідниковій пластині, що володіє електронною провідністю і має довжину a , ширину b і товщину d (рис. 5.1). На пластину в перпендикулярному напрямку діє магнітне поле, напруженість якого визначається вектором магнітної індукції B , а у поздовжньому напрямку протікає струм I_1 , зумовлений напруженістю електричного поля E_1 зовнішнього джерела. Магнітне поле, діючи на електрони, які рухаються з середньою швидкістю V , відхиляє їх у поперечному напрямку вздовж осі y . У результаті цього на одній з поперечних граней будуть накопичуватися електрони, а на протилежній – нескомпенсований позитивний заряд. Накопичення зарядів на протилежних гранях пластинки призводить до появи поперечного електричного поля з напруженістю E_2 , яке називається полем Холла. Накопичення зарядів на протилежних гранях пластини буде відбуватися доти, доки сила Лоренца F_1 що діє на електрони за рахунок магнітного поля B , не зрівноважить силу F_2 , яка діє на електрони за рахунок електричного поля E_2 . Після цього встановиться стаціонарний процес, при якому електрони рухаються паралельно до граней пластини, як і за відсутності магнітного поля. Зауважимо, що сказане справедливе для електронів, з середньою швидкістю руху. Умова рівноваги сил $F_1 = F_2$ запишеться в такому вигляді:

$$e[\mathbf{u} \times \mathbf{B}] = e\mathbf{e}_2, \quad (5.1)$$

де e – заряд електрона, рівний $1,602 \cdot 10^{-19}$ Кл.

Якщо врахувати, що вектори \mathbf{v} і \mathbf{B} взаємно перпендикулярні, то з (5.1) одержимо:

$$\mathbf{e}_2 = \mathbf{u} \cdot \mathbf{B}. \quad (5.2)$$

Підставляючи в (5.2) значення:

$$\mathbf{u} = m\mathbf{e}_1, \quad (5.3)$$

де μ – холлівська рухливість носіїв зарядів, маємо:

$$\mathbf{e}_2 = m\mathbf{e}_1 B, \quad (5.4)$$

Виразимо ε_1 через густину струму j_1 і електропровідність σ .

$$\mathbf{e}_1 = \frac{j_1}{\sigma} = \frac{j_1}{enm}, \quad (5.5)$$

де n – концентрація носіїв електронів.

Підставивши (5.5) у (5.4), отримаємо:

$$\mathbf{e}_2 = \frac{1}{en} j_1 B = R_H j_1 B. \quad (5.6)$$

Коефіцієнт R_H називається *коефіцієнтом Холла* (поширений також термін *стала Холла*, не зовсім правильний, оскільки R_H навіть для цього матеріалу не є сталою величиною, а залежить від ряду чинників, зокрема від індукції і температури).

Вираз (5.6) отримано для напівпровідникового матеріалу з електронною провідністю. Для напівпровідникового матеріалу зі змішаною провідністю, тобто за наявності у ньому дірок і електронів, ті й інші відхиляються магнітним полем до однієї і тієї ж грані зразка, тому що під дією зовнішнього електричного поля ξ_1 ці заряди рухаються в протилежних напрямках. Поперечні поля, створені накопиченням дірок і електронів, будуть взаємно послабляти один одного, і результуюче холлівське поле виявиться меншим, ніж у напівпровідниках з одним типом носіїв. Його значення в цьому випадку визначатимуться виразом:

$$\mathbf{e}_2 \sim \frac{\alpha (b^2 n - p)}{e (bn - p)^2} j_1 B. \quad (5.7)$$

де $b = \frac{m_n}{m_p}$ – відношення рухливостей електронів і дірок; n

і p – концентрація електронів і дірок; α – стала, яка залежить від механізму розсіювання носіїв заряду.

Отже, ефект Холла більш різко виражений у напівпровідниках, де переважають носії зарядів одного знака. Оскільки рухливість електронів μ_n звичайно більша, ніж рухливість дірок μ_p , то ефект Холла в напівпровідниках з електронною провідністю відчутніший, ніж у напівпровідниках із дірковою провідністю. У металах і діелектриках рухливість носіїв не перевищує десятків $\text{см}^2/(\text{Вс})$, у той час як в антимоніді індію рухливість сягає $80000 \text{ см}^2/(\text{Вс})$. Тому в металах

ефект Холла виявляється значно слабше, ніж у напівпровідниках.

Якщо врахувати, що поздовжній струм через пластину $I_1 = j_1 b a$, а поперечна напруга $U_{20} = \varepsilon_2 b$, то (5.6) можна записати в такому вигляді:

$$U_{20} = \frac{R_H}{d} I_1 B. \quad (5.8)$$

Відношення $R_H/d = \gamma_{20}$ називається питомою чутливістю ненавантаженого ідеального перетворювача Холла. Під ідеальним розуміється прямокутний перетворювач Холла, який має нескінченно велику довжину в поздовжньому напрямку і точкові холлівські електроди. Реальний перетворювач Холла, який має певні геометричні розміри, також характеризується питомою чутливістю до режиму неробочого ходу:

$$g'_{20} = g_{20} j_1. \quad (5.9)$$

Коефіцієнт φ_1 залежить від співвідношення геометричних розмірів і значення магнітної індукції. Отже, напругу Холла реального ненавантаженого перетворювача Холла можна подати рівнянням:

$$U_{20} = g'_{20} I_1 B. \quad (5.10)$$

З урахуванням того, що $\sigma = en\mu$ коефіцієнт Холла можна виразити через рухливість і електропровідність:

$$R_H = \frac{1}{ne} = \frac{m}{s}. \quad (5.11)$$

Отриманий вираз для коефіцієнта Холла в загальному випадку не зовсім точний, оскільки при його виведенні не враховувався розподіл електронів за швидкостями, механізм розсіювання носіїв, різниця між повною швидкістю електронів і дрейфовою швидкістю, яку електрон набуває під дією електричного поля. Для уточнення значення коефіцієнта Холла вираз (5.11) доповнюється множителем r_H , що називається *холл-фактором*:

$$R_H = r_H \frac{1}{ne} = r_H \frac{m}{s}. \quad (5.12)$$

У металах, сильновироджених напівпровідниках і в сильних магнітних полях, для яких справедлива умова $m^2 B^2 \gg 1$, холл-фактор дорівнює 1. У багатьох випадках перетворювачі Холла, які використовуються для технічних цілей, мають $r_H = 1$.

Розглядаючи результуюче електричне поле ε у напівпровідниковій пластині як геометричну суму полів ε_1 і ε_2 неважко зрозуміти, що воно повернуто щодо напрямку струму I_1 на кут θ , який визначається виразом:

$$\operatorname{tg} \theta = \frac{\varepsilon_2}{\varepsilon_1}$$

З урахуванням (5.4) одержимо:

$$\operatorname{tg} \theta = m^2 B. \quad (5.13)$$

Якщо вихід перетворювача Холла замкнутий на

зовнішній активний опір R , то напруга на холлівських електродах визначається виразом:

$$U_2 = U_{20} - I_2 R_2,$$

де I_2 – струм навантаження перетворювача; R_2 – опір перетворювача Холла між холлівськими електродами. Якщо врахувати, що

$$I_2 = \frac{U_{20}}{R_2 + R},$$

можна знайти напругу Холла для навантаженого перетворювача:

$$U_2 = \frac{U_{20}}{1 + \frac{R_2}{R}} = \frac{g_{20}}{1 + \frac{R_2}{R}} IB. \quad (5.14)$$

Коефіцієнт $g_2 = \frac{g_{20}}{1 + \frac{R_2}{R}}$ – чутливість навантаженого

перетворювача Холла – є функцією властивостей напівпровідникового матеріалу, магнітної індукції, геометрії перетворювача і значення навантаження на його виході R .

У разі проходження через холлівські електроди перетворювача струму навантаження I_2 виникає так звана *вторинна напруга Холла* U_{IX} , яка діє між керувальними електродами (рис. 5.1). Іншими словами, функції обох пар електродів взаємно змінюються – через холлівські електроди проходить струм I_2 , який, взаємодіючи з магнітним полем B , зумовлює вторинну напругу Холла U_{IX} , спрямовану проти прикладеної до керувальних електродів зовнішньої напруги U_1 . Вторинна напруга Холла аналогічна до первинної і визначається з виразу:

$$U_{IX} = g_1 I_2 B,$$

де g_1 – чутливість оберненого (для визначення вторинної напруги Холла) навантаженого перетворювача Холла. Відзначимо, що чутливості ненавантаженого перетворювача Холла для первинного і вторинного ефектів Холла однакові, тобто $g'_{10} = g'_{20}$. У той час як для навантаженого перетворювача Холла $g_{10} \neq g_{20}$.

5.2. Магніторезистивний ефект

Зміна електричного опору напівпровідника, в якому протікає електричний струм, під дією магнітного поля називається магніто-резистивним ефектом. Цей ефект, відомий також під назвою ефекту Гаусса, вивчався ще у 1858 р. Томсоном. Збільшення електричного опору під дією магнітного поля відбувається як у випадку, коли вектор магнітної індукції спрямований перпендикулярно до напрямку протікання струму в напівпровідниковій пластині (поперечний ефект магнітоопору), так і у випадку, коли вектор магнітної індукції паралельний до напрямку струму (поздовжній ефект магнітоопору). Проте зміна опору при поздовжньому ефекті магнітоопору є

незначною. Збільшення опору при поперечному ефекті магнітоопору пояснюється викривленням траєкторії руху частини електронів під дією магнітного поля.

Як було показано раніше, у довгій напівпровідниковій пластині, яка перебуває у магнітному полі, на електрони, які рухаються, діє сила Лоренца, що врівноважується дією електричного поля Холла. Тому електрони, що мають середню швидкість руху $u = m\mathbf{e}_1$, мають траєкторію, паралельно до поздовжніх граней пластини. Проте не всі електрони мають середню швидкість, яка визначається значенням зовнішнього електричного поля ξ_2 оскільки хаотичний тепловий рух електронів викликає значні відхилення від середнього значення швидкості. Поле Холла ε_2 вже не може цілком скомпенсувати силу Лоренца, що діє на електрони, які рухаються зі швидкістю, відмінною від середньої. Траєкторії останніх відхиляються як в один, так і в інший бік від напрямку, паралельного до поздовжніх граней пластини. У результаті для переносу заряду вздовж зовнішнього поля ε_1 такі електрони в поперечному магнітному полі повинні пройти більший шлях, що рівнозначно зростанню електричного опору.

Описане збільшення електричного опору напівпровідникового матеріалу в магнітному полі називають фізичною зміною опору. Ця зміна описується наближеною формулою:

$$\frac{r_B - r_0}{r_B} \alpha B^n$$

де ρ_B – питомий опір у магнітному полі; α – коефіцієнт, який залежить від матеріалу напівпровідника; n – показник, який залежить від значення магнітної індукції.

Для слабких магнітних полів ($m^2 B^2 = 1$) $n = 2$; в області сильних магнітних полів ($m^2 B^2 \gg 1$) показник степеня n наближається до 1.

Ефект магнітоопору залежить не тільки від матеріалу напівпровідникової пластини, але і від її форми, оскільки поле Холла визначається геометрією зразка. Найбільш наочним прикладом впливу геометрії зразка на ефект магнітоопору є диск Корбіно. Він являє собою напівпровідникову пластину у формі диска, до центра якого струм підводиться, а відводиться за допомогою електрода, що оперізує диск по зовнішній грані. Якщо до електродів такого диска прикласти напругу, то траєкторії електронів будуть мати вигляд радіальних променів, які розходяться від центра диска, а екіпотенційні лінії являтимуть собою концентричні кола (рис. 5.2, а).

При появі магнітного поля електричне поле Холла в диску Корбіно не виникає. Завдяки цьому під дією сили Лоренца лінії струму матимуть вигляд не найкоротшого шляху від електрода до електрода (як за відсутності магнітного поля), а кривих, дотичні до яких в кожній

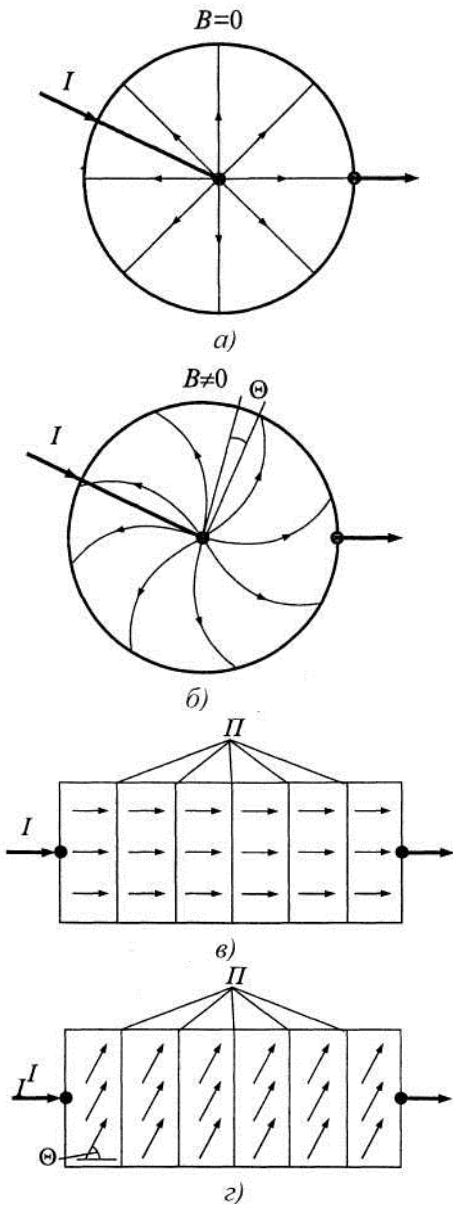


Рис. 5.2. Спотворення (подовження) ліній струму в диску Корбіно і напівпровідниковій пластині з поперечними провідними смугами П.

точці утворюють із напрямком радіальних променів кут Холла θ (рис. 5.2, б).

Еквіпотенційні лінії при цьому як і раніше мають вигляд концентричних кіл. Якщо виготовити диск Корбіно з напівпровідникового матеріалу з великою рухливістю носіїв зарядів, то збільшення його опору при $B = 1$ Тл може досягти 37-кратного значення. Проте диск Корбіно з матеріалу з високою рухливістю навіть при товщині 0,1 мм має малий опір, який не перевищує десятих частин Ом. Навантажувальна спроможність його по струму обмежена, що пояснюється високою густиною струму в зоні точкового електрода в центрі диска. Нарешті струмопідвід до центрального електрода призводить до потовщення диска Корбіно і неможливості розташування його в малих повітряних прошарках. Все це сильно обмежує практичне застосування диска Корбіно. Збільшення ефекту магнітоопору за рахунок геометрії напівпровідникового резистора може бути досягнуте також і в плоскій напівпровідниковій пластині, довжина якої a значно менша від її ширини b . При дії на таку пластину магнітного поля, напрямом якого перпендикулярний до площини пластини, поле Холла виявляється послабленим за рахунок дії шунтуючих струмових електродів. Завдяки цьому сила Лоренца, яка діє на електрони, буде скомпенсованою не цілком, і траєкторії руху електронів викривляться. Геометричний ефект магнітоопору в цьому випадку виявляється тим сильніше, чим менше відношення a/b . Отже, до вже описаного фізичного збільшення опору додається ще геометричне збільшення. Вплив обох чинників призводить до результуючої зміни опору, яка може бути подана виразом:

$$\frac{R_B}{R} = \frac{r_B - k_G}{r_0},$$

де k_G – величина, яка залежить від геометрії пластини і значення магнітної індукції.

Криві, що характеризують коефіцієнти k_G для прямокутних пластин із різноманітним співвідношенням a/b у функції від $\text{tg}\theta = \mu B$, зображено на рис. 5.3.

Виготовивши резистор у вигляді ряду послідовно з'єднаних напівпровідникових пластин із малим співвідношенням a/b , можна одержати магнітозалежні опори з напівпровідникових матеріалів, що мають високу рухливість носіїв струму, у яких абсолютне значення опору становить кілька сотень Ом. Зазначена ідея практично може бути реалізована у такий спосіб. На ізоляційну підкладку напильюється або приклеюється тонка пластинка подовженої прямокутної форми з напівпровідникового матеріалу з високою рухливістю носіїв струму. На поверхню пластини тим чи іншим методом наносять провідні електроди Π , які розділяють пластину на ряд послідовно сполучених ділянок з малим

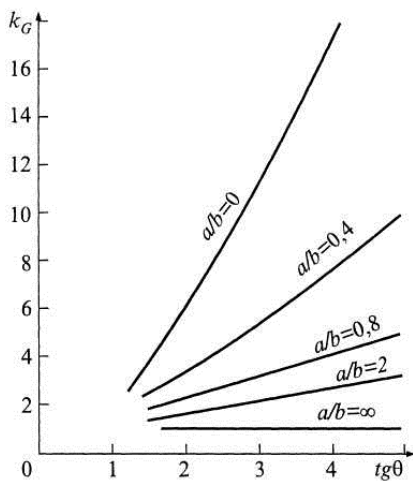


Рис. 5.3. Залежність коефіцієнта k_G від $\text{tg}\theta$ для різних відношень a/b .

відношенням a/b (рис. 5.2, *в*). На кожній з цих ділянок зменшити поле Холла до нуля (як у диску Корбіно) не вдасться, але воно виявляється значно послабленим. Завдяки цьому лінії струму на кожній ділянці виявляються повернутими відносно поздовжніх граней пластини на кут, близький до кута Холла θ (рис. 5.2, *г*). Зміною довжини, ширини і товщини кожної ділянки, а також вибором матеріалу напівпровідника можна домогтися 15-кратного збільшення опору в полі 1 Тл.

Інша можливість реалізації принципу використання послідовно сполучених ділянок з малим відношенням a/b може бути досягнута створенням напівпровідникового з'єднання з особливою структурою.

Основою цієї структури слугує антимонід індію InSb, який має високу рухливість носіїв. До цього основного матеріалу додається 1,8% антимоніду нікелю NiSb. Після сплавлення і наступного остигання антимонід нікелю викристалізовується в антимоніді індію у вигляді голок завтовшки близько 1 мкм і завдовжки близько 50 мкм. Спрямованим вирощуванням (зонним плавленням) вдається домогтися того, що голки розташовуються не довільно, а паралельно одна до одної. Відстань між осями голок лежить у межах від тисячних до десятих частин міліметра. Питома провідність голок з антимоніду нікелю дорівнює приблизно $7 \cdot 10^6 \text{ Ом}^{-1} \text{ м}^{-1}$ а антимоніду індію з власною провідністю $2 \cdot 10^4 \text{ Ом}^{-1} \text{ м}^{-1}$. Отже, голки відіграють роль добре провідних електродів, які розділяють основну масу напівпровідника на окремі зони з малим відношенням довжини до ширини. Це призводить до різкого послаблення поля Холла, а отже, до викривлення траєкторії електронів під дією магнітного поля, тобто до зростання ефекту магнітоопору.

5.3. Магнітоконцентраційний ефект

Розглянемо явища, які відбуваються у подовженому паралелепіпеді з власного напівпровідника, характерною рисою якого є рівність концентрацій електронів і дірок $n_0 = p_0$ (рис. 5.4). В умовах термодинамічної рівноваги у напівпровіднику одночасно з тепловим збудженням електронів тобто генерацією носіїв струму – електронів і дірок) відбувається і протилежний процес – рекомбінація носіїв струму, тобто повернення електронів із зони провідності у валентну зону, в результаті чого електрон «нейтралізує» дірку. Ці два процеси перебігають безупинно, і для кожної температури встановлюється рівновага що характеризується відповідною концентрацією електронів n_0 . З підвищенням температури рівноважна концентрацій носіїв струму напівпровідника зростає за експоненціальним законом. Внаслідок того, що на поверхні напівпровідникового паралелепіпеда є порушення структури ґратки і домішок, швидкість рекомбінації тут вища, ніж у середині його об'єму.

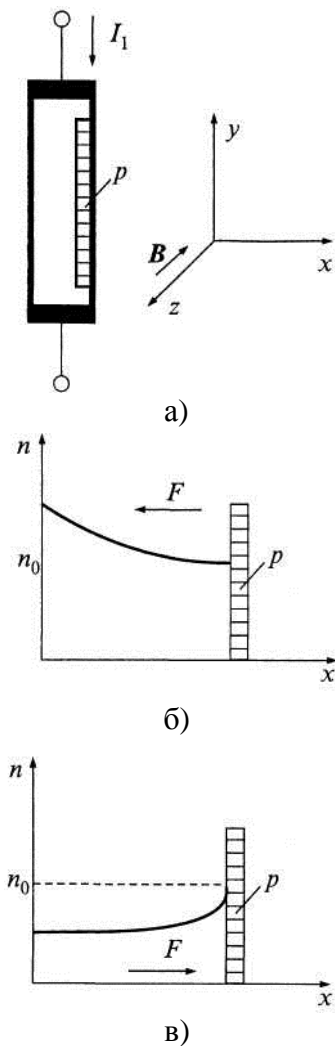


Рис. 5.4. Магнітоконцентраційний ефект: P – область з більшою швидкістю рекомбінації; F – сила Лоренца.

Пропустимо через напівпровідниковий паралелепіпед струм I_1 і помістимо його в магнітне поле, спрямоване перпендикулярно до напрямку протікання струму (рис. 5.4, а). Нехай на одній з граней напівпровідника штучно створена область P з більшою швидкістю рекомбінації. Під дією сили Лоренца F відбувається перерозподіл концентрації носіїв струму по перерізу паралелепіпеда в напрямку дії сили Лоренца. На грані з більшою швидкістю рекомбінації концентрація носіїв струму практично не змінюється, у той час як в області, протилежній до грані, залежно від напрямку сили Лоренца концентрація носіїв збільшується або зменшується. При цьому відбувається загальне збільшення або зменшення концентрації носіїв порівняно з рівноважною, що викликає відповідне зменшення або збільшення електричного опору напівпровідникового зразка.

Описаний ефект зміни опору напівпровідника під дією магнітного поля за рахунок зміни концентрації носіїв струму називається *магнітоконцентраційним* або *гальваноманітореконцентраційним ефектом*. Важливим для технічного використання цього ефекту є те, що він дозволяє створити напівпровідникові перетворювачі, чутливі не тільки до значення вектора індукції магнітного поля, але і до його напрямку. Зауважимо, що цією властивістю не володіють перетворювачі, які використовують магніторезистивний ефект. Водночас у зв'язку з необхідністю використання напівпровідникових матеріалів із власною провідністю, які мають велику температурну залежність, магнітоконцентраційний ефект також виявляється сильно залежним від температури. Крім того, сильна залежність ефекту від стану поверхні напівпровідникового зразка зумовлює нестабільність роботи і великий розкид від зразка до зразка параметрів перетворювачів магнітного поля, робота яких базується на магнітоконцентраційному ефекті.

5.4. Ефект Фарадея

Магнітне обертання площини поляризації лінійно поляризованого випромінювання було виявлено Фарадеєм в 1845 році. Цей експеримент вперше виявив зв'язок між оптичними та магнітними явищами. Ефект Фарадея полягає в обертанні площини поляризації лінійно поляризованого випромінювання при його поширенні через речовину, яка міститься в магнітному полі, паралельному напрямку поширення випромінювання (рис. 5.5).

Знак ефекту Фарадея визначається тільки напрямком магнітного поля, тобто обертання площини відбувається в один і той бік, незалежно від того поширюється випромінювання вздовж магнітного поля чи проти. На рис. 5.5 поворот площини поляризації в ефекті Фарадея відповідає позитивному куту обертання θ . Величина кута

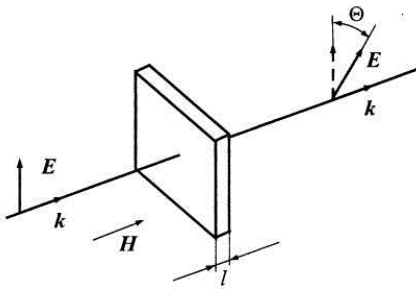


Рис. 5.5. Обертання площини поляризації лінійно поляризованого світла на кут θ в ефекті Фарадея в шарі напівпровідника завтовшки l (E, k – вектор напруженості електричного поля і хвильовий вектор електромагнітної хвилі, H – вектор напруженості постійного магнітного поля)

θ пропорційна величині магнітного поля.

Незалежність знака магнітного обертання площини поляризації від напрямку поширення впливає з симетрії магнітного поля. Застосування операції симетрії – відбивання в площині, яка перпендикулярна до магнітного поля, – не призводить до зміни знака θ . У цьому і полягає відмінність між магнітним обертанням і природним обертанням, яке спостерігається в оптично активних середовищах.

Феноменологічний розгляд магнітного обертання площини поляризації світла базується на уявленні про різницю характеристик середовища для двох неоднаковим чином поляризованих компонент випромінювання.

Нехай лінійно поляризована електромагнітна хвиля нормально падає на шар речовини. Лінійне коливання вектора ϵ хвилі можна подати як суперпозицію циркулярно право- і лівополяризованих компонент, кутові швидкості обертання яких однакові. При появі магнітного поля для цих компонент виникають неоднакові умови їх поширення. Частоти обертання електричних векторів двох компонент відрізняються від частоти хвилі ω на ларморівську частоту $\pm\omega_L = \frac{eH}{2mc}$. Внаслідок цього

показник заломлення, який є функцією частоти, буде відмінним для право- і лівополяризованих компонент, і ці компоненти будуть поширюватися з різною швидкістю.

Експериментально різна швидкість поширення циркулярно право- і лівополяризованих пучків світла в шарі речовини, яка міститься в магнітному полі, була показана Рігі і Бекерелем за зміщенням інтерференційних смуг. Зміщення виникало при появі магнітного поля, що свідчить про різну швидкість поширення через кристал, який міститься в магнітному полі, двох поляризованих по колу в різні боки пучків світла.

Брейсом було здійснено підтвердження розкладу лінійно поляризованого випромінювання, яке попадає на магнітноактивний кристал, на дві циркулярно поляризовані в різні боки компоненти. Як диспергуючий елемент використовувалася призма спеціальної конструкції. Оскільки право- і лівополяризовані компоненти мають різні швидкості поширення, то на виході з призми, яка міститься в магнітному полі, вони повинні відхилитися на різні кути. Різниця кутів відхилення цих променів є надзвичайно малою, бо показники заломлення для двох циркулярно поляризованих в різні боки компонент в магнітних полях напруженістю до 1 Тл відрізняються між собою на величину близько 10^{-5} . Для збільшення кута розведення компонент Брейс використовував спеціальну конструкцію двох призм, які розділені півхвильовою пластиною. Після двадцятиразового проходження через таку конструкцію

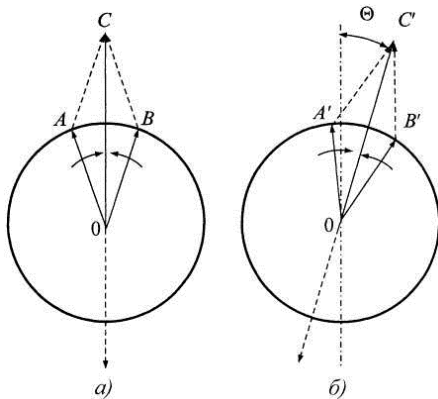


Рис. 5.6. Обертання площини поляризації лінійнополяризованого світла: а – до проникнення в магнітооптичне середовище; б – на виході із середовища.

початково лінійно поляризоване випромінювання вдалося геометрично розділити на право- і лівополяризовані компоненти.

Розглянемо процес поширення вздовж напрямку дії магнітного поля лінійно поляризованого випромінювання через шар речовини, який повертає площину поляризації на деякий кут θ (магнітоактивне середовище).

До проникнення в магнітооптичне середовище (рис. 5.6, а) вектори OA і OB , які зображають електричні вектори двох циркулярно поляризованих в різні боки компонент випромінювання, обертаються в різні боки з однаковою кутовою швидкістю. Результуючим вектором двох циркулярно поляризованих компонент є вектор OC , який зберігає свій напрямок в просторі і є лінійно поляризованим випромінюванням. При поширенні випромінювання через магнітооптичний активний кристал показники заломлення для ліво- і правополяризованого випромінювання є різними, тому компоненти поширюються з різною швидкістю. Внаслідок цього, при виході з активного середовища ці компоненти будуть володіти певною різницею фаз. Схема, що ілюструє позицію циркулярно поляризованих компонент на виході з середовища, подана на рис.5.6, б. Вектори OA' і OB' також обертаються в різні боки з однаковою кутовою швидкістю, однак результуючий вектор OC' коливається в площині, яка становить кут θ з його початковим положенням до активного середовища. У разі збігання напрямку OA' з початковим положенням результуючого вектора OC різниця фаз між компонентами OA' і OB' становить 2θ , тобто кут повороту площини поляризації в ефекті Фарадея дорівнює половині різниці фаз між двома циркулярно поляризованими в різні боки компонентами випромінювання.

Кут повороту площини поляризації можна виразити через показники заломлення $n_- = n(w + w_L)$ і $n_+ = n(w - w_L)$ для двох компонент, поляризованих по колу. Кількість обертів вектора індукції для правополяризованої по колу компоненти на одиниці довжини становитиме $1/I_+$, а для лівополяризованої $1/I_-$ де

$$I_{+,-} = \frac{2p}{w} \cdot \frac{c}{n_{+,-}},$$

тут ω – частота випромінювання; c – швидкість світла у вакуумі.

Різниця кількості обертів цих компонент і визначає різницю фаз на довжині l . Отже, для кута обертання площини поляризації отримаємо вираз:

$$q = \frac{j}{2} l = \frac{wl}{2c} (n_- - n_+).$$

Площина поляризації повертається в бік обертання циркулярно поляризованої компоненти з більшою швидкістю (у напівпровідниках $n_- - n_+ \sim 10^{-4} - 10^{-6}$).

У спектральному діапазоні $l \leq 50$ мкм, в якому переважно досліджують інфрачервоне фарадеївське обертання $w_L = w$, і для об'ємної різниці фаз:

$$n_- - n_+ \cong 2w_L \frac{dn}{dw}$$

Звідси випливає вираз:

$$q = \frac{w l}{c} w_L \frac{dn}{dw} = \frac{ew l H}{2mc^2} \frac{dn}{dw}.$$

Поворот площини поляризації в магнітному полі можна виразити через так звану сталу Верде V :

$$q = V H l.$$

Значення сталої Верде для різних напівпровідникових матеріалів коливається в широких межах ($V \sim 10^{-7} \dots 10^{-11}$ рад/Тлсм) залежно від величини ефективної маси і концентрації носіїв, g -факторів, довжини хвилі випромінювання, температури та ін. У напівпровідниках при частотах випромінювання, які далекі від частоти циклотронного резонансу, в магнітних полях напруженістю $H \leq 10$ Тл величина V не залежить від величини H .

5.5. Ефект Фогта

Подвійне променезаломлення в магнітному полі було передбачене 1899 р. і пізніше спостерігалось Фогтом і Віхером для D-ліній парів натрію. Цей ефект також називають ефектом Коттона-Мутона, оскільки подібний ефект в колоїдних розчинах заліза був виявлений Майораном 1902 р. і детально досліджений на більшій кількості розчинів Коттоном та Мутоном.

Ефект Фогта полягає в тому, що лінійно поляризоване випромінювання, яке напрямлене перпендикулярно до вектора індукції магнітного поля, після проходження шару ізотропної речовини, яка міститься в магнітному полі, стає еліптично поляризованим. Це зумовлене оптичною анізотропією, яка виникає в магнітному полі – у речовині з'являється виділений напрямок, який збігається з напрямком вектора індукції магнітного поля.

Явище магнітного двоприменезаломлення подібно до звичайного двоприменезаломлення в оптично анізотропних (одновісних) кристалах має таке пояснення. Нехай електричний вектор ϵ лінійно поляризованого випромінювання частоти ω , яке падає на ізотропний кристал, становить з напрямком зовнішнього магнітного поля деякий кут (рис. 5.7). Лінійне коливання вектора ϵ можна розкласти на дві компоненти, які мають паралельну ϵ_{\parallel} і перпендикулярну ϵ_{\perp} поляризації. За відсутності магнітного поля швидкості розповсюдження компонент ϵ_{\parallel} і ϵ_{\perp} однакові внаслідок ізотропності кристала. При появі магнітного поля умови поширення

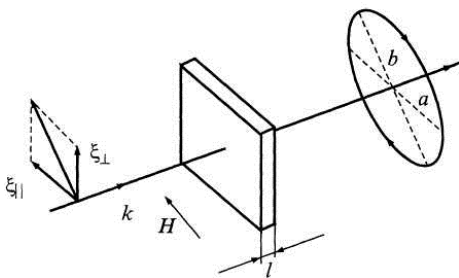


Рис. 5.7. Двоприменезаломлення в магнітному полі (ефект Фогта), яке виникає під час проходження лінійно поляризованого світла через шар ізотропного матеріалу, при поширенні випромінювання в напрямку, перпендикулярному напрямку магнітного поля H (стрілка на контурі еліпса показує напрямок обертання результуючого електричного вектора для $0 < \delta < \pi/2$, $n_{\perp} > n_{\parallel}$ при рівних амплітудах компонент x_{\parallel} і x_{\perp}).

цих двох компонент стають неоднаковими, і компоненти поширюються з різними швидкостями $u_{\parallel} = \frac{c}{n_{\parallel}}$ і $u_{\perp} = \frac{c}{n_{\perp}}$, де n_{\parallel} і n_{\perp} – показники заломлення для хвиль, які поляризовані паралельно і перпендикулярно магнітному полю.

Отже, після проходження шару речовини між компонентами e_{\parallel} і e_{\perp} з'являється різниця фаз, і результуюче випромінювання буде еліптично поляризованим, оскільки додавання двох взаємно перпендикулярних синусоїдальних доливаних однакової частоти дає еліптичну траєкторію. Фазовий зсув δ (величина ефекту Фогта), який виникає внаслідок різної швидкості поширення компонент на довжині зразка l , буде визначатися виразом:

$$d = 2pl \left(\frac{1}{I_{\perp}} - \frac{1}{I_{\parallel}} \right) = \frac{wl}{c} (n_{\perp} - n_{\parallel}) \text{ [рад]}.$$

Параметри еліпса, напрямок його головних осей визначаються різницею фаз δ , тобто різницею $n_{\perp} - n_{\parallel}$ і співвідношенням між амплітудами e_{\parallel} і e_{\perp} . При $\delta = 0$ випромінювання на виході залишається лінійно поляризованим. При $\delta = \pi/2$ лінійно поляризоване випромінювання на виході стає поляризованим по колу. При $\pi/2 < \delta < \pi$ більша вісь еліпса орієнтована під кутом 90° до площини поляризації лінійно поляризованого випромінювання.

5.6. Термомагнітні ефекти в напівпровідниках

Термічні і термомагнітні ефекти виникають при тих самих умовах, що і гальваномагнітні, і можуть спотворити вияв останніх. Відомі і досліджувалися три термомагнітні ефекти: Еттінгсгаузена, Нернста-Еттінгсгаузена і Рігі-Ледюка. Схематично вони наведені на рис. 5.8.

Ефект Еттінгсгаузена являє собою виникнення температурного градієнта dT/dy , який напрямлений перпендикулярно густині струму j_x і магнітній індукції B_z :

$$\frac{dT}{dy} = P j_x B_z. \quad (5.15)$$

Величина P називається коефіцієнтом Еттінгсгаузена. Цей ефект проявляється через виникнення термоелектричної напруги на холлівських контактах в магнітному полі.

Явище Нернста-Еттінгсгаузена являє собою ефект, зворотний до явища Еттінгсгаузена. Якщо в зразку створювати температурний градієнт в напрямку x , то в поперечному напрямку виникає напруга:

$$\frac{dU}{dy} = -Q \frac{dT}{dy} B_z, \quad (5.16)$$

де Q – коефіцієнт Нернста-Еттінгсгаузена.

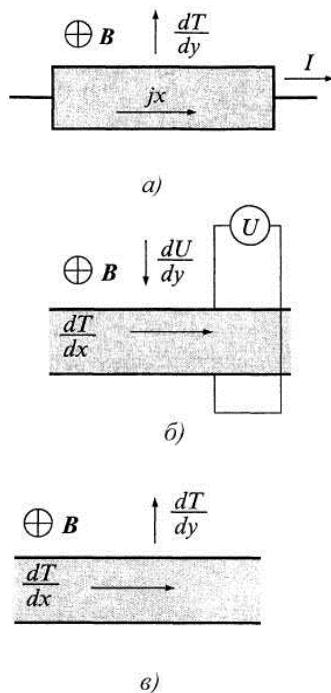


Рис. 5.8. Схема прояву ефектів Еттінгсгаузена (а), Нернста-Еттінгсгаузена (б), Рігі-Ледюка (в)

У випадку ефекту Рігі-Ледюка виникає поперечний температурний градієнт, перпендикулярний напрямку теплового потоку (напрямок x) і магнітного поля B_z :

$$\frac{dT}{dy} = S \frac{dT}{dx} B_z, \quad (5.17)$$

де S – коефіцієнт Рігі-Ледюка.

Для InSb з власною провідністю значення трьох термомагнітних коефіцієнтів P , Q і S при кімнатній температурі дорівнюють:

$$P = -5,3 \cdot 10^{-4} \text{ м}^3 \text{К}/(\text{Вт} \cdot \text{с});$$

$$Q = -3 \cdot 10^{-5} \text{ м}^2/(\text{К} \cdot \text{с});$$

$$|S| < 1 \cdot 10^{-2} \text{ м}^2/(\text{В} \cdot \text{с}).$$

У полі $B = 0,1$ Тл і при густині струму $j_x = 1 \cdot 10^3$ А/м² ефект Етінсгаузена викликає в напрямку y перепад температури близько 5,3 К/м. Якщо при такому самому магнітному полі створити градієнт температури 100 К/м в напрямку x , то в напрямку y виникне електричне поле з напруженістю близько 300 мкВ/м і градієнт температури менше 0,1 К/м.

Лекція 6

ПРОЦЕСИ ДИФУЗІЇ

План

- 6.1. Дифузійні рівняння
- 6.2. Рівняння Айнштейна

Література

- 6.1. Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.
- 6.2. Мікроелектронні сенсори фізичних величин: Науково-навчальне видання. В 3 томах. Том 1. / Вуйцик В., Готра З.Ю., Григор'єв В.В., Каліта В., Мельник О.М., Потенці Є.; за ред. З.Ю. Готри. – Львів: Ліга-Прес, 2002. – 475 с.
- 6.3. Мікроелектроніка. В 9 кн. / Под ред. Л. А. Коледова. Кн. 1. Физические основы функционирования изделий микроэлектроники / О.В. Митрофанов, Б.М. Симонов, Л.А. Коледов.: Высшая школа, 1987. – 168 с.
- 6.4. Степаненко И. П. Основы микроэлектроники. – М.: Сов. радио, 1980. – 286 с.

6.1. Дифузійні рівняння

У загальному випадку струм провідності складається з геометричної суми дрейфового і дифузійного струмів.

Дрейфові складові густини струму провідності визначаються згідно із законом Ома:

$$j_{ndp} = S_n e = enm_n \mathbf{E} = -enm_n \text{grad} \varphi,$$

$$j_{pdp} = S_p e = pm_p \mathbf{E} = -pm_p \text{grad} \varphi,$$

де φ – електростатичний потенціал; n і p – концентрації електронів і дірок, які в загальному випадку не є рівноважними.

Дифузійна складова залежить від градієнта концентрації вільних носіїв заряду. На рис. 6.1 показано два випадки на прикладі струму дірок. В першому з них $p_1 > p_2$ і дифузія, а отже, і відповідний дифузійний струм, напрямлений так само, як і дрейфовий. У другому випадку струми протилежні.

Дифузійну густину потоку можна записати у вигляді:

$$\Phi_n = -D_n \text{grad} n;$$

$$\Phi_p = -D_p \text{grad} p,$$

де D_n і D_p – коефіцієнти дифузії електронів і дірок, відповідно. Знак "мінус" означає, що перенесення йде в бік зменшення концентрації.

Дифузійним густинам потоків електронів і дірок відповідають дифузійні густини струмів у вигляді:

$$j_{ndиф} = -eD_n \text{grad} n;$$

$$j_{pдиф} = -eD_p \text{grad} p.$$

Тут врахований знак заряду.

Дифузійний струм створює просторовий розділ

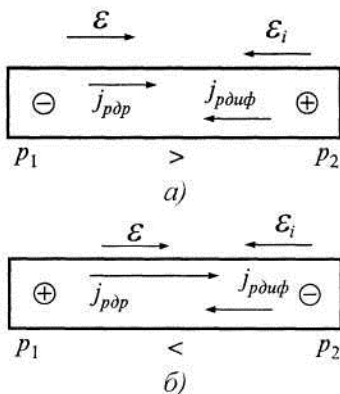


Рис. 6.1. Дифузійний і дрейфовий струми

зарядів, тобто внаслідок дифузії виникають об'ємні заряди, і концентрації носіїв заряду будуть відрізнятися від концентрацій n_0 і p_0 . Зокрема, на рис. 6.1, *a* ліва частина зразка буде заряджатися негативно, а права – позитивно; на рис. 6.1, *б* спостерігається зворотна полярність. Відповідно виникають внутрішні електричні поля ϵ_i , які разом із зовнішнім полем ϵ зумовлюють дрейфовий струм. Отже,

$$\mathbf{e}_n = \mathbf{e} + \mathbf{e}_i$$

Якщо зовнішнє поле дорівнює нулю і напівпровідник ізольований від інших тіл, то дифузійний струм урівноважується дрейфовим струмом, створеним полем ϵ_i і сумарний струм при термодинамічній рівновазі буде дорівнювати нулю. При цьому залишається незмінним розподіл температури в зразку зі зміною часу. Подібний випадок вже зустрічався при розгляді термо-ЕРС.

Отже, в загальному вигляді дифузійні рівняння записують так:

$$\begin{aligned} j_n &= enm_n \mathbf{e} + eD_n \operatorname{grad} n = -enm_n \operatorname{grad} j + eD_n \operatorname{grad} n; \\ j_p &= epm_p \mathbf{e} + eD_p \operatorname{grad} p = -epm_p \operatorname{grad} j + eD_p \operatorname{grad} p. \end{aligned} \quad (6.1)$$

Якщо мати на увазі проекції струмів на який-небудь напрям, то градієнти можна замінити відповідними частковими похідними.

У більшості випадків розглядають струми в якому-небудь одному напрямі, оскільки найчастіше градієнти відмінні від нуля для одного напрямку. Наприклад, для напрямку J_x :

$$\begin{aligned} j_{nx} &= enm_n \mathbf{e}_x + eD_n \frac{\partial n}{\partial x} = -enm_n \frac{\partial j}{\partial x} + eD_n \frac{\partial n}{\partial x}; \\ j_{px} &= epm_p \mathbf{e}_x + eD_p \frac{\partial p}{\partial x} = -epm_p \frac{\partial j}{\partial x} + eD_p \frac{\partial p}{\partial x}. \end{aligned}$$

Часто індекси x опускають, вважаючи їх неявними. При цьому мається на увазі одновимірний випадок, коли:

$$\begin{aligned} j_{ny} = j_{nz} = j_{pz} = i_{pz} &= 0; \quad \frac{\partial j}{\partial y} = \frac{\partial j}{\partial z} = 0; \\ \frac{\partial n}{\partial y} = \frac{\partial n}{\partial z} = \frac{\partial p}{\partial y} = \frac{\partial p}{\partial z} &= 0. \end{aligned}$$

Тоді дифузійні рівняння записують у вигляді:

$$\begin{aligned} j_n &= enm_n \mathbf{e} + eD_n \frac{\partial n}{\partial x}; \\ j_p &= epm_p \mathbf{e} + eD_p \frac{\partial p}{\partial x}. \end{aligned}$$

Повна густина струму:

$$j = j_n + j_p = e(nm_n + pm_p) \mathbf{e} + e \left(D_n \frac{\partial n}{\partial x} - D_p \frac{\partial p}{\partial x} \right). \quad (6.2)$$

Напрями струмів звичайно визначають залежно від напрямів градієнтів потенціалу ϕ і концентрацій. Відповідно, знаки струмів враховують при записі дифузійного рівняння для заданого випадку.

Дифузійні рівняння є найважливішими рівняннями фізики напівпровідників і діелектриків та широко використовуються при різних розрахунках.

6.2. Рівняння Айнштайна

Співвідношення (рівняння) Айнштайна справедливе для вільних носіїв заряду будь-якого типу: електронів, дірок, іонів, іонних вакансій. Воно справедливе як для рівноважних, так і для нерівноважних носіїв заряду в напівпровідниках і діелектриках.

Рівняння Айнштайна зв'язує коефіцієнт дифузії носіїв заряду з їх рухливістю в умовах термодинамічної рівноваги. Воно застосовується і для нерівноважних носіїв заряду, оскільки останні за вельми малий час приходять у теплову рівновагу з ґраткою. Відповідно розподіл за енергіями нерівноважних носіїв заряду в невідроджених напівпровідниках не відрізняється від розподілу рівноважних носіїв заряду.

Виведемо рівняння Айнштайна для напівпровідника p -типу з поперевжнім градієнтом концентрації носіїв заряду (рис. 6.2). Якщо концентрація дірок p_1 (ліворуч) більша, ніж концентрація дірок p_2 (праворуч), то дифузійний струм спрямований зліва направо. Дифузійний струм – перенесення дірок зліва направо триватиме доти, доки не виникне внутрішнє поле ϵ такої величини, при якій зустрічний створений ним дрейфовий струм не зрівноважить дифузійний. Це і буде умовою динамічної рівноваги. Отже,

$$j_p = -epm_p \frac{\partial j}{\partial x} - eD_p \frac{\partial p}{\partial x} = 0, \quad (6.3)$$

де φ – електростатичний потенціал, створений об'ємними зарядами.

Виберемо початок відліку φ на рис. 6.2 так, щоб на лівому кінці він дорівнював нулю.

Дірка, що бере участь в дифузійному струмі, повинна підійматися на потенційний бар'єр $e\varphi$.

Рівняння (6.3) можна переписати так:

$$\frac{\partial p}{p} = -\frac{m_p}{D_p} \partial j,$$

його розв'язок

$$p = p_1 e^{-\frac{m_p}{D_p} j}. \quad (6.4)$$

Згідно із законом Больцмана для заряду, що долає за рахунок теплового руху бар'єр $e\varphi$, можна записати:

$$p = p_1 e^{-\frac{e j}{kT}}. \quad (6.5)$$

Ця концентрація також є термодинамічно рівноважною, оскільки закон Больцмана відповідає рівноважному стану. Але оскільки тут є об'ємні заряди, визначимо цю концентрацію через p замість p_0 .

Прирівнюючи (6.4) і (6.5), отримаємо рівняння Айнштайна для дірок

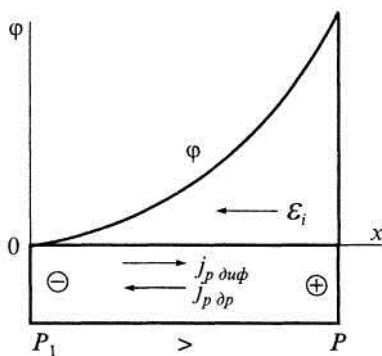


Рис. 6.2. Напівпровідник p -типу з поперевжнім градієнтом концентрації дірок.

$$\frac{m_p}{D_p} = \frac{e}{kT}. \quad (6.6)$$

Аналогічно отримаємо рівняння Айнштейна для електронів:

$$\frac{m_n}{D_n} = \frac{e}{kT}.$$

Такі самі співвідношення справедливі для вільних іонів.

Лекція 7

ПРОЦЕСИ ПОГЛИНАННЯ СВІТЛА

План

- 7.1. Власне поглинання світла
- 7.2. Домішкове поглинання світла
- 7.3. Екситонне поглинання світла
- 7.4. Поглинання світла вільними носіями заряду
- 7.5. Поглинання світла кристалічною ґраткою

Література

- 7.1. Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.
- 7.2. Мікроелектронні сенсори фізичних величин: Науково-навчальне видання. В 3 томах. Том 1. / Вуйцік В., Готра З.Ю., Григор'єв В.В., Каліта В., Мельник О.М., Потенцік Є.; за ред. З.Ю. Готри. – Львів: Ліґа-Прес, 2002. – 475 с.
- 7.3. Мікроелектроніка. В 9 кн. / Под ред. Л. А. Коледова. Кн. 1. Физические основы функционирования изделий микроэлектроники / О.В. Митрофанов, Б.М. Симонов, Л.А. Коледов.: Высшая школа, 1987. – 168 с.
- 7.4. Степаненко И.П. Основы микроэлектроники. – М.: Сов. радио, 1980. – 286 с.

7.1. Власне поглинання світла

Якщо енергія поглинутого фотона затрачується на перекидання електрона з валентної зони в зону провідності, то відповідне поглинання називають *власним* (або *фундаментальним*). Для зонної структури з екстремумами в $k = 0$ маємо:

$$h\nu = E_c(\mathbf{k}) - E_v(\mathbf{k}),$$

де $h\nu$ – енергія фотона, який забезпечує перекид електрона з рівня $E_v(\mathbf{k})$ на рівень $E_c(\mathbf{k})$, як показано на рис. 7.1. Такі переходи електронів, коли \mathbf{k} залишається незмінним, називають *прямими* або *вертикальними*. Насправді оптичні "прямі" переходи не зовсім "вертикальні", оскільки поглинутий фотон передає електрону і свій імпульс. Проте цей імпульс порівняно малий, так що можна вважати перехід вертикальним.

Для параболічних законів дисперсії електронів, $E_c(\mathbf{k})$ і дірок, $E_v(\mathbf{k})$, відрахованих від, відповідно, дна, $E_c(0)$, і стелі, $E_v(0)$, зон, а саме:

$$E_c(\mathbf{k}) = E_c(0) + \frac{\hbar^2 k^2}{2m_n^*};$$

$$E_v(\mathbf{k}) = E_v(0) + \frac{\hbar^2 k^2}{2m_p^*},$$

отримаємо:

$$h\nu = E_c(0) - E_v(0) + \frac{\hbar^2 k^2}{2m_n^*} + \frac{\hbar^2 k^2}{2m_p^*}.$$

Ввівши так звану *приведену ефективну масу*

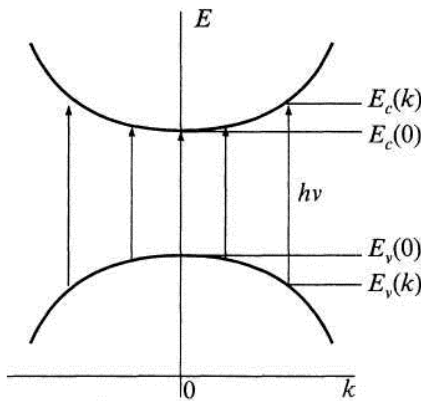


Рис. 7.1. Прямі переходи.

електронів і дірок:

$$m_r^* = \frac{m_p^* m_n^*}{m_p^* + m_n^*}$$

і враховуючи, що $E_c(0) - E_v(0) = E_g$, – ширина забороненої зони,

$$h\nu = E_g + \frac{\hbar^2 k^2}{2m_r^*}.$$

Отже, при параболічних зонах власне поглинання для прямих переходів починається з частоти, що задовольняє умову:

$$h\nu_{\min} = E_g. \quad (7.1)$$

Ця умова визначає мінімальну частоту (максимальну довжину хвилі) фотона, здатного перекинути електрон через заборонену зону. У спектрі поглинання буде спостерігатися зростання поглинання, починаючи з вказаної частоти, при її збільшенні. Аналогічна умова справедлива і для простих непараболічних зон в напівпровідниках.

Очевидно, в співвідношеннях (7.1) залишаються справедливими закони збереження енергії і імпульсу.

Коефіцієнт поглинання α , яким звичайно характеризують поглинання, є коефіцієнтом пропорційності між інтенсивністю падаючого світла та інтенсивністю поглиненого світла dl при проходженні променем шляху dx .

$$dI = -\alpha I dx. \quad (7.2)$$

Як відзначалося, при переході до власного поглинання повинне спостерігатися зростання α .

В антимоніді індію (InSb), зони якого непараболічні, але ефективна маса ізотропна, спектр власного поглинання характеризується крутим зростанням α , як показано на рис. 7.2. При цьому положення краю смуги власного поглинання залежить від концентрації домішки, що зумовлюється малою ефективною масою електронів ($m_n^* = 0,013m$). Оскільки зона провідності InSb характеризується малою густиною станів, то навіть при невеликій концентрації вільних електронів вони заповнюють досить високі рівні (рис. 7.3). При зростанні концентрації електронів край власного поглинання буде зсуватися в область коротких хвиль, тому що всі рівні поблизу дна зони провідності будуть зайняті, і поглинуті кванти повинні перекидати електрони на вищі рівні.

Зокрема, при зміні концентрації електронів від 10^{17} до 10^{19} см^{-3} оптична "ширина" забороненої зони в InSb збільшується на декілька десятків eV.

У напівпровідниках зі складними енергетичними зонами (германій, кремній і ін.) власне поглинання ускладнюється. Тут спостерігаються як прямі, так і *непрямі переходи*, хоч ймовірність останніх менша.

Якщо при прямому оптичному переході імпульс

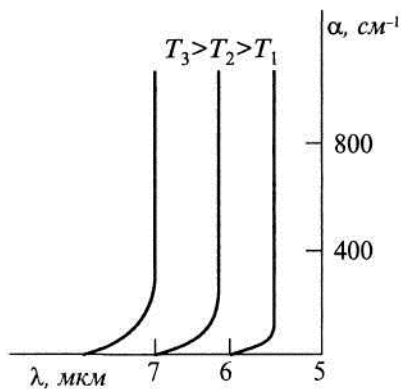


Рис. 7.2. Спектр власного поглинання InSb.

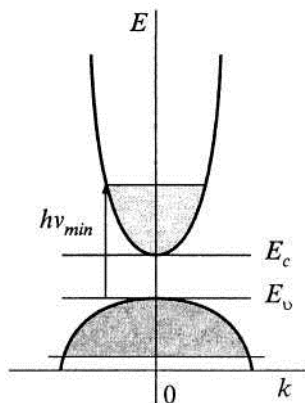


Рис. 7.3. Частина зонної структури InSb.

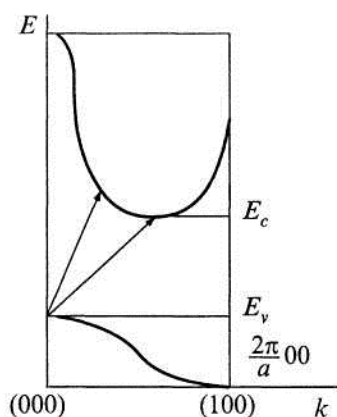


Рис. 7.4. Непрямі переходи в кремнії (одна з кривих зони провідності і одна з кривих валентної зони).

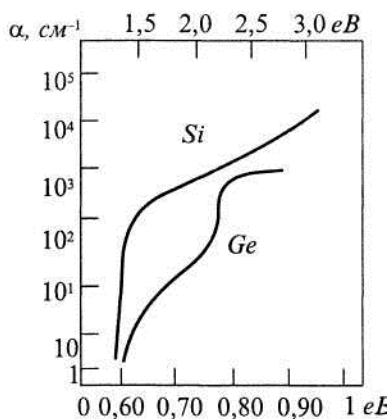


Рис. 7.5. Спектр поглинання кремнію (енергія фотонів відкладається на верхній шкалі) і германію (енергія фотонів відкладається на нижній шкалі).

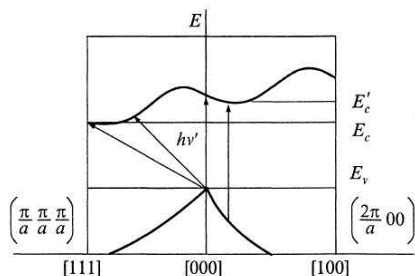


Рис. 7.6. Прямі і непрямі переходи в германію.

електрона практично не змінюється, оскільки імпульс поглинутого фотона $h\nu/c$ малий, то при *непрямих переходах* імпульс електрона значно змінюється. Така зміна імпульсу відбувається за рахунок передачі його від фонона (тобто третьої квазічастинки) електрона. При поглинанні фонона електроном разом з імпульсом електрон отримує від фонона і енергію. При непрямих переходах можливий і обернений процес – випускання фонона. У всіх цих процесах залишаються справедливими закони збереження енергії та імпульсу.

На рис. 7.4 показано переходи електронів з валентної зони в зону провідності Si.

Ширина забороненої зони, визначена оптичними вимірюваннями, не завжди збігається з шириною забороненої зони, визначеною на основі температурних вимірювань електропровідності, тобто оптична і термічна ширини забороненої зони можуть відрізнятися одна від одної. Це пов'язано із складною структурою зон і за участю фононів при взаємодії пари фотон – електрон.

Вивчення краю основної смуги оптичного поглинання монокристалів германію і кремнію, взятих у вигляді вельми тонких зразків, дало криві для коефіцієнта поглинання, зображені на рис. 7.5. Тут на нижній горизонтальній осі відкладена енергія фотонів для германію, а на верхній – для кремнію.

Крива $\alpha = f(h\nu)$ для германію має дві ділянки різкого зростання. Перша – починається при $h\nu = 0,62$ eV і відповідає непрямим переходам, друга – при $h\nu = 0,81$ eV і відповідає прямим переходам. Таке трактування узгоджується з зонною схемою для германію. Прямий перехід електрона при $k = 0$ вимагає витрати енергії (рис. 7.6):

$$h\nu' = E'_c - E_v,$$

в той час як термічна ширина забороненої зони менша і становить $E_c - E_v$. Для непрямого переходу потрібна менша енергія. Зокрема при переході на дно долини:

$$h\nu = E_c - E_v$$

Крива $\alpha = f(h\nu)$ для кремнію має круту ділянку, яка починається близько 1 eV, але потім пологіше зростає при подальшому збільшенні $h\nu$. Це пов'язано з особливостями зонної структури.

У табл. 7.1 наведено значення оптичної ширини забороненої зони для Si і Ge (рис. 7.6) і ряду інших напівпровідників і діелектриків.

Таблиця 7.1.

Оптична ширина забороненої зони для деяких напівпровідників і діелектриків.

| Напівпро- відник | E_g , eV при 300 К | E_g , eV при 0 К | Діелектрик | E_g , eV |
|---------------------|-------------------------|-----------------------|--------------------------------|------------|
| Si | 1,09 | 1,14 | Al ₂ O ₃ | 5,80 |
| Ge | 0,66 | 0,75 | TiO ₂ | 3,67 |
| InP | 1,25 | 1,34 | SnO ₂ | 3,71 |
| InAs | 0,35 | 0,45 | BaTiO ₃ | 3,0-3,1 |
| InSb | 0,175 | 0,25 | PbTiO ₂ | 3,0 |
| GaP | 2,24 | 2,40 | | |
| GaAs | 1,45 | 1,53 | | |
| GaSb | 0,70 | 0,80 | | |
| AlSb | 1,60 | 1,70 | | |

На закінчення зазначимо, що ширина забороненої зони залежить від температури і зовнішнього тиску. Відповідно зміщується край смуги власного поглинання.

У разі зростання температури змінюється амплітуда теплових коливань атомів, у зв'язку з чим зменшується ширина забороненої зони. Крім того, внаслідок зміни міжатомних відстаней, тобто об'єму тіла, з підвищенням температури може спостерігатися зменшення або збільшення ширини забороненої зони. Проте переважно із зростанням температури спостерігається зменшення ширини забороненої зони і зміщення краю смуги власного поглинання в область довгих хвиль. У деякому інтервалі температури для ряду напівпровідників можна вважати ширину забороненої зони лінійною функцією температури

$$E_g = (E_g)_{T=0} - gT. \quad (7.3)$$

Для кремнію величина $\gamma = 4 \cdot 10^{-4}$ eV/K, а для германію $\gamma = 4,5 \cdot 10^{-5}$ eV/K.

Однак іноді спостерігається зворотний ефект. Наприклад, для PbS із зростанням температури ширина забороненої зони збільшується.

При стиску в більшості випадків спостерігається збільшення ширини забороненої зони і зсув краю смуги власного поглинання в область коротких хвиль. Але можуть бути винятки (наприклад, ефект має складний характер в CdS). На рис. 7.7 наведена залежність ширини забороненої зони від тиску для деяких напівпровідників. Додатним значенням відповідає збільшення ширини забороненої зони. Як видно з рисунка, більшість залежностей немонотонні.

7.2. Домішкове поглинання світла

Поглинання світла внаслідок передавання енергії (імпульсу) від фотонів до електронів, що призводить до іонізації відповідних домішкових центрів, називають *домішковим*. Якщо поглинуті фотони перекидають електрони з рівнів домішки в зону провідності, то буде спостерігатися смуга поглинання, зсунута порівняно з

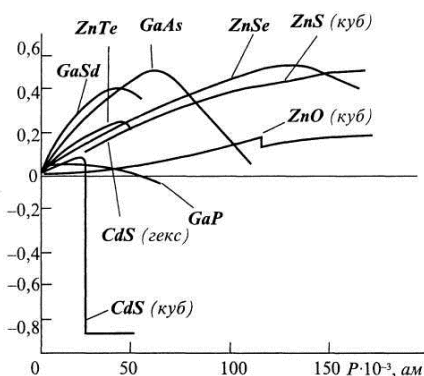


Рис. 7.7. Зміна ширини забороненої зони в деяких напівпровідниках залежно від тиску P .

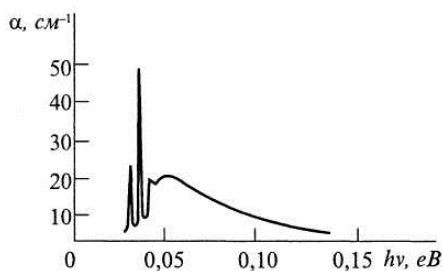


Рис. 7.8. Спектр поглинання, зв'язаний з домішкою бору в кремнії.

власним поглинанням в бік довгих хвиль. В той же бік зсунута смуга поглинання, що відповідає перекиданню електронів з валентної зони на рівні домішки.

На рис. 7.8 як приклад наведений спектр поглинання, зумовлений іонізацією атомів бору (акцептора) в кремнії. Звичайно спектри домішкового поглинання містять декілька максимумів. Це пов'язано з тим, що електрони на домішкових рівнях можуть знаходитися, в основному, і в збуджених станах, подібно до електрона в атомі водню. Спектри домішкового поглинання звичайно вимірюють при низькій температурі ($T < 20$ К), оскільки енергія іонізації домішок може бути невеликою і необхідно зменшити термічну генерацію. Крім того, внаслідок взаємодії електронів з фононами смуги поглинання розширюються. Необхідно також обмежувати концентрацію домішки, бо взаємодія їх призводить до розширення смуг поглинання.

Спектри домішкового поглинання дають змогу отримувати енергію іонізації дрібних домішкових центрів з дуже високою точністю.

7.3. Екситонне поглинання світла

Екситони – збуджені стани електронної системи кристала, що складаються з пов'язаних між собою електрона і дірки. Якщо під час поглинання світла електрон покидає валентну зону, але не переходить в зону провідності, а має енергію, відповідну тому або іншому рівню в забороненій зоні, він може блукати кристалом, захоплюючи за собою дірку валентної зони. Очевидно, цей випадок особливий, оскільки електрон не закріплюється на тому або іншому рівні в забороненій зоні. Крім того, рухи такого електрона не створюють електричного струму, оскільки разом з електроном переміщується і дірка.

Екситон може руйнуватися в результаті теплового "дозбудження", тобто термічної дисоціації, при якій створюються вільний електрон у зоні провідності та вільна дірка у валентній зоні. Може також статися зникнення екситону з випусканням фононів або фотона, електрон при цьому повернеться у валентну зону.

Якщо розміри екситона великі, порівняно зі сталою ґратки, то взаємодію електрона і дірки можна подати як кулонівську взаємодію двох точкових зарядів. Отже, екситон нагадує вод неподібну систему.

Рівні енергії екситону наведені на рис. 7.9. Тут $h\nu_{\infty} = E_g$ відповідає випадку, коли екситон зникає внаслідок розриву зв'язку електрон – дірка і утворення двох вільних носіїв заряду.

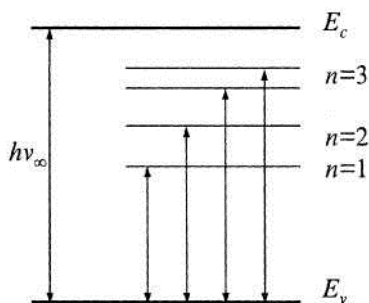


Рис. 7.9. Рівні енергії екситону

Експериментально спектри поглинання, відповідні екситонам, виявлені в CuO_2 , $CdSe$, ZnS , HgI_2 , PbI_2 та інших напівпровідниках.

7.4. Поглинання світла вільними носіями заряду

При поглинанні світла вільні електрони переходять на більш високі рівні в зоні провідності, тобто "розігріваються". Відповідно енергія світлової хвилі зменшується. Поглинання світла може бути і на вільних дірках.

Розглянемо основні закономірності поглинання світла вільними електронами напівпровідника, що має просту будову зон в k -просторі.

Взаємодія світла з вільними електронами напівпровідника призводить до зменшення амплітуди падаючої електромагнітної хвилі, так що напруженість електричного поля можна подати у вигляді:

$$e = A \exp \left[i \omega \left(t - \frac{\bar{n}}{c} x \right) \right], \quad (7.4)$$

де \bar{n} – показник заломлення напівпровідника. Амплітуда хвилі:

$$A = \mathbf{e}_0 e^{-\frac{s}{2nc\epsilon_0}},$$

де \bar{n} – відстань, пройдена світлом у напівпровіднику; σ – питома електропровідність напівпровідника; c – швидкість світла у вакуумі, ϵ_0 – амплітуда падаючої електромагнітної хвилі.

Отже, потік енергії хвилі експоненціально зменшується з відстанню x . З іншого боку, інтенсивність світла I , яке поширилося в середовище на глибину x , пропорційна до e^{-ax} , де a – коефіцієнт поглинання. Беручи до уваги, що $I \sim A^2$, отримуємо:

$$a = \frac{s}{\bar{n}cE_0} = \frac{e^2 \langle t \rangle}{\bar{n}c\mathbf{e}_0 m^*} n. \quad (7.5)$$

Отже, коефіцієнт поглинання світла вільними електронами визначається їх концентрацією n , ефективною масою m^* , середнім часом релаксації $\langle t \rangle$, а також показником заломлення середовища \bar{n} . Оскільки останній залежить від довжини світлової хвилі, то і $a = f(\lambda)$. При складній будові зон картина ускладнюється через виникнення переходів між окремими підзонами.

7.5. Поглинання світла кристалічною ґраткою

Поглинання світла кристалічною ґраткою відбувається через взаємодію електромагнітного поля світлової хвилі з рухомими атомами у вузлах ґратки, тобто поглинання світла ґраткою пов'язане зі зміною енергії коливань атомів напівпровідника. Оскільки кристалічна ґратка може поглинати електромагнітне поле лише за певних значень енергії фотонів, спектр характеризується рядом піків поглинання, які звичайно накладаються на поглинання вільними носіями заряду. На рис. 7.10 наведено спектр поглинання світла кристалічними ґратками германію і кремнію. На нижній шкалі осі абсцис відкладене значення $1/(\lambda \cdot 10^3)$, виражене в см^{-1} для германію, на верхній – для кремнію. За довжинами хвиль

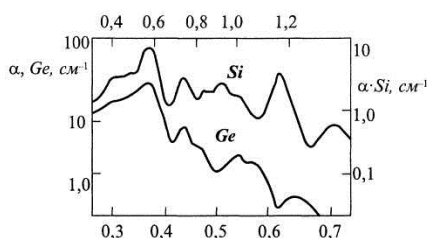


Рис. 7.10. Залежність коефіцієнта поглинання світла кристалічною ґраткою германію і кремнію від величини $\frac{1}{\lambda \cdot 10^3} \text{см}^{-1}$

шкала для германію відповідає інтервалу від 14 до 33 мкм, для кремнію – від 7 до 20 мкм. Коефіцієнт поглинання германію більший, ніж кремнію, оскільки густина речовини германію перевищує густину кремнію. Докладний аналіз показує, що світло поглинається тільки поперечними оптичними коливаннями ґратки.

Лекція 8

ЯВИЩЕ ФОТОПРОВІДНОСТІ

План

- 8.1. Квантовий вихід фотоелектру.
- 8.2. Релаксація фотопровідності і стаціонарна фотопровідність.
- 8.3. Ефект Дембера.
- 8.4. Вентильна фото-ЕРС.
- 8.5. Фотомагнітоелектричний ефект (ефект Кікоїна-Носкова).

Література

- 7.5. Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.
- 7.6. Мікроелектронні сенсори фізичних величин: Науково-навчальне видання. В 3 томах. Том 1. / Вуйцік В., Готра З.Ю., Григор'єв В.В., Каліта В., Мельник О.М., Потенці Є.; за ред. З.Ю. Готри. – Львів: Ліга-Прес, 2002. – 475 с.
- 7.7. Микроэлектроника. В 9 кн. / Под ред. Л. А. Коледова. Кн. 1. Физические основы функционирования изделий микроэлектроники / О.В. Митрофанов, Б.М. Симонов, Л.А. Коледов.: Высшая школа, 1987. – 168 с.
- 7.8. Степаненко И. П. Основы микроэлектроники. – М.: Сов. радио, 1980. – 286 с.

8.1. Квантовий вихід фотоелектру

Для монохроматичного світла інтенсивність I можна виразити через потік фотонів $q = \frac{I}{h\nu}$. Концентрація носіїв заряду, що генеруються, пропорційна до інтенсивності поглинутого випромінювання αI (α – коефіцієнт поглинання). Отже, наприклад, концентрація надлишкових електронів, що створюються протягом часу t , буде:

$$\Delta n = \alpha a q t = b' a q h \nu t, \quad (8.1)$$

тобто

$$b' = \frac{b}{h\nu}.$$

Для власного поглинання $\Delta n = \Delta p$; тому кількість пар носіїв заряду в одиниці об'єму, що генеруються світлом за одиницю часу, визначають через інтенсивність світла, що поглинається, і коефіцієнт пропорційності β який називають *квантовим виходом*. Очевидно, квантовий вихід визначає кількість пар носіїв заряду, утворених одним поглинутим фотоном (квантом), якщо інтенсивність світла I вимірювати кількістю фотонів (квантів) за секунду. Звичайно величина $b \leq 1$. Однак високоенергетичні фотони можуть створювати ударну іонізацію і в цьому випадку $\beta > 1$.

При домішковому поглинанні утвориться один тип носіїв заряду, і квантовий вихід β визначає кількість носіїв заряду, утворених одним поглинутим фотоном.

Якщо інтенсивність світла однакова в усьому об'ємі напівпровідника і не змінюється з часом, то встановлюється динамічна рівновага процесів генерації і рекомбінації, якому відповідають незмінні значення надлишкових концентрацій:

$$\Delta n_{CT} = \frac{b'a}{t_n}, \Delta p_{CT} = \frac{b'a}{t_p},$$

де t_n і t_p – час життя електронів і дірок, відповідно. Тоді стаціонарна фотопровідність:

$$S_{f,CT} = eb'aI(t_n m_n + t_p m_p). \quad (8.2)$$

При значній перевазі одного з доданків буде спостерігатися *монополярна фотопровідність*.

Характеристикою речовини є *світлочутливість* S_ϕ , що визначається як відношення фотопровідності до інтенсивності світла. Отже, якщо як фотопровідність використовувати стаціонарну, $\sigma_{f,CT}$, то

$$S_f = \frac{S_{f,CT}}{I} = eb'a(t_n m_n + t_p m_p). \quad (8.3)$$

У СІ світлочутливість S_ϕ вимірюють у $\frac{\text{м}}{\text{Ом} \cdot \text{Вт}}$.

Зазначимо також, що величину фотоструму j_ϕ називають *фотовідгуком*, причому часто вимірюють її у відносних одиницях.

Квантовий вихід, який дорівнює одиниці, можна спостерігати на широкозонних кристалах, наприклад, типу алмазу, який є діелектриком ($E_g = 5,3$ еВ). Крім того, фотони з великою енергією, як вже відзначалося, можуть дати $\beta > 1$. Більше того, в кремнії, наприклад, спостерігається залежність квантового виходу від енергії поглинутого фотона.

Оскільки ефективна маса електронів і дірок у кремнії порівняно мало відрізняється одна від одної, можна вважати, що у разі виникнення пари електрон-дірка за рахунок ударної іонізації надмірна кінетична енергія розподіляється між електроном і діркою, а також первинним електроном порівну. У зв'язку з цим потрібно чекати істотного зростання квантового виходу при $h\nu \geq 3E_g$, тобто приблизно з 3,3 еВ для кремнію при кімнатній температурі (що й спостерігається експериментально, рис. 8.1).

При поглинанні фотонів великої енергії, що відповідають рентгенівському або γ -випромінюванню, квантовий вихід β відповідно збільшується, що показано на рис. 8.2. Тут же наведено квантовий вихід для випадку опромінення електронами і α -частинками відповідних енергій.

8.2. Релаксація фотопровідності і стаціонарна фотопровідність

Перехідні процеси після вмикання або вимкнення постійного освітлення характеризує її релаксація. Якщо не враховувати рекомбінацію на поверхні і якщо випромі-

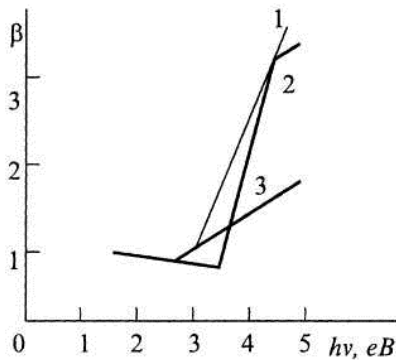


Рис. 8.1. Квантовий вихід β в кремнії за температури 1 – 100 K; 2 – 300 K; 3 – 400 K.

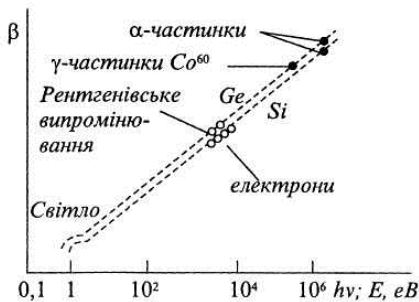


Рис. 8.2. Квантовий вихід β у кремнії і германії при опроміненні квантами $h\nu$ і частинками різної енергії.

нювання поглинається рівномірно в усьому об'ємі, зміна надлишкової концентрації електронів визначається рівнянням неперервності. Крім того, вважаємо, що напівпровідник ізольований від інших тіл і дрейф носіїв заряду відсутній. Отже, маємо:

$$\frac{\partial \Delta n}{\partial t} = b'aI - g_r(np - n_0p_0). \quad (8.4)$$

При низькому рівні інжекції рекомбінаційний доданок можна записати через $\frac{\Delta n}{t_n}$; тоді отримаємо

$$\frac{\partial \Delta n}{\partial t} = b'aI - \frac{\Delta n}{t_n}.$$

Отже, у дірковому напівпровіднику концентрація надлишкових неосновних носіїв заряду при миттєвому (ступінчастому) вмиканні освітлення згідно з розв'язком рівняння (8.4) з початковою умовою $\Delta n = 0$ при $t = 0$ приймає позначення Δn_{CT} і змінюється за законом

$$\Delta n = \Delta n_{cm} \left(1 - e^{-\frac{t}{t_n}} \right),$$

$$\Delta n_{CT} = t_n b'aI.$$

Якщо, навпаки, в напівпровіднику вже є стаціонарна надмірна концентрація Δn_{CT} і в деякий новий момент $t = 0$ проводиться вимкнення освітлення, то, починаючи відлік з цього часу, надлишкова концентрація неосновних носіїв заряду змінюється за законом:

$$\Delta n = \Delta n_{CT} e^{-\frac{t}{t_n}}.$$

На рис. 8.3, б наведено криві $\Delta n = f(t)$ при вмиканні і вимкненні світла, в припущенні, що процеси вмикання і вимкнення миттєві, тобто світловий імпульс прямокутний (рис. 8.3, а), що використано при виведенні виразів для Δn .

При високому рівні інжекції маємо:

$$\frac{\partial \Delta n}{\partial t} = b'aI - g_r (\Delta n)^2. \quad (8.5)$$

У цьому рівнянні враховані як рекомбінаційна складова, так і генераційна.

Інтегруючи рівняння (8.5) з урахуванням початкової умови ($\Delta n = 0$ при $t = 0$), отримаємо:

$$\Delta n = \sqrt{b'aI/g_r} \operatorname{th}(t\sqrt{g_r b'aI}),$$

тобто наростаюча величина Δn визначається постійним коефіцієнтом $\sqrt{b'aI/g_r}$ і змінним з часом співмножником у вигляді гіперболічного тангенса. У разі вимкнення освітлення отримують:

$$\Delta n = \sqrt{\frac{b'aI}{g_r}} \cdot \frac{1}{1 + t\sqrt{g_r b'aI}},$$

тобто гіперболічну від часу залежність.

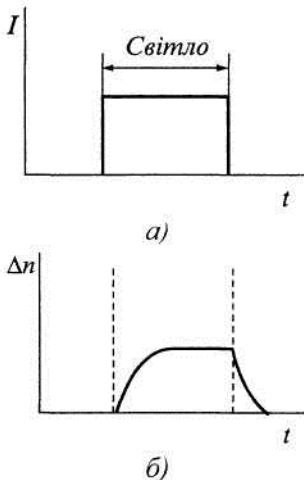


Рис. 8.3. Залежність надлишкової концентрації від часу після вмикання і вимкнення світла

Якщо використати миттєвий час життя, то рівняння (8.5) можна застосувати і до випадку високого рівня інжекції.

При розрахунку встановленого значення Δn_{CT} потрібно покласти

$$\frac{\partial \Delta n}{\partial t} = 0,$$

так що за (8.4) маємо:

$$t_n = \frac{\Delta n_{CT}}{b'aI}.$$

Причому останнє співвідношення виражає час життя електронів як неосновних носіїв заряду в дірковому напівпровіднику при лінійній рекомбінації і миттєвий час життя електронів при квадратичній рекомбінації.

Отже,

$$\Delta n_{CT} = t_n b'aI,$$

тобто при лінійній рекомбінації Δn_{CT} маємо лінійну функцію інтенсивності світла I , а при квадратичній рекомбінації Δn_{CT} – нелінійну функцію I , оскільки $\tau_n = f(I)$.

На рис. 8.4 наведено залежності Δn_{CT} від інтенсивності світла I для лінійної (1) і квадратичної (2) рекомбінацій.

Вище не враховувався вплив пасток захоплення. Заповнення останніх може затягувати наростання Δn при вмиканні освітлення, а звільнення – затягувати її спад при вимкненні освітлення. Зокрема, при вмиканні світла може спостерігатися S-подібне наростання Δn (рис. 8.5). Нехай N_t є концентрація пасток захоплення, які здатні захоплювати і потім упродовж часу θ втримувати носій заряду. Якщо $q = t_n$, то пастка протягом часу життя нерівноважного електрона здатна захопити його неодноразово. Такі пастки називають пастками α -типу. Якщо ж $q > t_n$, то пастка може захоплювати електрон не більше ніж один раз. В останньому випадку пастки позначають як β -тип. Якщо пастки вільні, то в перший момент вони будуть захоплювати електрони, зменшуючи швидкість наростання Δn і затягуючи час наростання. Аналогічно при спустошенні пасток після вимкнення світла буде спостерігатися затягнення Δn спаду. Початковий етап наростання і спаду Δn визначається величиною θ у випадку α -пасток і τ_n – в випадку β -пасток. Але у міру насичення або спустошення пасток їх роль у вимірюванні Δn зменшується.

Якщо побудувати залежність фотовідгуку від енергії фотонів або довжини світлової хвилі, що поглинається, то виходять складні криві, які в області домішкового поглинання залежать від наявності домішок. Так, на рис. 8.6 і 8.7 наведено дві криві спектральної залежності фотовідгуку (фотовідгуку, відкладеного на осі ординат у відносних одиницях) для германію, легованого міддю і цинком. В області власного поглинання (ліва

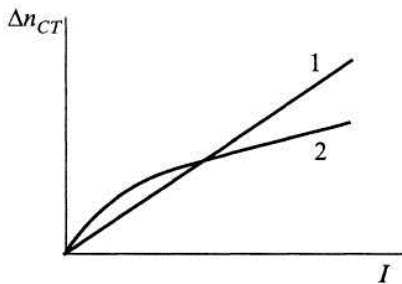


Рис. 8.4. Залежність надлишкової концентрації Δn_{CT} , що встановилася, від інтенсивності світла I при лінійній (1) і квадратичній (2) рекомбінації.

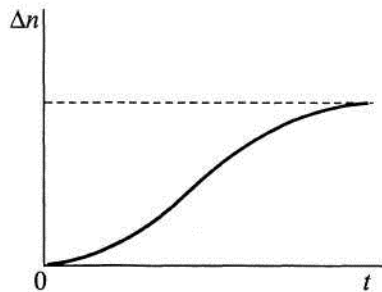


Рис. 8.5. Зміна надлишкової концентрації після вмикання світла, що затягується процесами захоплення носіїв заряду пастками.

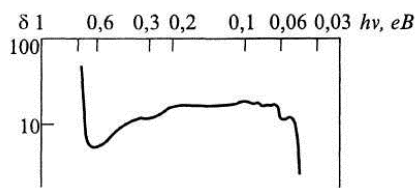


Рис. 8.6. Спектральна залежність фотопровідності в германію, легованого міддю.

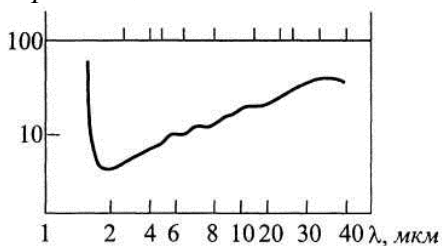


Рис. 8.7. Спектральна залежність фотопровідності в германію, легованого цинком.

частина рисунка) криві збігаються, а в домішковій області сильно відрізняються.

Якщо спектр власного поглинання, круто наростаючи з боку його довгохвильової границі, може мати постійне або зростаюче значення а при подальшому зменшенні довжини хвилі в досить широкому діапазоні, то після проходження довгохвильової границі власного поглинання внутрішній фотоэффект досягає максимуму і потім швидко зменшується, хоч інтервал довжин хвиль продовжує відповідати власному поглинанню. Тут позначається фактичне зменшення швидкості генерації в міру заглиблення в об'єм у зв'язку з сильним поглинанням світла поверхневими шарами. У цьому випадку фотопровідність буде сильно залежати від розмірів зразка.

Розглянемо вплив поверхневої рекомбінації носіїв заряду на фотопровідність. Поверхнева рекомбінація спричиняє зменшення концентрації нерівноважних носіїв заряду в приповерхневому шарі.

Якщо поверхнева рекомбінація є значною, то їй відповідає малий час життя пари електрон-дірка в приповерхневому шарі τ_s . Для тонких зразків це може визначати темп всієї рекомбінації, оскільки довільний парціальний час життя визначається як величина, обернена ймовірності відповідного процесу, а сумарний процес визначається сумою ймовірностей $1/\tau = 1/\tau_V + 1/\tau_s$, де τ_V – час життя пари електрон-дірка в об'ємі кристала.

У випадку, якщо швидкість поверхневої рекомбінації $s \ll 2D/d$, де d – товщина зразка, D – біполярний коефіцієнт дифузії, можна вважати, що $\tau_s = d/2s$.

Якщо час життя пари електрон-дірка в об'ємі τ_V значно більший за τ_s , то для ефективного часу життя пари електрон-дірка отримуємо $\tau \approx \tau_s$.

Отже, згідно з виразом $1/\tau = 1/\tau_V + 1/\tau_s$, за умови $\tau_s = d/2s$, $1/\tau = 1/\tau_V + d/2s$, де s – швидкість поверхневої рекомбінації, d – в напрямку променя світла. Звідси

$$\frac{\tau_V}{\tau} = 1 + \frac{2s\tau_V}{d}. \quad (8.6)$$

З (8.6) видно, що, якщо $2s\tau_V \ll d$, то $\tau = \tau_V$, тобто поверхнева рекомбінація практично не позначається на часі життя. Навпаки, при $2s\tau_V \gg d$ маємо:

$$\frac{1}{\tau} = \frac{2s}{d} \text{ або } \tau = \frac{d}{2s} = \tau_s,$$

тобто фотопровідність визначається не об'ємним часом життя, а поверхневим. Відповідно, σ_ϕ виявляється меншою, ніж в першому випадку.

8.3. Ефект Дембера

При освітленні напівпровідника світлом, що сильно поглинається, в приповерхневому шарі надмірна концентрація нерівноважних носіїв заряду буде більшою, ніж в об'ємі, оскільки туди доходить лише невелика частина падаючого на поверхню світла. Тому буде

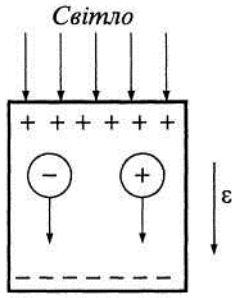


Рис. 8.8. Схема виникнення ефекту Дембера.

відбуватися дифузія електронів і дірок від поверхні в глибину зразка, тобто в напрямі падаючого променя.

Електрони, що мають більшу рухливість, ніж дірки, будуть їх випереджати, що призведе до розділення зарядів. Безпосередньо біля поверхні буде створюватися позитивний заряд, а неосвітлена протилежна грань зразка (рис. 8.8) буде заряджатися негативно.

Якщо позначити напруженість внутрішнього електричного поля Дембера ϵ , а густини струмів електронів і дірок відповідно j_n, j_p то:

$$j_n = enm_n \mathbf{e} + m_n kT \frac{dn}{dx};$$

$$j_p = enm_p \mathbf{e} - m_p kT \frac{dp}{dx};$$

$$j = j_n + j_p = e(nm_n + pm_p) \mathbf{e} + kT \left(m_n \frac{dn}{dx} - m_p \frac{dp}{dx} \right).$$

У випадку ізолюваного напівпровідника $j = 0$, так що:

$$\mathbf{e} = -\frac{kT}{e} \cdot \frac{m_n \frac{dn}{dx} - m_p \frac{dp}{dx}}{nm_n + pm_p}.$$

Вважаючи $\frac{dn}{dx} = \frac{dp}{dx} = \frac{d\Delta n}{dx} = \frac{d\Delta p}{dx}$ знаходимо

абсолютне значення ЕРС Дембера між точками x_1 і x_2 .

$$\int_{x_1}^{x_2} \mathbf{e} dx = \frac{kT}{e} \int_{x_1}^{x_2} \frac{m_n - m_p}{nm_n + pm_p} d\Delta n = V_{12}. \quad (8.7).$$

Оскільки $s = e(nm_n + pm_p)$ і $\Delta s = e(m_n + m_p)\Delta n$, то

$$d\Delta n = \frac{d\Delta s}{e(m_n + m_p)}. \quad (8.8)$$

Отже, з (8.7) і (8.8) маємо:

$$U_{12} = \frac{kT}{e} \frac{m_n - m_p}{m_n + m_p} \int_{x_1}^{x_2} \frac{d\Delta s}{s}.$$

Оскільки $d\Delta s = ds$, звідси отримуємо:

$$U_{12} = \frac{kT}{e} \frac{m_n - m_p}{m_n + m_p} \ln \frac{s_2}{s_1}. \quad (8.9)$$

Отже, ЕРС Дембера тим більша, чим більше відрізняються рухливості електронів і дірок. При $\mu_n = \mu_p$ ефект відсутній. Якщо напівпровідник має монополярну провідність, то ЕРС Дембера, наприклад, при $p = 0$, дорівнює:

$$U_{12} = \frac{kT}{e} \ln \frac{n_2}{n_1}$$

Повернемося до біполярної провідності, тобто до співвідношення (8.9), вважаючи $\Delta s = s_2 - s_1$ малою величиною $s = (s_2 - s_1)/2$. Тоді, розкладаючи $\ln \frac{s_2}{s_1}$ в ряд і обмежуючись першим членом розкладу, отримуємо:

$$U_{12} = \frac{kT}{e} \frac{m_n - m_p}{m_n + m_p} \frac{\Delta\sigma}{S}$$

Для значень $\sigma = 1 \text{ Ом}^{-1}\text{см}^{-1}$, $T = 300 \text{ К}$, $\mu_n = 3900 \text{ см}^2/(\text{В}\cdot\text{с})$, $\mu_p = 1900 \text{ см}^2/(\text{В}\cdot\text{с})$ германій), $\Delta\sigma = 0,2 \text{ Ом}^{-1}\text{см}^{-1}$, отримаємо $U_{12} = 1,7 \cdot 10^{-3} \text{ В}$. Отже, навіть при порівняно великій величині $\Delta\sigma$ отримуємо мале значення ЕРС Дембера.

8.4. Вентильна фото-ЕРС

Якщо р-п-перехід розташований поблизу поверхні напівпровідника так, що падаючі на поверхню кванти світла генерують пари електрон-дірка досить близько від р-п-переходу, то в останньому спостерігається розділення пар і пониження контактної різниці потенціалів. Необхідно, щоб місце генерації пари знаходилося не далі, як на відстані приблизно дифузійної довжини. У такому випадку носії заряду не встигають рекомбінувати, а дифузійний рух підводить частину з них до р-п-переходу. Електричне поле, створене контактною різницею потенціалів, переводить надлишкові дірки з n-області в р-область, а надлишкові електрони, навпаки, з р-області в n-область, що призводить до зниження контактного потенціального бар'єра на величину, відповідну вентильній фото-ЕРС. Якщо р-п-перехід освітлюється так, що промені світла напрямлені вздовж границь розділу областей, то вказана генерація пар спостерігається з обох боків гід переходу (рис. 8.9, а) але при цьому діє лише приповерхнева його частина, бо світло поглинається в приповерхневому шарі і не доходить до глибинних областей.

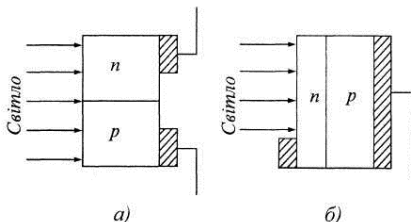


Рис. 8.9. Схема виникнення вентильної фото-ЕРС.

У випадку, коли світло падає перпендикулярно до р-п-переходу (рис. 8.9, б), спостерігається генерація електронно-діркових пар в області одного типу електропровідності (приповерхневої області), і якщо границя р-п-переходу залягає неглибоко, то виникає вентильна фото-ЕРС за рахунок пересування через р-п-перехід неосновних носіїв заряду з вказаної області.

Якщо р-п-перехід до освітлення знаходиться в рівновазі, то повний струм через нього дорівнює нулю, тобто густина струму електронів і дірок з протилежних областей однакові між собою:

$$j_n^{(n)} = j_n^{(p)}; j_p^{(p)} = j_p^{(n)},$$

тут $j_n^{(n)}$ і $j_n^{(p)}$ – густини електронних потоків з n- та – p-областей відповідно; $j_p^{(p)}$ і $j_p^{(n)}$ – густини діркових потоків з p- та n- областей відповідно;

За рахунок опромінення створюється фотострум j_f , який збігається зі струмом неосновних носіїв заряду, тобто сума густин струмів неосновних носіїв заряду:

$$j_f = j_n^{(p)} + j_p^{(n)} = j_{\text{неосн}},$$

причому суму густин струмів неосновних носіїв заряду без фотоструму можна замінити:

$$j_n^{(p)} + j_p^{(n)} = j_{nc} + j_{ps} = \frac{eD_n n_p}{L_n} + \frac{eD_p n_n}{L_p},$$

де L_n і L_p – дифузійні довжини, відповідно, електронів і дірок.

Фотострум знижує контактний потенціальний бар'єр на величину eU_f , внаслідок чого струм основних носіїв заряду через р-п-перехід збільшується, так що:

$$j_n^{(n)} = j_{ns} e^{\frac{eU_f}{kT}}; \quad j_p = j_{ps} e^{\frac{eU_f}{kT}}. \quad (8.10)$$

Загальна алгебраїчна сума густини струмів:

$$j_f + j_n^p + j_p^n - j_n^n - j_p^p = j. \quad (8.11)$$

Якщо коло замкнене, то струм j буде протікати в ньому. Якщо ж воно розімкнене, то $j = 0$.

Підставимо (8.10) в (8.11) і виразимо j_f у вигляді

$$j_f = j + j_s \left(e^{\frac{eU_f}{kT}} - 1 \right).$$

Звідси різниця потенціалів створена на р-п-переході внаслідок опромінення:

$$U_f = \frac{kT}{e} \ln \left(1 + \frac{j_f - j}{j_s} \right).$$

При розімкненому колі φ_f є вентильною фото-ЕРС, тому:

$$U_{f \text{ вент}} = \frac{kT}{e} \ln \left(1 + \frac{j_f}{j_s} \right).$$

Отже, вентильна фото-ЕРС визначається температурою T і величиною густини струму j_f , яка, в свою чергу, залежить від інтенсивності світла.

Вентильна фото-ЕРС є джерелом напруги в напівпровідникових перетворювачах енергії світла в електричну енергію (*сонячні батареї* і ін.). За наявності мікрогетеропереходів і бар'єрних шарів типу напівпровідник – квазіметал або інших потенціальних бар'єрів в об'ємі неоднорідного напівпровідника сумарна вентильна фото-ЕРС може бути досить значною.

8.5. Фотомагнітоелектричний ефект (ефект Кікоїна-Носкова)

Якщо поверхня напівпровідника опромінюється світлом, то в приповерхневому шарі генеруються надмірні електронно-діркові пари. У глибині напівпровідника створюється дифузійний потік електронів і дірок. Якщо зразок вмістити у зовнішнє постійне магнітне поле, напрямлене перпендикулярно до променів світла, то дифундуючі носії заряду різних знаків будуть відхилятися силою Лоренца до протилежних граней (рис. 8.10), створюючи ЕРС.

На рис. 8.11, б показано фізичні напрями струмів електронів і дірок, відповідних напрямку руху зарядів. Якщо магнітне поле перпендикулярне до площини рисунка і напрямлене від нас, то дірки відхиляються до

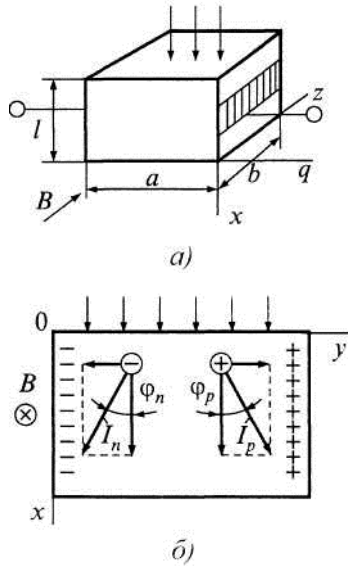


Рис. 8.10. Схема виникнення ефекту Кікоїна – Носкова.

правої грані, а електрони – до лівої. Електричне поле, що виникає, протидіє вказаному відхиленню і процес досягає рівноваги, коли струм провідності, зумовлений дією поля, компенсує магнітодифузійний струм. Цій рівновазі і відповідає фотоманітна ЕРС.

Уявимо, що швидкість поверхневої рекомбінації мала. При слабкому магнітному полі кути Холла φ_p і φ_n дорівнюють:

$$j_p \approx Am_p B \approx \operatorname{tg} j_p;$$

$$j_n \approx Am_n B \approx \operatorname{tg} j_n.$$

Проекції повних густин струмів дірок і електронів на вісь y (рис. 8.10. б) будуть:

$$j_{py} = j_{px} \operatorname{tg} j_p = Am_p B j_{px};$$

$$j_{ny} = j_{nx} \operatorname{tg} j_n = Am_n B j_{nx}.$$

Густина магнітодифузійного струму:

$$j_y = j_{py} + j_{ny} = AB(m_p j_{px} - m_n j_{nx}). \quad (8.12)$$

Вважаючи, що $\Delta n = \Delta p$, дифузійні густини токів дірок та електронів:

$$j_{px} = -j_{nx} = -eD \frac{\partial \Delta n}{\partial x}. \quad (8.13)$$

Тут D – коефіцієнт біполярної дифузії, який визначається для напівпровідника р-типу $D = D_n$, для n – типу $D = D_p$.

Підставивши (8.12) в (8.13), отримаємо

$$-j_y = eABD(m_p + m_n) \frac{\partial \Delta n}{\partial x}.$$

Нехай контакти на зразку нанесені на бокові поверхні, площа кожної дорівнює bl , тобто розглядаємо всі розрізи зразка в площині xz (на рис. 8.10, а показана частина контакту з площею bdx). Всі величини, що входять в останнє рівняння, за винятком Δn , вважаються постійними і повний струм при роз'єднаному зовнішньому колі:

$$I = \int_0^l j_x b dx = -eABDb(m_p + m_n) \int_0^l d\Delta n =$$

$$= eABDb(m_p + m_n) [\Delta n(0) - \Delta n(l)]. \quad (8.14)$$

В умовах рівноваги цей струм дорівнює струму фотоманітної ЕРС, так що його можна виразити через $U_{\text{фм}}$, та повну провідність зразка Ω :

$$I_{\text{фм}} = \Omega U_{\text{фм}}. \quad (8.15)$$

Оскільки абсолютні значення струмів (8.14) і (8.15) однакові (хоч струми мають протилежні знаки), прирівнюємо їх і отримуємо:

$$U_{\text{фм}} = \frac{eABDb(m_p + m_n) [\Delta n(0) - \Delta n(l)]}{\Omega}. \quad (8.16)$$

Вираз для повної провідності можна записати у вигляді

$$\Omega = \frac{bl}{a} S_0 + \frac{b}{a} \int_0^l \Delta dx.$$

Тут перший доданок є темною електропровідністю, а другий – визначається надлишковими носіями заряду. Оскільки $\Delta n = \Delta p$, то останній вираз можна подати як

$$\Omega = \frac{bl}{a} S_0 + \frac{eb}{a} (m_p + m_n) \int_0^l \Delta n dx.$$

Якщо врахувати, що:

$$\Delta n = (\Delta n)_0 e^{-\frac{x}{L_n}} dx = \Delta n(0) e^{-\frac{x}{L_n}},$$

де L_n – дифузійна довжина (для р-типу напівпровідника), отримаємо:

$$\int_0^l \Delta n dx = \Delta n(0) \int_0^l e^{-\frac{x}{L_n}} dx = -\Delta n(0) e^{-\frac{x}{L_n}} L_n \Big|_0^l.$$

При $l = L_n$, для товстого зразка маємо:

$$\int_0^l \Delta n dx = L_n \Delta n(0).$$

Отже, у цьому випадку отримаємо:

$$U_{\phi_m} = \frac{eaABD(m_p + m_n)\Delta n(0)}{lS_0 + e(m_p + m_n)L_n\Delta n(0)}.$$

При високому рівні інжекції перший доданок у знаменнику можна відкинути і тоді:

$$U_{\phi_m} = \frac{aABD}{L_n}. \quad (8.17)$$

Навпаки, при низькому рівні інжекції в знаменнику (8.17) можна знехтувати другим доданком, так що отримаємо фотомагнітну ЕРС:

$$U_{\phi_m} = eaABD \frac{(m_p + m_n)\Delta n(0)}{lS_0}.$$

Якщо швидкістю поверхневої рекомбінації знехтувати, то концентрація надлишкових носіїв на поверхні визначається інтенсивністю збуджуючого світла I , квантовим виходом β і часом життя носіїв заряду τ . Отже,

$$L_n = \Delta n(0) = b'tI.$$

Отже, при низькому рівні інжекції і малій швидкості поверхневої рекомбінації з урахуванням рівності $L_n^2 = Dt$; отримуємо:

$$U_{\phi_m} = eaABDb'tI \frac{m_p + m_n}{lS_0 L_n} = eaABb'I L_n \frac{m_p + m_n}{lS_0}.$$

Графік залежності фотомагнітної ЕРС з від інтенсивності освітлення для германію наведений на рис. 8.11, де лінійна частина відповідає низькому рівню інжекції, насичення – високому рівню інжекції.

Ефект Кікоїна-Носкова, з одного боку, аналогічний до

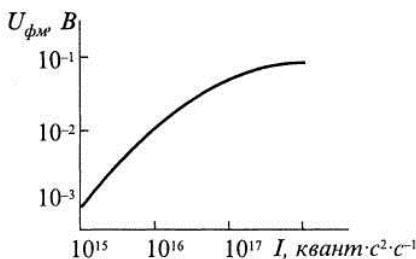


Рис. 8.11. Залежність фотомагнітної ЕРС від інтенсивності освітлення германію.

явища Холла, з іншого боку – до термомагнітного ефекту Нернста-Еттінгсгаузена. Фотомагнітоелектричний ефект використовують для дослідження різних нерівноважних процесів у напівпровідниках, зокрема, для визначення малого часу життя, наприклад близько 10^{-10} с.

Лекція 9

ФІЗИЧНІ ПРОЦЕСИ В ПОВЕРХНЕВИХ ШАРАХ

План

- 9.1. Поверхневі стани
- 9.2. Приповерхневий просторовий заряд
- 9.3. Поверхнева провідність і ефект поля
- 9.4. Поверхнева рекомбінація

Література

- 9.1. Штернов А. А. Физические основы конструирования, технология РЗА и микроэлектроники. – М: Радио и связь, 1981. – 248 с.
- 9.2. Киреев И. П. Физика полупроводников. – М.: Высшая школа, 1975. – 247 с.

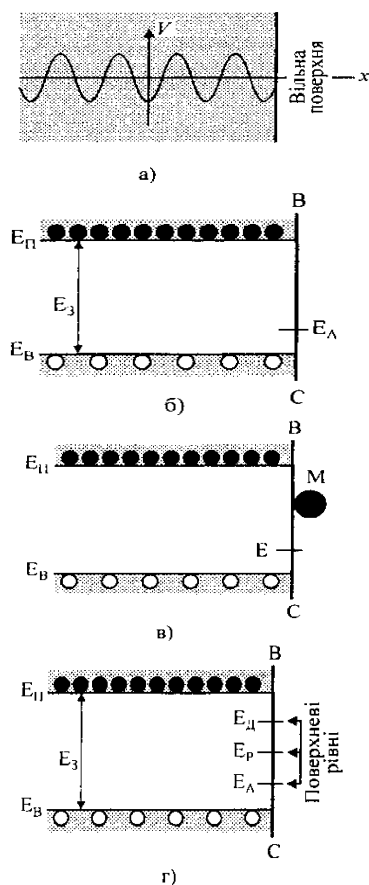


Рис. 9.1 Виникнення поверхневих станів: а – обрив періодичного потенціалу ґратки поблизу поверхні кристала; б – поверхневі стани, які виникають в результаті обриву ґратки; в – атоми М на поверхні; г – типи поверхневих станів (E_D – донорні, E_A – акцепторні, E_P – рекомбінаційні).

9.1. Поверхневі стани

Реальні кристали мають скінченні розміри, тобто, існує деяка поверхня розділу кристал – зовнішнє середовище, на якій відбувається розрив кристалічної ґратки періодичності її потенціалу (рис. 9.1, а). Тому енергетичні спектри обмеженого та необмеженого кристалів істотно відрізняються. В останньому виникають нові дискретні дозволені рівні (стани), розміщені в зоні заборонених значень енергій для необмеженого кристала, які одержали назву рівнів Тамма (поверхневі рівні або поверхневі стани, рис. 9.1, б). Відповідні їм хвильові функції електронів мають максимум на границі кристала і швидко затухають у міру віддалення від поверхні як вглиб кристала, так і в бік зовнішнього середовища (вакууму). Це значить, що електрони, які займають рівні Тамма E_T не можуть проникати вглиб кристала – вони локалізуються поблизу поверхні. густина таммівських станів визначається Густиною поверхневих атомів, тобто становить $10^{14} - 10^{15} \text{ см}^{-2}$. Аналогічні поверхневі стани виникають внаслідок адсорбції атомів інших речовин.

Наочне уявлення про виникнення поверхневих станів можна одержати, розглядаючи зв'язки, що діють між атомами в об'ємі та на поверхні кристала. На рис. 9.2. зображено плоску модель ґратки германію. Атом в об'ємі кристала оточений чотирма найближчими сусідами, зв'язок з якими здійснюється попарним узагальненням валентних електронів. У атомів, розміщених на вільній поверхні АА, один валентний зв'язок виявляється розірваним, а електронна пара недокомплектованою. Прагнучи вкомплектувати цю пару та заповнити свою зовнішню оболонку до стійкої восьмиелектронної конфігурації, поверхневі атоми поводять себе як типові акцептори, яким у забороненій зоні відповідають акцепторні рівні E_A (рис. 9.1, б). Електрони, що попали на ці рівні з валентної зони, не проникають вглиб кристала і локалізуються на відстані близько сталої ґратки від

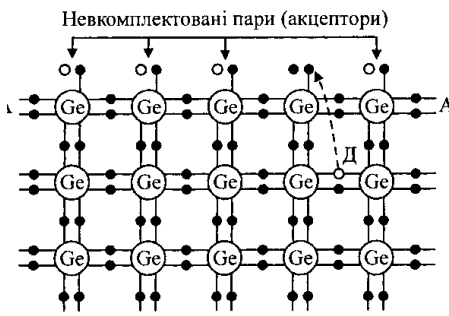


Рис. 9.2. Схема виникнення акцепторних поверхневих станів на чистій поверхні напівпровідника.

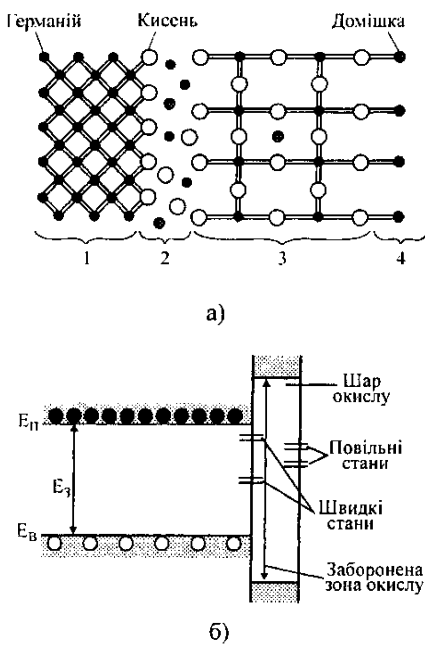


Рис. 9.3. Структура поверхневого шару германію (а) і розміщення швидких і повільних поверхневих станів (згин зон поблизу поверхні не показано) (б): 1- германій; 2 – перехідний шар; 3 – окисний шар; 4 – адсорбовані домішки.

поверхні. У валентній зоні виникають при цьому дірки, а в поверхневому шарі напівпровідника – діркова провідність.

Розглянуті поверхневі стани виникають на ідеально чистій бездефектній поверхні, одержати яку практично неможливо. В реальних умовах поверхневі стани напівпровідників визначаються поверхневими станами, створеними, в основному, сторонніми атомами (молекулами) на поверхні. На рис. 9.1, в показана зонна структура напівпровідника. Вертикальною прямою ВС позначається одна з вільних його поверхонь. Припустимо, що на цій поверхні хімічно сорбувалася частинка М. При такій сорбції хвильові функції ґратки та частинки перекриваються настільки, що частинку можна розглядати як домішку, яка локально порушує періодичність потенціалу ґратки та призводить до виникнення в забороненій зоні поверхневого рівня.

Характер таких рівнів залежить від природи поверхні та частинок. Вони можуть бути акцепторними, донорними та рекомбінаційними (рис. 9.1, г). Так, кисень, сорбований на поверхні германію, створює акцепторні рівні, вода - донорні. Якщо рівні Р акцепторні, то вони захоплюють електрони і заряджають поверхню напівпровідника негативно з поверхневою густиною $S^- = qNf_{\phi-D}^-$, де N – кількість молекул, адсорбованих одиницею поверхні кристала; $f_{\phi-D}^-$ – функція Фермі-Дірака, що виражає ймовірність заповнення поверхневих рівнів електронами; q – заряд електрона. Якщо рівні Р донорні, то вони, віддаючи електрони кристалу, заряджають поверхню напівпровідника позитивно з густиною $S^+ = qNf_{\phi-D}^+$, де $Nf_{\phi-D}^+$ – ймовірність того, що поверхневі рівні порожні, тобто частинки М іонізовані.

При високій густині поверхневих станів вони, взаємодіючи один з одним, можуть розмитися в поверхневу зону. Електрони в цій зоні можуть рухатися лише вздовж поверхні.

Поверхневі стани, обумовлені дефектами поверхні та адсорбцією на ній сторонніх частинок, розміщуються безпосередньо поблизу поверхні та мають хороший контакт з об'ємом напівпровідника. Тому час встановлення рівноваги цих станів з об'ємом напівпровідника (час переходу електронів з енергетичних зон на поверхневі рівні або назад) виявляється дуже невеликим ($\sim 10^{-7}$ с). Такі стани називаються швидкими поверхневими станами. Вони мають густину близько 10^{15} м^{-2} , яка залежить від характеру обробки поверхні та можуть характеризуватися великими перерізами захоплення як для електронів, так і для дірок, внаслідок чого можуть слугувати ефективними центрами рекомбінації.

За нормальних умов поверхня напівпровідника

вкрита шаром окису товщиною, по меншій мірі, в десятки ангстрем (рис. 9.3, а). На зовнішній поверхні окису сорбуються домішкові атоми, які створюють зовнішні або повільні поверхневі стани (рис. 9.3, б). Час встановлення рівноваги таких станів з об'ємом напівпровідника значно більший, ніж для швидких станів і коливається від мікросекунд до хвилин, годин і навіть діб, що зумовлюється дуже низькою ймовірністю проходження електронів крізь окисний шар, який є ізолятором.

Повільні поверхневі стани можуть створюватися також атомами домішки в самому окисі. Густина повільних станів значно більша, ніж швидких ($10^{17}-10^{19} \text{ м}^{-2}$) і значно залежить від їх природи та стану навколишнього середовища.

9.2. Приповерхневий просторовий заряд

Зарядження поверхні напівпровідника при заповненні поверхневих станів супроводжується виникненням біля поверхні шару об'ємного заряду, нейтралізуючого поверхневий заряд. Нейтралізація відбувається притяганням до поверхні носіїв зі знаком заряду, протилежним знаку заряду поверхні, та відштовхуванням носіїв одного знаку. Тому поверхневий шар напівпровідника стає збідненим носіями одного знаку зі знаком поверхневого заряду і збагаченим носіями протилежного знаку.

Товщина шару об'ємного заряду, вздовж якої відбувається нейтралізація поверхневого заряду, залежить від концентрації носіїв. В той час як у металах, що мають дуже високу концентрацію носіїв ($\sim 10^{28} \text{ м}^{-3}$), нейтралізація відбувається вже на відстані кількох параметрів ґратки, в напівпровідниках область просторового заряду може простягатися на значну глибину ($\sim 10^{-6} \text{ м}$ і більше). Звичайно її приймають рівною так званій дебаєвській довжині екранування L_D , яка має такий фізичний зміст: L_D є відстанню, протягом якої потенціал поля у речовині з вільними носіями зменшується в e разів. Розрахунок показує, що при невеликому відносному збагаченні приповерхневої області напівпровідника вільними носіями заряду:

$$L_D = \left(\frac{e_0 e j_T}{q^2 n_0} \right)^{1/2} \quad (9.1)$$

де n_0 – концентрація носіїв заряду в об'ємі; j_T – температурний потенціал. Фізичний зміст цієї величини полягає в тому, що в приповерхневому шарі товщиною L_D змінюються поле, потенціал і концентрація носіїв. Просторовий заряд, який екранує зовнішнє поле, обмежується відстанню L_D від поверхні і саме тут відновлюється електронейтральність.

Таким чином, поблизу поверхні напівпровідника існує область, електричні властивості якої визначаються не об'ємними концентраціями домішки, а величиною

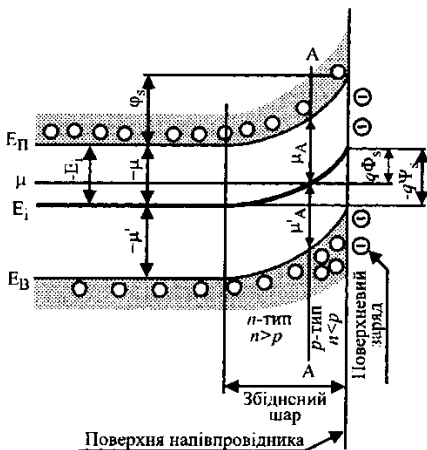


Рис. 9.4. Вигин зон поблизу поверхні напівпровідника n-типу при негативному зарядженні поверхні.

поверхневого заряду. В цій області концентрація носіїв може істотно відрізнятись від об'ємної концентрації. Наявність такої області має суттєвий вплив на багато властивостей напівпровідника: електропровідність, роботу виходу, фото-ЕРС тощо, а також на параметри приладів.

Зарядження поверхні напівпровідника викликає виникнення різниці потенціалів між поверхнею і об'ємом напівпровідника і, отже, викривлення енергетичних зон. При зарядженні поверхні негативно енергетичні зони вигинаються вгору, оскільки при переміщенні електрона з об'єму на поверхню його потенціальна енергія збільшується (рис. 9.4). У випадку позитивного заряду поверхні зони вигинаються вниз. Вигин простягається вглиб напівпровідника приблизно на дебаївську довжину екранування L_D .

На рис. 9.4 показано зонну структуру n-напівпровідника, поверхня якого заряджена негативно, і наведені позначення основних величин, що характеризують поверхню: U_s - різниця потенціалів між поверхнею та об'ємом напівпровідника; $f_s = -qU_s$ - вигин зон поблизу поверхні; E_I - середина забороненої зони; $q\Phi_s$ - відстань від середини забороненої зони до рівня Фермі на поверхні; Φ_s - називається поверхневим потенціалом. З рис. 9.4 видно, що в об'ємі напівпровідника відстань від дна зони провідності до рівня Фермі ($-\mu$) менша від відстані до рівня Фермі від стелі валентної зони ($-\mu'$). Тому рівноважна концентрація електронів n_0 більша від концентрації дірок: $n_0 > p_0$, як і має бути у випадку n-напівпровідників. У поверхневому шарі об'ємного заряду відбувається викривлення зон, і відстань від дна зони провідності до рівня Фермі у міру переміщення до поверхні неперервно збільшується, а відстань від рівня Фермі до стелі валентної зони неперервно зменшується. В перерізі AA ці відстані стають однаковими ($|\mu| = |\mu'|$) напівпровідник стає власним: $n = p$; Правіше від перерізу AA $|\mu| > 1$, внаслідок чого $p > n$, і напівпровідник стає напівпровідником р-типу. В цьому випадку поблизу поверхні утворюється поверхневий р-п-перехід.

Часто вигин зон поблизу поверхні виражають в одиницях kT/q і позначають через Y_s . Тоді:

$$Y_s = \frac{qU_s}{kT} = -\frac{j_s}{kT}$$

На рис. 9.5 зображено структуру енергетичних зон поблизу поверхні n-напівпровідника, вкритого шаром окису, вказані швидкі та повільні стани і вигин зон у припущенні, що поверхня заряджена негативно.

При формуванні приповерхневої області напівпровідника можуть зустрітись три важливих випадки: збіднення, інверсія та збагачення цієї області носіями заряду. Ці випадки для напівпровідників n- і р-

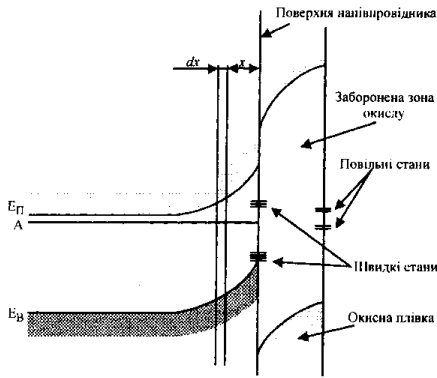
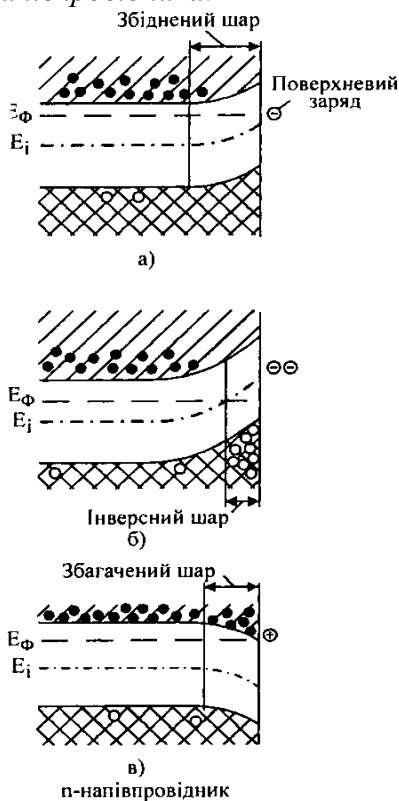


Рис. 9.5. Структура енергетичних зон поблизу поверхні напівпровідника.



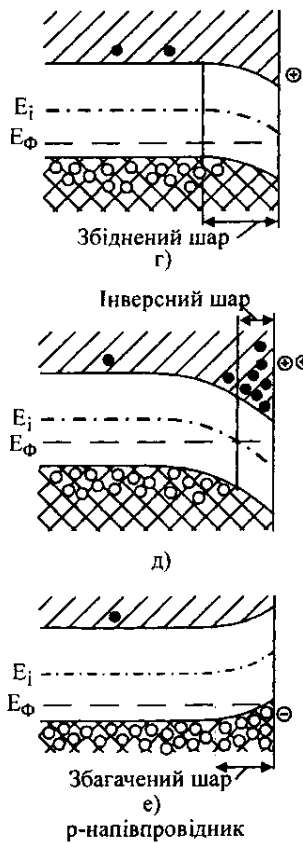


Рис. 9.6. Утворення приповерхневого шару об'ємного заряду під впливом зарядженої поверхні у напівпровідниках n - і p -типу: а, г – збіднений шар; б, д – інверсний; в, е- збагачений.

типу показані на рис. 9.6.

Збіднена область проявляється в тому випадку, коли на поверхні напівпровідника виникає поверхневий заряд, що збігається за знаком зі знаком основних носіїв заряду (рис. 9.6, а, г). Викликаний таким зарядом вигин зон призводить до збільшення відстані від рівня Фермі до дна зони провідності в напівпровіднику n -типу і до вершини валентної зони у напівпровіднику p -типу.

Збільшення цієї відстані супроводжується збідненням приповерхневої області основними носіями: їх концентрація виявляється набагато меншою від концентрації некомпенсованої домішки, що визначає тип провідності напівпровідника.

При високій густині поверхневого заряду, який збігається за знаком зі знаком основних носіїв, відстань від рівня Фермі до стелі валентної зони в напівпровіднику p -типу виявляється меншою від віддалі до дна зони провідності, внаслідок чого концентрація неосновних носіїв заряду (дірок) поблизу поверхні напівпровідника перевищує концентрацію основних носіїв і тип провідності в цій області змінюється. Це явище одержало назву інверсії, а шари, в яких воно спостерігається, називаються інверсійними шарами (рис. 9.6, б, д).

Якщо знак поверхневого заряду протилежний знаку заряду основних носіїв струму в напівпровіднику, то під його впливом відбувається притягання до поверхні основних носіїв та збагачення ними приповерхневого шару. Такі шари називаються збагаченими (рис. 9.6, в, г).

Наявність поверхневих станів призводить до того, що поверхня кристала може бути зарядженою. Наприклад, якщо на поверхні виникають донорні рівні, то вони можуть бути або електронейтральними, або зарядженими позитивно (при переході електрона в зону провідності), тобто, при наявності донорних станів поверхня кристала може бути заряджена позитивно. У випадку наявності акцепторних станів поверхня може мати негативний заряд, оскільки вони можуть бути заповнені електронами. Величина заряду поверхні Q_S визначається концентрацією поверхневих рівнів N_S і взаємним розміщенням рівня Фермі та поверхневого рівня.

Відповідно до принципу електронейтральності заряд повинен бути нейтралізованим. Нейтралізація відбувається внаслідок того, що електричне поле, утворене зарядом, викликає перерозподіл рухомих носіїв у приповерхневій області: носії зі знаком, протилежним знаку заряду поверхні, підтягуються до неї з об'єму, носії зі знаком, однойменним заряду поверхні, відштовхуються вглиб об'єму. Очевидно, незалежно від типу провідності матеріалу, енергетичні зони будуть зігнуті вниз, якщо поверхня буде заряджена позитивно. Якщо ж поверхня заряджена негативно, то зони поблизу неї будуть вигнуті вгору (рис. 9.2). Звичайно вигин зон вимірюється різницею

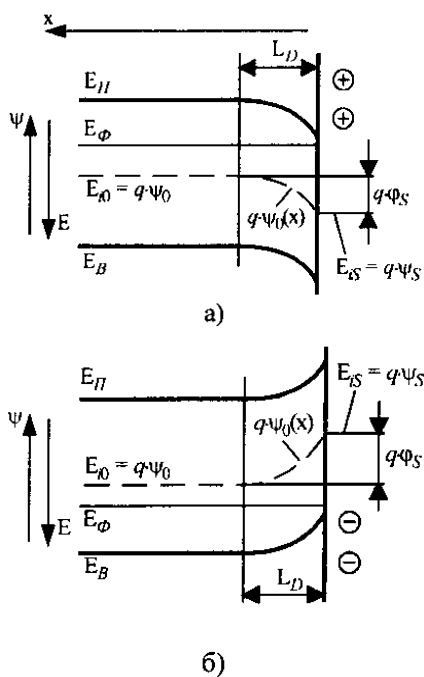


Рис. 9.7. Додатний (а) та від'ємний (б) вигин зон.

енергій середини забороненої зони (рівнів електростатичних потенціалів) у приповерхневій області та в глибині об'єму. Він вважається додатним, якщо зони вигнуті вниз, та від'ємним, якщо вони вигнуті вгору.

При малій густині від'ємних поверхневих станів на напівпровіднику n-типу утворюється збіднений шар (рис. 9.6, а), тому що основні носії заряду – електрони – відштовхуються від'ємним поверхневим зарядом у глибину напівпровідника. Електростатичне поле поверхневих зарядів проникає в напівпровідник на певну глибину, яка залежить від питомого опору цього напівпровідника.

При високій густині від'ємних поверхневих станів поблизу поверхні напівпровідника n-типу утворюється шар з протилежним типом електропровідності – інверсний шар. Границя інверсного шару в глибині напівпровідника розміщена там, де рівень Фермі перетинає середину забороненої зони. Інверсним шаром у напівпровіднику є збіднений шар.

Якщо на поверхні напівпровідника n-типу переважають додатні поверхневі стани, то поверхня напівпровідника збагачується електронами – основними носіями заряду, тобто утворюється збагачений шар. Границя збагаченого шару в глибині напівпровідника знаходиться там, де починається вигин енергетичних рівнів, тобто визначається глибиною проникнення електростатичного поля поверхневих зарядів. На поверхні напівпровідника р-типу при інших знаках поверхневих зарядів можуть утворюватися аналогічні поверхневі шари (рис. 9.7).

На реальному напівпровіднику завжди є шар оксиду. Тому поверхневі стани можуть знаходитися не тільки безпосередньо на напівпровіднику, але також у шарі оксиду і на його поверхні. При зміні зовнішнього електричного поля та при відповідній зміні енергетичної діаграми поблизу поверхні напівпровідника повинне відбуватися заповнення або спустошення електронами деяких поверхневих станів. Поверхневі стани, розміщені поблизу границі розділу напівпровідник – оксид, заповнюються або спустошуються відносно швидко, оскільки в напівпровіднику є достатня кількість вільних електронів у зоні провідності та дірок (порожніх енергетичних рівнів) у валентній зоні. Внаслідок малого часу релаксації таких поверхневих станів їх називають швидкими поверхневими станами.

В шарі оксиду концентрація носіїв заряду мізерно мала. Тому зі зміною зовнішнього електричного поля поверхневі стани, розміщені в шарі оксиду або на його поверхні, перезаряджаються повільно (час їхньої релаксації великий). Такі поверхневі стани називають повільними поверхневими станами.

Деякі поверхневі стани створюють поблизу середини

забороненої зони енергетичні рівні, які називаються рівнями рекомбінаційних пасток. Рекомбінаційними пастками можуть бути тільки швидкі поверхневі стани, тому що час переходу носіїв на повільні поверхневі стани досить великий. Явище поверхневої рекомбінації прийнято характеризувати швидкістю поверхневої рекомбінації носіїв заряду, яка визначається відношенням густини потоку носіїв заряду біля поверхні напівпровідника до надлишкової концентрації цих носіїв біля поверхні.

Таким чином, при виготовленні напівпровідникових приладів необхідно вибирати метод обробки поверхні кристалів напівпровідника, при якому швидкість поверхневої рекомбінації мінімальна.

Отже, поблизу поверхні виникає подвійний електричний шар, що екранує об'єм напівпровідника від дії поля. Цей шар виявляється збідненим носіями заряду, які мають той самий знак, що й заряд поверхні, та збагачений носіями протилежного знака. Іншими словами, поблизу поверхні існує область, електричні властивості якої визначаються не концентрацією домішки в об'ємі, а величиною поверхневого заряду.

Однак товщина цієї області визначається об'ємними характеристиками матеріалу. Звичайно її приймають рівною дебаєвській довжині екранування L_D .

Оскільки L_D визначається концентрацією носіїв у об'ємі матеріалу, то для металів з їх великою концентрацією носіїв ($n_0 \approx 10^{23} \text{ см}^{-3}$) довжина екранування дуже мала ($L_D \approx 10^{-8} \text{ см} = 0,1 \text{ нм}$) – менше міжатомної відстані. Тому просторовий заряд практично зосереджується на поверхні. У напівпровідниках концентрація носіїв значно менша, крім того, вона може змінюватися в широких межах, тому для них довжина екранування набагато більша, ніж для металів. Наприклад, для кремнію при $n_0 = n_i \approx 10^{10} \text{ см}^{-3}$, $L_D = 20 \text{ мкм}$, для власного германію $L_D = 1,4 \text{ мкм}$.

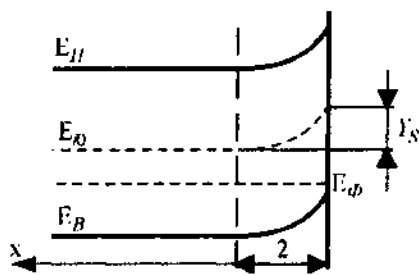
Вигин зон у приповерхневому шарі призводить до того, що в напрямку, нормальному до поверхні, концентрація рухомих носіїв залежить від координати (на відстані від поверхні, рівній L_D). Це значить, що поблизу поверхні можуть виникнути області з провідністю, яка значно відрізняється від провідності об'єму.

Оцінимо зміну концентрації носіїв у приповерхневому шарі. Для невідроджених напівпровідників відповідно до розподілу Больцмана:

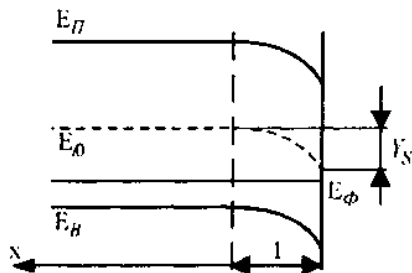
$$n = n_0 \exp\left(\frac{q\phi}{kT}\right); \quad p = p_0 \exp\left(-\frac{q\phi}{kT}\right), \quad (9.2)$$

де n_0, p_0 – рівноважні концентрації носіїв у об'ємі.

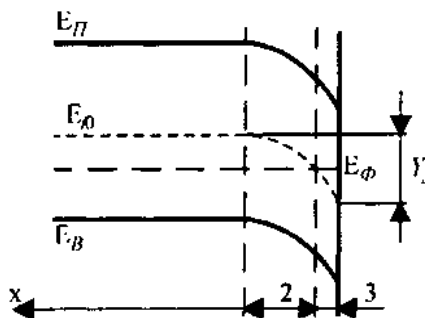
З наведених співвідношень видно, що при додатному вигині зон у приповерхневій області збільшується концентрація електронів, а при від'ємному – концентрація дірок.



а)



б)



в)

Рис. 9.8. Збагачення (а), збіднення (б), інверсія (в) приповерхневого шару напівпровідника р-типу.

Густина об'ємного заряду в приповерхневому шарі можна знайти з виразу:

$$r(x) = -q(n - p - N_A - N_D) \quad (9.3)$$

якщо припустити, що домішки N_A і N_D повністю іонізовані. Оскільки об'єм матеріалу електронейтральний, то виконується умова електронейтральності $n_0 - p_0 - N_A - N_D$. Тому (9.3) з урахуванням (9.2) можна записати у вигляді:

$$r(x) = qn_0 \left(1 - \exp \frac{qj}{kT} \right) - qp_0 \left(1 - \exp \frac{qj}{kT} \right). \quad (9.4)$$

Повний заряд у приповерхневій області можна знайти зі співвідношення:

$$Q_{s1} = \int_0^{\infty} r(x) dx = -\frac{e_0 e k T}{q} \int_0^{\infty} \frac{d^2 Y}{dx^2} dx = \frac{e_0 e k T}{q} \left(\frac{dY}{dx} \right)_{x=0}$$

Одержимо:

$$Q_{s1} = 2qn_i L_D F(Y_s, 1) \quad (9.5)$$

де $Y_s = qj_s / kT$, як це видно з рис. 9.3.

Одержане співвідношення свідчить, що величина заряду у приповерхневій області визначається значенням поверхневого потенціалу Y_s та мірою легування напівпровідника А. Відповідно до цього розрізняють три важливих випадки: збагачення, збіднення та інверсія приповерхневого шару (рис. 9.8).

При збагаченні знак поверхневого заряду протилежний знаку основних носіїв заряду: $Y_s > 0$ для напівпровідників n-типу, $Y_s < 0$ для напівпровідників р-типу. В першому випадку $Q_{s1} > 0$, в другому $-Q_{s1} > 0$. Концентрація основних носіїв у приповерхневому шарі перевищує їх концентрацію в об'ємі (рис. 9.8, а).

Збіднена область виникає в тому випадку, коли заряд поверхні за знаком збігається зі знаком заряду основних носіїв. В цьому випадку $Y_s < 0$ для напівпровідників n-типу, для напівпровідників р-типу, навпаки, $Y_s > 0$ (рис. 9.8. б).

Якщо поверхневий заряд, який збігається за знаком із зарядом основних носіїв, досить великий, а викликане ним викривлення рівнів таке, що рівень електростатичного потенціалу перетинає рівень Фермі, то поблизу поверхні утворюється інверсійний шар (рис. 9.8, в). При цьому поблизу поверхні концентрація неосновних носіїв перевищує концентрацію основних носіїв, отже, тип провідності у приповерхневому шарі змінюється.

9.3. Поверхнева провідність і ефект поля

Зміна концентрації рухомих носіїв заряду в приповерхневому шарі змінює його провідність. Якщо припустити, що рухомість носіїв поблизу поверхні не відрізняється від рухомості в глибині об'єму, то поверхневу провідність можна записати як:

$$S_s = q(\Delta n m_n + \Delta p m_p) = q m_p (\Delta p + b \Delta n) \quad (9.6)$$

де $b = m_n / m_p$, а Δn і Δp – зміна концентрації носіїв поблизу поверхні.

Матимемо:

$$\begin{aligned} \Delta n &= \int_0^{\infty} [n(x) - n_0] dx = n_0 \int_0^{\infty} (e^Y - 1) dx = \\ &= n_0 \int_0^{Y_s} (e^Y - 1) \left(\frac{dY}{dx} \right) dY = n_0 L_D \int_0^{Y_s} \frac{e^Y - 1}{F(Y, I)} \left(\frac{dY}{dx} \right) dY \end{aligned}$$

Аналогічно

$$\Delta p = p_0 L_D \int_0^{Y_s} \frac{e^{-Y} - 1}{F(Y, I)} dY$$

Підставляючи вирази для Δn і Δp в (9.6), одержуємо для поверхневої провідності вираз:

$$S_s = q m_p n_i L_D g \quad (9.7)$$

де $g = 0,5 \int_0^{\infty} [I(e^{-Y} - 1) + b(e^Y - 1)I^{-1}] [F(Y, I)]^{-1} dY -$

табульована для різних значень I .

Поверхнева провідність є досить складною функцією поверхневого електростатичного потенціалу та ступеня легування матеріалу. Якісно оцінити поверхневу провідність можна так.

У випадку додатного вигину зон поверхнева провідність напівпровідника n-типу підвищується внаслідок утворення біля поверхні збагаченого шару. При досить великих від'ємних вигинах зон внаслідок утворення поблизу поверхні інверсійного шару поверхнева провідність напівпровідника n-типу також велика. Мінімального значення вона досягає, коли існує незначний від'ємний вигин зон, при якому поверхнева провідність визначається провідністю збідненого шару.

У напівпровіднику p-типу, навпаки, поверхнева провідність буде зростати внаслідок утворення збагаченого шару при від'ємних вигинах зон та інверсійного шару, що виникає при великих додатних вигинах зон. Мінімальне значення вона матиме в цьому випадку при незначному додатному вигині зон, бо, як і в попередньому випадку, буде визначатися провідністю збідненого шару.

Сказане можна пояснити за допомогою рис. 9.9, відносно якого слід зробити деякі зауваження. По-перше, мінімальне значення провідності досягається при $Y_{s \min} = \ln I^2 / b$. Це значить, що положення мінімуму на кривих $S_s(Y_s)$ визначається лише об'ємними властивостями напівпровідника: зі збільшенням ступеня легування мінімальне значення поверхневої провідності зміщується в бік зростання абсолютного значення Y_s . По-друге, у власних напівпровідниках мінімум поверхневої

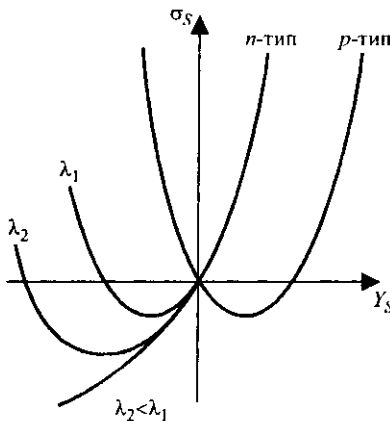


Рис. 9.9. Залежність провідності приповерхневого шару від вигину зон.

провідності зміщений у бік від'ємних значень вигинів зон, бо рухомість електронів більша від рухомості дірок. Нарешті, очевидно, що при дуже великому вигині зон рівень Фермі біля поверхні може опинитися в одній з дозволених зон – біля поверхні виникає шар напівпровідника виродженого типу. Поверхнева провідність у цьому випадку буде особливо великою.

9.4. Поверхнева рекомбінація

На поверхні напівпровідника може виникати велика кількість станів. Деякі стани, а саме, швидкі, можуть відігравати роль рівнів прилипання та рекомбінаційних пасток аналогічно до об'ємних рекомбінаційних пасток, енергетичні рівні яких розмішуються поблизу середини забороненої зони. Наявність поверхневих центрів рекомбінації визначає можливість рекомбінації носіїв заряду на поверхні. Оскільки на поверхні завжди існують додаткові рівні, то швидкість поверхневої рекомбінації може виявитися значно вищою, ніж об'ємної.

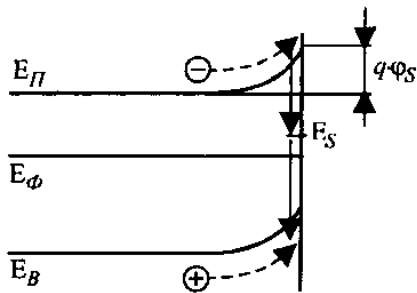


Рис. 9.10. Поверхнева рекомбінація.

Роль поверхневої рекомбінації особливо помітна в тонких зразках, що мають велике відношення поверхні до об'єму, а також у чистих кристалах, де концентрація рекомбінаційних пасток мала. Очевидно, повільні стани не можуть слугувати рекомбінаційними центрами, оскільки час переходу на них настільки великий, що неможливо забезпечити рекомбінацію, порівняну за темпом з об'ємною.

Припустимо, що на поверхні існує деяка концентрація дефектів, яка характеризується рекомбінаційною пасткою E_s (рис. 9.10), через яку здійснюється рекомбінація. Цей процес відбувається так. В об'ємі генеруються нерівноважні носії Δn та Δp . Наявність на поверхні рекомбінаційних центрів призводить до виникнення напрямлених потоків носіїв до поверхні f_n і f_p . У стані термодинамічної рівноваги (струм через поверхню відсутній) потоки електронів і дірок однакові й визначаються кількістю носіїв, що рекомбінують за одиницю часу на одиниці площі:

$$f_n = f_p = f_s = s_n \Delta n = s_p \Delta p$$

Коефіцієнт пропорційності s – відношення кількості актів рекомбінації за одиницю часу на одиниці площі до концентрації надлишкових носіїв – має розмірність швидкості і називається швидкістю поверхневої рекомбінації.

Слід відзначити, що потік електронів визначає електронну складову струму в напрямку, нормальному до поверхні, тобто $j_{ns} = q \Delta n s_n$. При цьому за додатний напрямок струму приймається напрямок від поверхні вглиб напівпровідника, який відповідає руху електронів до поверхні.

Швидкість поверхневої рекомбінації значною мірою залежить від величини поверхневого потенціалу (вигину зон). Ця залежність має дзвоноподібний вигляд і

якісно її можна пояснити так. При великих від'ємних поверхневих потенціалах (вигин зон вгору) в зоні провідності концентрація електронів мізерно мала – швидкість рекомбінації низька. При великих позитивних потенціалах (вигин зон вниз) мала концентрація дірок у валентній зоні – швидкість рекомбінації також невисока. Найсприятливіші умови створюються для рекомбінації в області Y_s , близькій до нуля. Тут швидкість рекомбінації максимальна.

Швидкість поверхневої рекомбінації дуже сильно залежить від стану поверхні. Різні адсорбційні процеси на поверхні можуть значно змінити поверхневий потенціал і, отже, швидкість поверхневої рекомбінації. Крім того, механічна обробка поверхні порушує кристалічну структуру поверхневого шару. Це призводить до утворення великої кількості поверхневих рекомбінаційних центрів, що може різко збільшити швидкість поверхневої рекомбінації (наприклад, для германію зі шліфованою поверхнею $s = 10^5$ — 10^6 см/с). Травлення поверхні в спеціальних розчинах знижує швидкість поверхневої рекомбінації до 50-100 см/с, однак окислення поверхні та адсорбція газів збільшують її, і вона через деякий час досягає звичайного значення (порядку 10^3 см/с).

Необхідно зазначити, що поверхнева рекомбінація може істотно змінювати час життя носіїв. Так, сумарний час життя, який визначається процесами в об'ємі τ_v і на поверхні τ_s можна подати у вигляді суми:

$$\frac{1}{t} = \frac{1}{t_v} + \frac{1}{t_s}$$

де $t_s = d / 2s$ (d – товщина зразка). Як видно, при малій товщині час життя буде визначатися в основному значенням τ_s .

Лекція 10

КОНТАКТНІ ЯВИЩА В НАПІВПРОВІДНИКАХ .

План

- 10.1 Контакт напівпровідник-напівпровідник
- 10.2 Нерівноважний стан переходу
- 10.3. Вольт-амперна характеристика і ємність переходу
- 10.4. Неінжектуючі контакти, тунельні переходи і гетеропереходи
 - 10.4.1. Неінжектуючі контакти
 - 10.4.2. Тунельні переходи
 - 10.4.3. Гетеро переходи
- 10.5. Контакти
 - 10.5.1. Контакт метал-напівпровідник
 - 10.5.2. Випрямний контакт
 - 10.5.3. Омичний контакт
 - 10.5.4. Контакт метал-метал

Література

- 10.1. Штернов А. А. Физические основы конструирования, технология РЗА и микроэлектроники. – М: Радио и связь, 1981. – 248 с.
- 10.2. Киреев И. П. Физика полупроводников. – М.: Высшая школа, 1975. – 247 с.

10.1 Контакт напівпровідник-напівпровідник

Контакти двох напівпровідників широко використовуються в електроніці для створення різноманітних активних і пасивних елементів та для їх ізоляції. Процеси, що відбуваються в таких контактах, часто визначають параметри діодів і транзисторів. Розрізняють контакти на межі розділу областей з різним типом провідності в об'ємі одного напівпровідника і контакти на межі розділу двох різнорідних напівпровідників. Контакти першого типу називають гомопереходами, контакти другого типу – гетеропереходами. Зазвичай замість терміну "гомоперехід" використовують термін "р-п-перехід" або "електронно-дірковий перехід".

Електронно-дірковим, або р-п-переходом називають поверхню розділу між областями р- і п-типу, які одночасно існують у монокристалі. Практично це деякий перехідний шар, що розділяє області напівпровідників з різним типом провідності.

Якщо концентрація основних носіїв у суміжних шарах приблизно однакова ($p_{p0} \approx n_{n0}$), то перехід називається симетричним. Якщо ж $p_{p0} \gg n_{n0}$ або $n_{n0} \gg p_{p0}$, то перехід називається несиметричним. Якщо концентрації основних носіїв відрізняються на кілька порядків, говорять про односторонній перехід. Такі переходи позначають символами p^+-n (або n^+-p).

У напівпровіднику р-типу існує деяка концентрація

акцепторів N_A , а також основних p_{p0} і неосновних n_{n0} носіїв. Іонізовані атоми акцепторів та основної речовини створюють нерухомий негативний заряд, вільні електрони і дірки – відповідно рухомий негативний та позитивний заряди. При цьому $p_{p0} \approx N_A + n_{p0}$. Аналогічно напівпровідник n-типу характеризується концентрацією донорів N_D основних n_{n0} і неосновних p_{p0} носіїв. В ньому також є позитивний і негативний об'ємні заряди, які компенсують один одного: $n_{n0} \approx N_D + p_{n0}$. Для визначеності будемо вважати, що $N_D \gg N_A$. Це значить, що шар n – низькоомний, а шар p – високоомний.

Оскільки концентрація електронів у n-шарі значно перевищує їх концентрацію в p-шарі, частина електронів дифундує з шару n у шар p , внаслідок чого в n-шарі виникають некомпенсовані позитивні заряди іонів донорів.

З іншого боку, в p-шарі поблизу межі містяться надлишкові електрони, які будуть рекомбінувати з дірками доти, доки не буде виконуватися умова рівноваги. Отже, і в цій області зменшиться концентрація рухомих носіїв і також виникнуть некомпенсовані негативні заряди акцепторів. Необхідно відзначити, що дірки будуть дифундувати з шару p у шар n , що також призводить до деякого зниження концентрації рухомих носіїв у областях, прилеглих до межі розділу, та утворення додаткових некомпенсованих просторових зарядів. Однак у несиметричному переході дифузія з високоомного шару не має великого значення, оскільки градієнти концентрацій та дифузійні потоки визначаються різницями $n_{n0} - n_{p0}$ і $p_{p0} - p_{n0}$. У розглянутому випадку очевидно, що $n_{n0} - n_{p0} \gg p_{p0} - p_{n0}$.

Таким чином, у результаті дифузії носіїв у деяких областях по обидва боки межі розділу утворюються просторові некомпенсовані заряди іонів домішки. Ця область з шириною l_0 і є p-n-переходом. Оскільки область переходу збіднена основними носіями заряду, її опір буде набагато більшим від опору електронейтральних p- і n-шарів. Тому часто p-n-перехід називають збідненим, або виснаженим шаром. Якщо припустити, що в p-n-переході немає рухомих носіїв заряду, то опір його виявиться нескінченно великим. Такі переходи називають ідеальними.

Необхідно відзначити, що перехід загалом електронейтральний, бо позитивний та негативний заряди в суміжних шарах однакові. Це свідчить про те, що відмінність між концентрацією домішок у суміжних областях спричиняє різницю в ширині областей, зайнятих просторовими зарядами: в шарі з меншою концентрацією домішки ширина області просторового заряду більша. Іншими словами, несиметричний перехід розташований, в основному, у високоомному шарі.

Некомпенсовані заряди іонів домішок зумовлюють

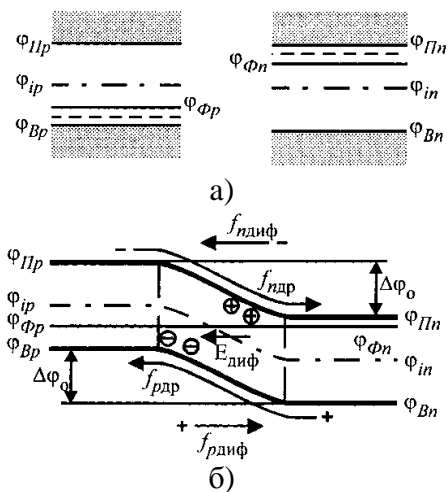


Рис. 10.1. Зонні діаграми напівпровідника р- і n-типу (а), утворення потенціального бар'єра (б).

виникнення електричного поля, скерованого від позитивних зарядів до негативних, тобто з шару n у шар р. Це поле перешкоджає подальшій дифузії: в рівноважному стані дифузійні потоки, викликані наявністю градієнта концентрації, будуть зрівноважуватися зустрічними (дрейфовими) потоками носіїв того ж знака, викликаними електричним полем переходу. Одночасно з виникненням електричного поля виникає різниця потенціалів між суміжними шарами, яка називається потенціальним бар'єром.

Розглянемо зонну діаграму переходу. Для зручності перейдемо від енергії (в електрон-вольтах) до потенціалу (у вольтах). Для переходу від даної розмірності до іншої досить розділити відповідні значення енергії на величину елементарного заряду q . За відсутності контакту р- і n-області характеризуються діаграмою, наведеною на рис. 10.1, а. Тут φ_B – потенціал стелі валентної зони, φ_{Π} – потенціал дна зони провідності, ($\varphi_D, \varphi_A, \varphi_F, \varphi_i$ – відповідно потенціали донорів, акцепторів, Фермі та електростатичний потенціал. Позначимо $kT/q = \varphi_T$.

При наявності контакту рівні Фермі в суміжних шарах повинні бути єдиними, тому відбувається викривлення зон, що призводить до різниці електростатичних потенціалів та утворення потенційного бар'єра (рис. 10.1, б). Висота бар'єра:

$$\Delta j = j_{ip} - j_{in} \quad (10.1)$$

Потенціальний бар'єр створює різні умови для переходу носіїв у суміжні шари. Наприклад, електрон з шару n може перейти в шар р лише в тому випадку, коли він має енергію, достатню для подолання "сходинки" висотою $\Delta\varphi_0$, тобто, коли він може подолати сили електричного поля, що виштовхують його з переходу назад у n-шар. В той же час перехід електронів з р-шару в n-шар здійснюється безперешкодно, більше того, електричне поле переходу допомагає їм (електрони неначе "скочуються" з р-шару). У рівноважному стані ці потоки носіїв взаємно зрівноважуються.

Аналогічна ситуація складається у валентній зоні. Дірки для переходу з шару р в шар n повинні "опуститися" на глибину $\Delta\varphi_0$. Оскільки рух дірок викликаний протилежним переміщенням електронів, це означає, що дірки також повинні мати відповідну енергію для подолання бар'єра висотою $\Delta\varphi_0$ під час переходу з р-шару в n-шар. Іншими словами, дірки, які переходять з р- в n-шар, повинні характеризуватися енергією, більшою від енергії діючого в переході електричного поля. Зворотний рух дірок (з n-шару в р-шар) здійснюється безперешкодно.

Таким чином, у рівноважному стані в переході протікають чотири складові струму (рис. 10.2, б). Дві з них протікають у зоні провідності: дрейфовий потік електронів (скерований з р-шару в n-шар) та зрівноважуючий його дифузійний потік електронів (скерований з n-шару в р-

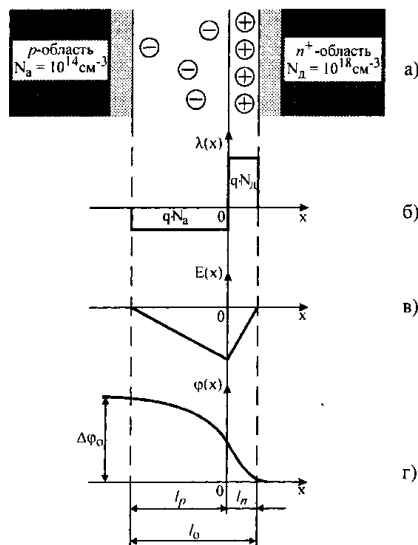


Рис. 10.2. Розподіл домішок і зарядів (а), густини об'ємного заряду (б), поля (в) і потенціалу (г) в ідеальному переході.

шар). Дві інші протікають у валентній зоні: дрейфовий потік дірок напрямлений з n-шару в р-шар, а зрівноважуючий його дифузійний потік дірок – з р-шару в n-шар. Тому густина струму, що проходить через перехід у рівноважному стані, дорівнює нулю: $j_{\Sigma} = j_{n\text{др.}} + j_{n\text{диф}} + j_{p\text{диф}} + j_{p\text{др.}} = 0$

Оцінимо висоту потенційного бар'єра. Оскільки кожний шар перебуває в рівноважному стані, повинна виконуватись умова рівноваги. Наприклад, для n-шару $n_{n0}p_{n0} = n_i^2$. Тоді, можна записати:

$$n_{n0} = n_i \frac{n_i}{p_{n0}} = n_i \exp \frac{j_{\Phi n} - j_{in}}{j_T},$$

звідки знаходимо значення електростатичного потенціалу для n-шару:

$$j_{in} = j_{\Phi n} - j_T \ln \left(\frac{n_{n0}}{n_i} \right).$$

Аналогічно значення електростатичного потенціалу для р-шару:

$$j_{ip} = j_{\Phi n} + j_T \ln \left(\frac{p_{p0}}{n_i} \right).$$

Враховуючи (10.1), одержимо:

$$\Delta j_0 = j_T \ln \frac{p_{p0} n_{n0}}{n_i^2} = j_T \ln \frac{p_{p0}}{p_{n0}} = j_T \ln \frac{n_{n0}}{n_{p0}}. \quad (10.2)$$

Наприклад, для германієвого переходу ($n_i \approx 2,5 \cdot 10^{13} \text{ см}^{-3}$) при $N_D = 10^{15} \text{ см}^{-3}$ і $N_A = 10^3 \cdot N_D$, припускаючи $n_{n0} \approx N_D$ і $p_{p0} \approx N_A$, одержуємо $\Delta \varphi_0 \approx 0,4 \text{ В}$. Для кремнієвого переходу ($n_i \approx 1,6 \cdot 10^{10} \text{ см}^{-3}$) при $N_D = 4 \cdot 10^{14} \text{ см}^{-3}$ і $N_A = 10^{16} \text{ см}^{-3}$ одержуємо $\Delta \varphi_0 \approx 0,6 \text{ В}$.

Відзначимо, що вираз (10.2) дозволяє оцінити співвідношення однойменних носіїв на межах переходу:

$$p_{n0} = p_{p0} \exp \left(-\frac{\Delta j_0}{j_T} \right), \quad n_{p0} = n_{n0} \exp \left(-\frac{\Delta j_0}{j_T} \right) \quad (10.3)$$

Оцінимо тепер ширину переходу, тобто протяжність області некомпенсованих зарядів у суміжних шарах. Найпростіше це зробити для ступінчастого переходу. Розподіл домішок і зарядів у цьому випадку буде таким, як показано на рис. 10.2, а, а розподіл густини об'ємного заряду $\lambda(x)$ таким, як показано на рис. 10.2, б. Аналітичний вираз для розподілу електричного поля і потенціалу в переході одержимо, розв'язавши рівняння Пуассона, яке зв'язує потенціал поля та густину нерухомого об'ємного заряду, який створює це поле:

$$\frac{d^2 j(x)}{dx^2} = -\frac{l(x)}{e_0 e}, \quad (10.4)$$

де для області $x \geq 0$ необхідно прийняти $\lambda(x) = \lambda_1 = qN_D$, а для області $x \leq 0$, $\lambda(x) = \lambda_2 = qN_A$. Граничні умови будуть такими:

$$j(-l_p) = \Delta j_0, \left(\frac{dj_1}{dx} \right)_{x=-l_p} = 0, j_2(-l_p) = 0, \\ \left(\frac{dj_2}{dx} \right)_{x=-l_n} = 0.$$

Тоді:

$$j_1(x) = \Delta j_0 - \frac{qN_a}{2e_0e} (x+l_p)^2, x \leq 0 \\ j_2(x) = \Delta j_0 - \frac{qN_d}{2e_0e} (x+l_n)^2, x \geq 0 \quad (10.5)$$

Вирази для напруженості поля по обидва боки межі розділу будуть мати вигляд:

$$E_1(x) = \frac{dj_1}{dx} = -\frac{qN_a}{e_0e} (x+l_p) \\ E_2(x) = \frac{dj_2}{dx} = -\frac{qN_d}{e_0e} (x+l_n) \quad (10.6)$$

Оскільки електричне поле не має розриву на межі розділу $E_1(0) = E_2(0)$ (рис. 10.2, в.), то з (10.4) одержуємо:

$$\frac{N_d}{N_a} = \frac{l_p}{l_n} \quad (10.7)$$

Співвідношення (10.7) показує, що симетричний перехід ($N_A = N_D$) має однакову протяжність у суміжних шарах ($l_p = l_n$), а несиметричний перехід міститься, в основному, у високоомному шарі (наприклад, при $N_D > N_A$ маємо $l_p > l_n$).

Оскільки $\lambda_1(0) = \lambda(0)$ (рис. 10.2, г), то, враховуючи, що $l_0 = l_n + l_p$, з (10.5) одержуємо вираз для ширини переходу:

$$l_0 = \left[2e_0e\Delta j_0 \frac{1}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right) \right]^{1/2} \quad (10.8)$$

Для несиметричного переходу цей вираз можна спростити. Припускаючи, наприклад, $N_D \gg N_A$, одержуємо:

$$l_0 \approx \left(\frac{2e_0e\Delta j_0}{qN_A} \right)^{1/2} \approx (2e_0er_p m_p \Delta j_0)^{1/2} \quad (10.9)$$

Так, для розглянутих вище германієвого ($\varepsilon = 16,5$) і кремнієвого ($\varepsilon = 11,7$) переходів маємо відповідно $l_0 \approx 0,85$ мкм і $l_0 \approx 1,4$ мкм. Для симетричного переходу вираз (10.8) також спрощується: $l_0 \approx (4e_0erm\Delta j_0)^{1/2}$.

Слід врахувати деякі обставини. Зі співвідношення видно, що в логарифмічному масштабі криві розподілу концентрації носіїв будуть симетричні відносно рівня n_i (рис. 10.3). В середині переходу існує ділянка з власною провідністю. Це значить, що перехід є найбільш високоомною частиною структури. Точка, що характеризується власною провідністю (фізичний

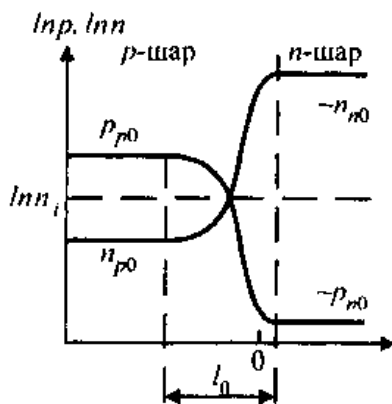


Рис. 10.3. Розподіл концентрації носіїв при рівноважному стані переходу.

перехід), в загальному випадку не збігається з межею розділу р- і n-шарів (металічний перехід) – вона зсунута у високоомний шар. У симетричному переході фізичний та металічний переходи збігаються.

10.2 Нерівноважний стан переходу

При під'єднанні джерела напруги U між шарами р і n прикладене зовнішнє електричне поле порушить рівновагу системи та викличе протікання струму. При цьому висота потенціального бар'єра повинна змінитися. Оскільки опір переходу значно вищий від опору прилеглих шарів, можна вважати, що зовнішня напруга практично повністю падає на переході. Отже, зміна висоти потенційного бар'єра практично дорівнює прикладеній напрузі. Тому, якщо джерело напруги під'єднане плюсом до р-шару, а мінусом до n-шару, висота потенційного бар'єра зменшується. Таке вмикання переходу називається прямим. Якщо ж джерело під'єднане плюсом до n-шару, а мінусом до р-шару, то висота бар'єра збільшується. Таке вмикання переходу називають зворотним.

Таким чином, під'єднання до переходу зовнішньої напруги змінює висоту бар'єра:

$$\Delta j = \Delta j_0 \pm U \quad (10.10)$$

де знак "мінус" стосується прямого, а знак плюс – зворотного вмикання переходу.

Зміна висоти бар'єра призводить до зміни ширини переходу та концентрації носіїв на межі переходу. Підставляючи (10.10) в (10.8), маємо:

$$l = l_0 \left(1 \pm \frac{U}{\Delta j_0} \right) \quad (10.11)$$

де знак "мінус" стоїть у випадку прямого зміщення, знак плюс – при зворотному. З одержаного виразу видно, що перехід звужується у випадку прямого вмикання і розширюється при зворотному. Стосовно виразу (10.11) необхідно зробити таке зауваження. Для зворотного увімкнення він виявляється дуже точним і має широке практичне використання. При $|U| \gg \Delta \varphi_0$ вираз спрощується:

$$l = \left(l_0 \frac{|U|}{\Delta j_0} \right)^{1/2}$$

Для прямого зміщення це зауваження правильне лише якісно (при $U > \Delta \varphi_0$ вираз не має змісту). Це пояснюється тим, що в реальному переході завжди існує деяка концентрація рухомих носіїв заряду. Ідеалізація ж переходу під час виведення співвідношення (10.8) призводить до неточності.

Підставивши тепер у формулу (10.3) замість рівноважного значення висоти бар'єра вираз (10.10), одержимо:

$$p_n = p_{p0} \left(-\frac{\Delta j}{j_T} \right) = \left[p_{p0} \exp \left(-\frac{\Delta j_0}{j_T} \right) \right] \exp \left(\pm \frac{|U|}{\Delta j_T} \right);$$

$$n_p = n_{n0} \left(-\frac{\Delta j}{j_T} \right) = \left[n_{n0} \exp \left(-\frac{\Delta j_0}{j_T} \right) \right] \exp \left(\pm \frac{|U|}{\Delta j_T} \right);$$
(10.12)

де для прямого вмикання потрібно взяти знак "плюс", для зворотного – "мінус". Ці вирази дозволяють оцінити зміну концентрації носіїв заряду на межах переходу в n- і p-шарах. Маємо відповідно:

$$\Delta p_n = p_n - p_{n0} = p_{n0} \left(\exp \frac{U}{j_T} - 1 \right);$$

$$\Delta n_p = n_p - n_{p0} = n_{p0} \left(\exp \frac{U}{j_T} - 1 \right).$$
(10.13)

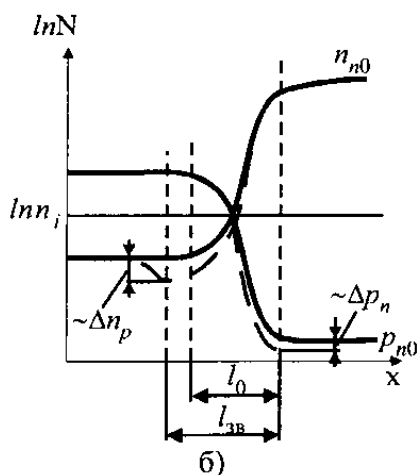
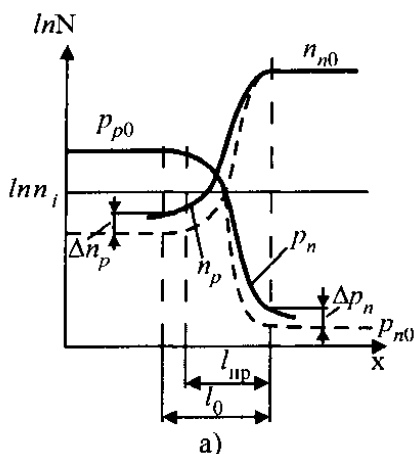


Рис. 10.4. Розподіл концентрації носіїв при прямому (а) і зворотному (б) включенні переходу.

Звідси видно, що у випадку прямого ввімкнення переходу ($u > 0$) в кожному з шарів з'являються надлишкові носії заряду ($\Delta p_n > 0$, $\Delta n_p > 0$), причому концентрація їх збільшується зі зростанням прикладеної напруги. Це означає, що у випадку прямого ввімкнення переходу відбувається інжекція неосновних носіїв заряду в суміжні шари. Розподіл концентрації носіїв заряду при прямому ввімкненні показано на рис. 10.4, а. З (10.13) видно, що при зворотному ввімкненні переходу ($u < 0$) суміжні шари збіднюються неосновними носіями заряду ($\Delta p_n < 0$, $\Delta n_p < 0$), тобто відбувається екстракція неосновних носіїв. Розподіл концентрації носіїв заряду при зворотному вмиканні переходу показано на рис. 10.4, б. Ця обставина є принциповою особливістю р-п-переходу – він належить до інжектуючих контактів.

Порівняємо прирости концентрацій неосновних носіїв на межах переходу в суміжних шарах:

$$\frac{\Delta n_p}{\Delta p_n} = \frac{\Delta n_{p0}}{\Delta p_{n0}} = \frac{\Delta n_{n0}}{\Delta p_{p0}}$$

Маємо свідчення того, що в несиметричному переході як інжекція, так і екстракція неосновних носіїв мають напрямлений характер: інжекція відбувається, в основному, у високоомний шар, а екстракція – в основному, з високоомного шару (наприклад, якщо $N_D \gg N_A$, то $n_n \gg p_{p0}$ і $\Delta n_p \gg \Delta p_n$).

Зазвичай шар з меншим питомим опором називають емітером, а з більшим – базою. Таким чином, при прямому ввімкненні переходу відбувається інжекція неосновних носіїв у базу, що викликає проходження прямого струму, при зворотному – екстракція цих носіїв з бази, що викликає проходження зворотного струму.

Із співвідношень (10.13) видно, що у випадку прямого ввімкнення переходу надлишкові концентрації завжди додатні і необмежено зростають зі збільшенням прямої напруги. При зворотному вмиканні надлишкові концентрації стають від'ємними, однак вони обмежені

величинами $\Delta n_p = \Delta n_{p0}$; $\Delta p_0 = \Delta p_{n0}$, і це обмеження настає практично при $|U| > (2\div 3)\phi_T$. Отже прямий струм через перехід за абсолютною величиною більший від зворотного.

Сказане можна якісно пояснити за допомогою зонних діаграм переходу. При прямому вмиканні потенційний бар'єр зменшується на величину прикладеної напруги, отже, дифузійні складові електронного та діркового струмів зростають. Разом з цим, оскільки результуюча напруженість поля в переході зменшується, зменшуються і дрейфові складові струмів. Це означає, що через перехід буде проходити напрямлений потік носіїв, що визначається переважанням дифузійної складової над дрейфовою:

$$j_{\Sigma} = -j_{n\text{др.}} + j_{n\text{диф.}} + j_{p\text{диф.}} + j_{p\text{др.}} > 0$$

Для зворотного вмикання переходу картина буде іншою. Потенційний бар'єр зростає на величину прикладеної напруги. Тому дифузійна складова струму зменшується, а дрейфова, навпаки, зростає. Через перехід буде проходити струм, густина якого визначається переважанням дрейфової складової над дифузійною:

$$j_{\Sigma} = j_{n\text{др.}} + j_{n\text{диф.}} + j_{p\text{диф.}} + j_{p\text{др.}} < 0$$

Очевидно, що зі збільшенням зворотної напруги дифузійна складова буде зменшуватися, прямуючи до нуля, дрейфова ж – зростати, залишаючись, однак, обмеженою концентраціями неосновних носіїв в р- і n-шарах. Таким чином, при досить великому зворотному зміщенні $|U| > (2-3)\phi_T$ зворотний струм буде мати деяке максимальне значення. Цей струм називається струмом насичення і визначається власною провідністю (концентрацією власних носіїв) матеріалу.

10.3 Вольт-амперна характеристика і ємність переходу

Як і раніше, будемо вважати, що $N_D \gg N_A$. Припустимо, що перехід нескінченно вузький (не будемо враховувати явищ, які відбуваються в р-n-переході) та рівень інжекції низький: $d = \frac{\Delta p_n}{n_{n0}} \ll 1$ (це дозволяє

вважати надлишкову концентрацію невеликою і враховувати лише дифузійну складову струму). З урахуванням цих припущень рівняння неперервності для електронейтральних областей напівпровідника буде мати

розв'язок $\Delta n_p(x) = \Delta n(0) \exp\left(\frac{x}{L_n}\right)$ для шару р (тут $x < 0$,

рис. 10.5) і $\Delta p_n(x) = \Delta p(0) \exp\left(-\frac{x}{L_n}\right)$ для шару n.

Підставляючи замість $\Delta n(0)$ і $\Delta p(0)$ відповідні вирази зі співвідношень (10.13), одержуємо:

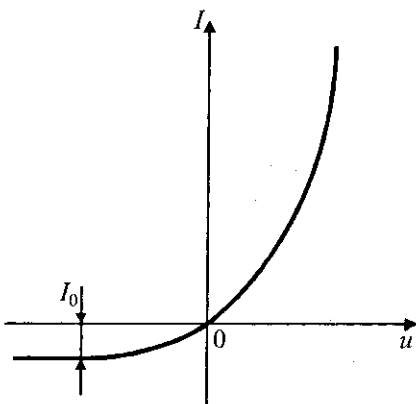


Рис. 10.5. Вольт-амперна характеристика переходу в безрозмірних координатах.

$$\Delta n_p = n_{p0} \left[\exp\left(\frac{|U|}{j_T} - 1\right) \right] \exp\frac{x}{L_n};$$

$$\Delta p_n = p_{n0} \left[\exp\left(\frac{|U|}{j_T}\right) - 1 \right] \exp\left(-\frac{x}{L_n}\right).$$

Продиференціювавши одержані вирази за координатою, одержимо розподіл густини електронного струму в р-шарі та діркового в n-шарі:

$$j_n(x) = q \frac{D_n n_{p0}}{L_n} \left[\exp\left(\frac{|U|}{j_T} - 1\right) \right] \exp\frac{x}{L_n};$$

$$j_p(x) = q \frac{D_p p_{n0}}{L_p} \left[\exp\left(\frac{|U|}{j_T}\right) - 1 \right] \exp\left(-\frac{x}{L_n}\right).$$
(10.14)

Як видно, дірковий та електронний струми залежать від координати і при $x \gg L$ набувають нульового значення, що є наслідком рекомбінації електронів у базі та дірок у емітері.

Беручи максимальні значення струмів (припускаючи в (10.14) $x = 0$), множачи обидві складові на площу переходу S і додаючи, одержуємо вольт-амперну характеристику р-п-переходу:

$$I_0 = qS \left(\frac{D_p}{L_p} p_{n0} + \frac{D_n}{L_n} n_{p0} \right) \left(\exp\frac{U}{j_T} - 1 \right).$$
(10.15)

Очевидно, що при досить великих зворотних напругах струм через перехід не залежить від прикладеної напруги і набуває деякого постійного значення:

$$I_0 = qS \left(\frac{D_p}{L_p} p_{n0} + \frac{D_n}{L_n} n_{p0} \right).$$
(10.16)

Цей струм називається зворотним струмом насичення, або тепловим струмом.

Вольт-амперна характеристика переходу (рис. 10.5), як видно з (10.15), істотно нелінійна, причому прямий струм значно перевищує зворотний. Разом з тим, що р-п-перехід є інжектуючим контактом, це також принципова особливість подібного контакту.

Відзначимо також, що, як видно з (10.15), невелика зміна напруги приводить до досить значних змін струму. З цієї причини р-п-перехід часто використовують у так званому режимі заданих струмів. У цьому випадку (10.15) доцільно переписати у вигляді:

$$U = j_T \ln\left(\frac{I}{I_0} + 1\right).$$
(10.17)

Звідси видно принципові недоліки переходу як вентиляного елемента: будь-які зміни зворотного струму насичення змінюють пряму напругу в протилежному напрямку.

Оцінити нелінійність вольт-амперної характеристики переходу доцільно порівнянням опору

переходу при прямому та зворотному ввімкненні. Звичайно говорять про диференційний опір $R_j = du/di$ та опір постійному струму $R_n = U/I$. З (10.17) та рис. 10.5 видно, що перехід характеризується дуже малим опором при прямому ввімкненні та неперервно зростаючим при зворотному.

Вище вказувалось, що ширина переходу i , отже, величина об'ємного заряду змінюються зі зміною прикладеної напруги, тобто перехід характеризується деякою ємністю. Зазвичай її називають бар'єрною. Оцінімо величину цієї ємності. Перехід подамо у вигляді конденсатора, заряди на обкладках якого дорівнюють $|-Qp| = Qn = Q$. Якщо перехід несиметричний і $l \approx l_p$, то $Q \approx qN_A S l$, де, як і раніше, S – площа переходу, l – його ширина. Враховуючи (10.12), маємо:

$$Q \approx qN_A S l_0 \left(\frac{|U|}{\Delta j_0} \right)^{1/2}$$

Тоді диференційне значення бар'єрної ємності буде дорівнювати:

$$C_\sigma = \frac{dQ}{dU} = S \left(\frac{e_0 e q N_A}{2|U|} \right)^{1/2} = \frac{e_0 e S}{l_0} \left(\frac{\Delta j_0}{|U|} \right)^{1/2} = C_0 \left(\frac{\Delta j_0}{|U|} \right)^{1/2} \quad (10.18)$$

Тут C_0 можна інтерпретувати як ємність переходу в рівноважному стані. Як видно, ємність переходу обернено пропорційна $|U|^{1/2}$, а також прямо пропорційна квадратному кореню з концентрації домішок у високоомному шарі. Оцінімо значення бар'єрної ємності для реального переходу. При $l_0 \approx 0,8$ мкм, $\Delta \varphi_0 \approx 0,4$ В і $|U| = 1$ В для германію $C_\sigma \approx 120$ пФ, а при $|U| = 1,5$ В маємо $C_\sigma \approx 95$ пФ.

Залежність ємності переходу від прикладеної напруги дає змогу застосовувати перехід як керовану ємність, що використовується, наприклад, у параметричних діодах та помножувачах частоти з електронним перестроюванням.

Що таке пробій переходу. Це явище полягає в тому, що при деякому значенні зворотної напруги струм через р-п-перехід різко зростає, що супроводжується істотним зменшенням диференційного опору при незначному прирості напруги. Типова залежність зворотного струму від напруги в області пробою показана на рис. 10.6. Залежно від причин, що викликають пробій р-п-переходу, розрізняють польовий, лавинний, тепловий і поверхневий пробій.

Перші два пробіи – польовий і лавинний – пов'язані зі зростанням енергії та концентрації електронів при збільшенні напруженості поля в переході. Ці види пробою відбуваються внаслідок ударної або електростатичної іонізації. Хід вольт-амперної характеристики в області польового і лавинного пробіюв приблизно однаковий і показаний лініями 1 і 2 на рис. 10.6. Різка залежність

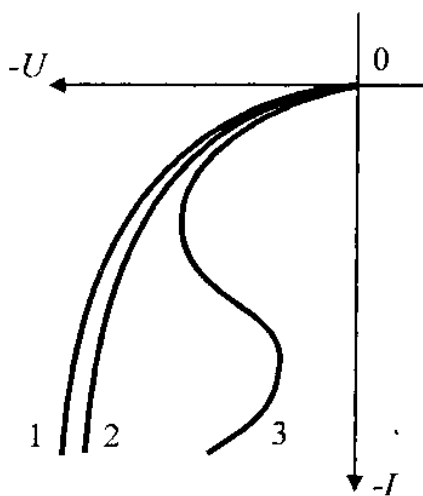


Рис. 10.6. Вольт-амперна характеристика області пробою.

струму від напруги при зворотному ввімкненні переходу в передпробійній області використовується для стабілізації напруги. Діоди, призначені для цього, називаються стабілітронами.

Тепловий пробій зумовлюється виділенням тепла при проходженні струму через р-п-перехід. Цей вид пробою може мати місце при набагато нижчих зворотних напругах, ніж польовий та лавинний, якщо не здійснюється необхідне відведення тепла від переходу, що працює в режимі великих струмів. У подібній ситуації перехід може нагрітися до температури, при якій можливий розрив валентних зв'язків внаслідок теплового збудження. Це викликає підвищення концентрації вільних носіїв, і, отже, подальше збільшення струму. В результаті перехід буде нагріватися ще більше, а струм зростатиме. Вольт-амперна характеристика в області теплового пробою зображена на рис. 10.6 лінією 3.

Поверхневий пробій р-п-переходу зумовлений діелектричними властивостями поверхні напівпровідника, прилеглої до переходу. На ці властивості впливають шари окислів, що утворюються на поверхні, захисні шари, нанесені для захисту від вологи та забруднень на поверхню напівпровідника, а також поверхневий заряд, який практично завжди утворюється на поверхні напівпровідника і місці виходу на поверхню р-п-переходу.

10.4. Неінжектуючі контакти, тунельні переходи і гетеропереходи

10.4.1. Неінжектуючі контакти. Розглянемо контакт типу n^+-n , тобто перехід на основі електронного напівпровідника, одна частина якого легована донорною домішкою дещо сильніше (рис. 10.7).

Легко помітити, що градієнт концентрації однойменних носіїв по обидва боки переходу значно менший, ніж у класичному переході. Отже, висота бар'єра в розглянутому випадку може бути дуже незначною. Це видно і з формули (10.2), куди замість n_{p0} потрібно підставити величину n_{n0} , що визначає концентрацію основних носіїв у n-шарі.

Зниження градієнта концентрації призводить також до того, що електрони з n^+ -шару проникають на меншу глибину в n-шар. В результаті в n^+ -шарі виникає невеликий позитивний заряд іонів донорів, а в n-шарі – невеликий негативний заряд рухомих носіїв. Отже, ширина такого переходу порівняно з класичним зменшується. Це видно також і з формули (10.8), в яку замість $\Delta\phi_0$ необхідно підставити $\Delta\phi_0' < \Delta\phi_0$, а замість N_A – концентрацію донорів у високоомному n-шарі, припускаючи при цьому, що $N_D' > N_A$.

З наведених міркувань ясно, що нелінійність вольт-амперної характеристики подібного переходу набагато менша, ніж нелінійність вольт-амперної характеристики класичного переходу. Очевидно також, що при $N_D' = N_A$

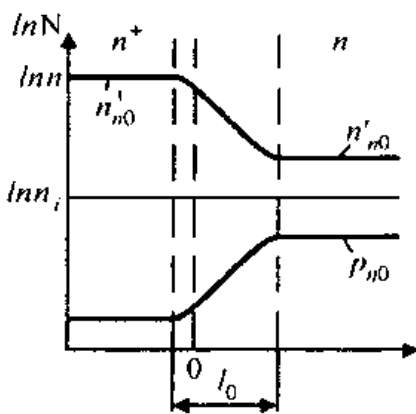


Рис. 10.7. Розподіл концентрації носіїв у контакті n^+-n .

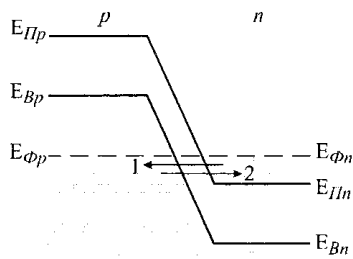


Рис. 10.8. Тунельний перехід у рівноважному стані.

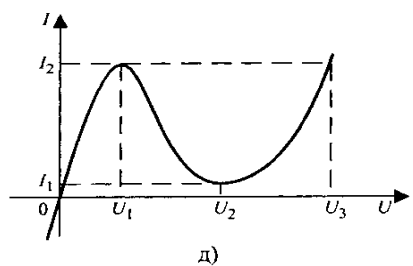
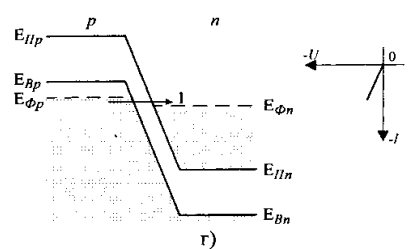
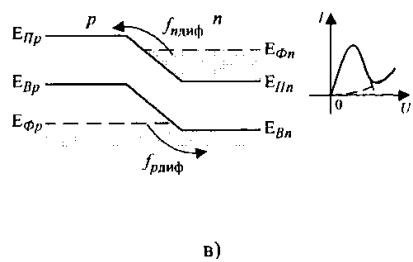
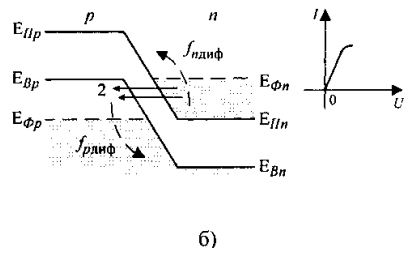
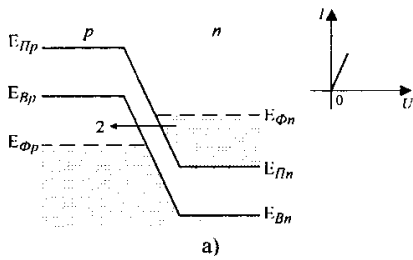


Рис. 10.9. Протікання прямого (а, б, в) і зворотного (г) струмів і вольт-амперна характеристика (д) тунельного переходу.

перехід зникає і виникає однорідний, досить сильно легований напівпровідник. Усе сказане справедливе і для переходів типу p^+-p .

Принциповою особливістю описаних переходів є те, що основні носії по обидва боки межі розділу однакові. Тому в таких переходах не відбуваються інжекція та екстракція неосновних носіїв заряду при прямому та зворотному вмиканні. Такі переходи називаються неінжектуючими. Вони характеризуються незначними вентильними властивостями і можуть використовуватися для створення омичних контактів.

10.4.2. Тунельні переходи. Робота тунельних діодів побудована на тунелюванні електронів через дуже тонкий p - n -перехід, виготовлений з вироджених напівпровідників. Зонна діаграма тунельного переходу в стані рівноваги показана на рис. 10.8.

Необхідно підкреслити, що утворене внаслідок виродження перекриття зон є необхідною умовою для тунелювання електронів крізь потенціальний бар'єр вузького переходу. Перекриття зон та надзвичайно мала ширина переходу призводять до виникнення деякої ймовірності переходу електронів крізь потенціальний бар'єр із зони провідності n -шару у валентну зону p -шару, і навпаки. Це явище має квантово-механічний характер і полягає в тому, що електрон з енергією, меншою від висоти потенціального бар'єра, може все ж проникнути крізь бар'єр.

Якісно вольт-амперну характеристику тунельного переходу можна пояснити так. У рівноважному стані (рис. 10.8) рівні Фермі в p - і n -шарах єдині. З'являється можливість тунельного переходу електронів з валентної зони p -шару в зону провідності n -шару і навпаки. Потік електронів з зони провідності n -шару створює прямий струм, потік електронів з валентної зони p -шару – зворотний. У рівноважному стані ці потоки зрівноважують один одного, оскільки всі стани нижче рівня Фермі заповнені і перехід деякої кількості електронів зліва направо можливий лише в тому випадку, коли справа наліво перейде така ж кількість електронів.

У випадку вмикання переходу в прямому напрямку потенціальний бар'єр знижується, ступінь перекриття зон зменшується. Навпроти вільних рівнів валентної зони p -шару опиняються заповнені рівні зони провідності n -шару. (рис. 10.9, а). Виникає напрямлений потік електронів, що забезпечує прямий струм. Очевидно, цей потік досягне максимального значення, коли навпроти всіх вільних рівнів валентної зони p -шару будуть розміщуватися всі заповнені рівні зони провідності n -шару (рис. 10.9, б). При подальшому збільшенні напруги прямого зміщення тунельний струм буде зменшуватися і перетвориться на нуль, коли навпроти стелі валентної зони p -шару буде розташоване дно зони провідності n -шару.

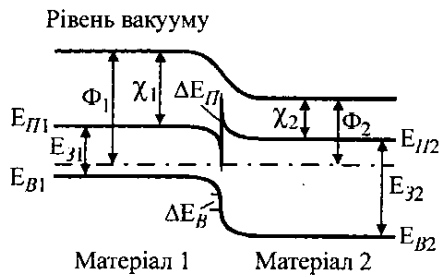


Рис. 10.10. Зонна діаграма гетеро переходу.

Однак, оскільки зі збільшенням напруги прямого зміщення неперервно зменшується висота потенціального бар'єра, починає зростати роль дифузійного струму через перехід (штрихові криві на рис. 10.9, б). При подальшому збільшенні прямої напруги через перехід протікає тільки дифузійний струм (рис. 10.9, в).

Зворотне ввімкнення переходу призводить до того, що висота потенціального бар'єра зростає, отже зростає і перекриття зон. Навпроти вільних рівнів зони провідності n-шару опиняються заповнені рівні валентної зони p-шару. Різко зростає тунельний струм, що визначається потоком електронів, які переходять з валентної зони (рис. 10.9, г).

Таким чином, вольт-амперна характеристика тунельного переходу має вигляд, зображений на рис. 10.9, д. При напругах, менших U_2 , через перехід протікає тунельний струм, при напругах, більших від U_2 – звичайний дифузійний струм, зумовлений процесами, що відбуваються в класичному p-n-переході.

Основною особливістю переходу є наявність спадаючої ділянки на вольт-амперній характеристиці. Ця ділянка характеризується від'ємним значенням диференціальної провідності $di/du < 0$ на ділянці $U_1 < U < U_2$.

10.4.3. Гетеропереходи. Гетероперехід – перехід, утворений напівпровідниковими матеріалами з різною шириною забороненої зони. На основі подібних переходів з їх дуже специфічними властивостями створено ряд напівпровідникових приладів, характеристики яких значно відрізняються від характеристик приладів на основі класичного переходу. Наприклад, на основі гетеропереходів виготовлені швидкодіючі перемикаючі p-n-діоди, напівпровідникові лазери, фототранзистори тощо.

З зонної діаграми гетеропереходу (рис. 10.10) видно, що два напівпровідники мають різну ширину забороненої зони E_{z1} і E_{z2} , різну роботу виходу Φ_1 і Φ_2 та різну спорідненість до електрона χ_1 і χ_2 . Роботою виходу називається енергія, яку необхідно надати електрону для того, щоб він вилетів у вакуум з нульовою початковою швидкістю. Іншими словами – це енергетичний проміжок між рівнем вакууму та рівнем Фермі. Аналогічно визначається і енергія спорідненості до електрона – це енергетичний проміжок між рівнем вакууму та дном зони провідності.

У рівноважному стані розташування рівнів Фермі по обидва боки межі розділу однакове. Рівень вакууму завжди неперервний і розміщений паралельно межах зон, тому вигин зон, що виникає у приконтактній зоні, повторює вигин рівня вакууму. Завдяки цьому в зоні провідності та валентній зоні виникають енергетичні розриви.

Величина розриву ΔE_{II} у зоні провідності визначається різницею в енергіях спорідненості до

електрона. Величина розриву ΔE_{II} у валентній зоні визначається, крім того, і різницею у величині заборонених зон. Це свідчить про те, що висота потенціальних бар'єрів для електронів та дірок у таких переходах різна. Попри те, умови проходження бар'єрів носіями також різні: наприклад, електрони можуть нагромаджуватися у "западині", а потім проходити "пік" за рахунок тунельного ефекту. Ширина області об'ємного заряду в кожному з матеріалів, що утворюють гетероперехід, залежить від ступеня легування; емність переходу, в основному, також визначається більш високоємним матеріалом.

Різниця висоти потенціальних бар'єрів для електронів та дірок – принципова особливість гетеропереходів. При подачі напруги зміщення на перехід висота бар'єра змінюється, однак різниця в умовах руху носіїв через перехід зберігається. Ця особливість дає змогу одержувати на основі гетеропереходів структури типу n-n або p-p з несиметричною вольт-амперною характеристикою, тобто випрямними властивостями. Вибираючи відповідні матеріали для створення гетеропереходу, тобто, підбираючи певним способом співвідношення енергій спорідненості до електрона, робіт виходу, тип провідності та ступінь легування матеріалу, можна одержати структури із заданими електрофізичними характеристиками.

10.5. Контакти

10.5.1. Контакт метал-напівпровідник.

Встановлення рівноваги у приконтактній області. Властивості контакту метал-напівпровідник залежать, головним чином, від взаємного розміщення рівнів Фермі в матеріалах, що контактують. Розглянемо ідеалізовану модель контакту, припускаючи, що поверхні металу та напівпровідника не містять ніяких сторонніх фаз або поверхневих електронних станів і не розділяються шаром окислу. Для певності розглянемо напівпровідник n-типу, робота виходу якого Φ_n , і метал, робота виходу якого Φ_m .

На рис. 10.11, а зображено випадок $\Phi_n < \Phi_m$. Це значить, що після дотикання частина електронів з напівпровідника перейде в метал. Напрявлений потік електронів буде тривати доти, доки рівні Фермі в обох частинах системи не розташуються однаково. Тоді у напівпровіднику поблизу межі розділу виникне позитивний просторовий заряд, а в металі – негативний. Це призведе до викривлення енергетичних рівнів у напівпровіднику та виникнення потенціального бар'єра висотою $\Delta\phi_0$. Як видно з рис. 10.11, висота потенціального бар'єра дорівнюватиме $\Delta\phi_0 = (\Phi_m - \Phi_n)/q$.

Особливість розглянутого випадку полягає в тому, що в напівпровіднику концентрація електронів значно менша, ніж у металі. Тому об'ємний заряд розповсюджується в металі на глибину близько 10^7 см, у

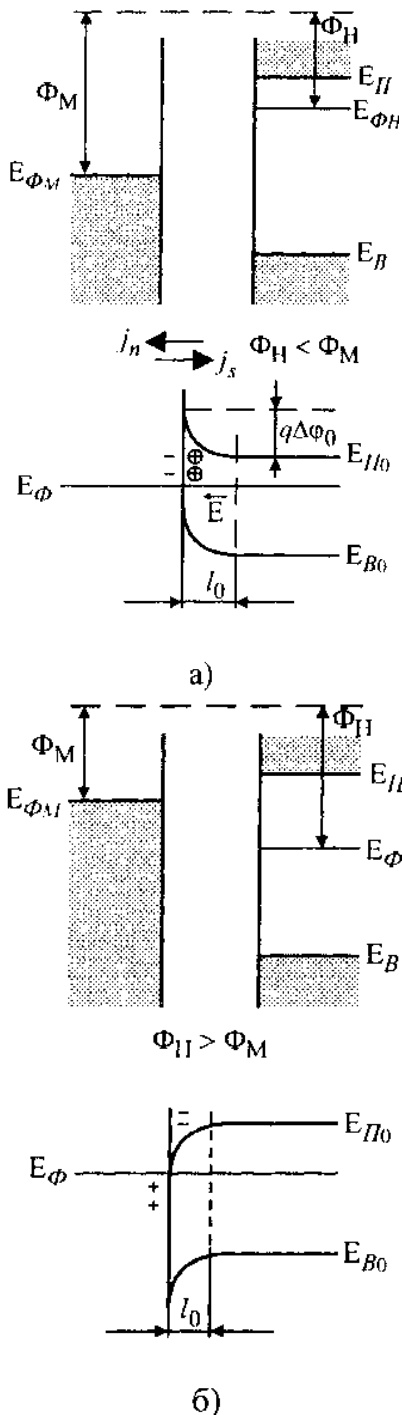


Рис. 10.11. Контакт метал-напівпровідник n-типу.

напівпровіднику ж він розповсюджується значно глибше. Приповерхневий шар напівпровідника шириною l_0 збіднюється основними носіями заряду. Отже, питомий опір цього шару буде більшим від питомого опору прилеглих областей і визначатиме опір структури загалом. Шар з підвищеним питомим опором іноді називається запірним.

Розглянемо тепер рис. 10.11, б. При встановленні контакту частина електронів з металу перейде в напівпровідник, рівні Фермі металу і напівпровідника розташуються однакою. Енергетичні рівні напівпровідника у приграничній області викривляються вниз, ця область буде збагачена основними носіями заряду. Таким чином, питомий опір приграничної області буде меншим від питомого опору решти напівпровідника. Тому опір приграничного шару мало впливатиме на опір структури загалом. Отже, під'єднання зовнішнього джерела зміщення в тому чи іншому напрямку не змінить істотно опору структури. Це дозволяє вважати розглянутий контакт не випрямним. Такі контакти називають також антизапірними.

Очевидно, контакт металу з напівпровідником р-типу буде мати запірні властивості, якщо робота виходу металу буде меншою від роботи виходу напівпровідника. В іншому випадку контакт буде антизапірним.

При контакті металу з власним напівпровідником викривлення енергетичних зон призводить до утворення в напівпровіднику приконтатної області з підвищеною провідністю як у випадку $\Phi_n > \Phi_m$, так і у випадку $\Phi_n < \Phi_m$.

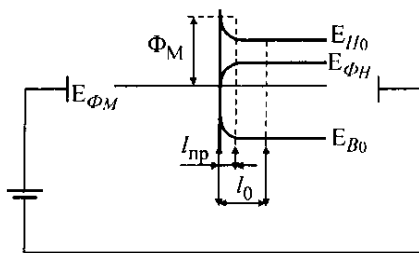
10.5.2. Випрямний контакт. Розглянемо випрямний контакт метал-напівпровідник n-типу (рис. 10.12, а). Якщо припустити, що всі атоми домішки в запірному шарі іонізовані і звільнені електрони перейшли в метал, а весь запірний шар розміщений у напівпровіднику, то рівняння Пуассона (10.4) буде мати розв'язок:

$$j(x) = \left(\frac{qN_D}{2e_0e} \right) (x - l_0)^2$$

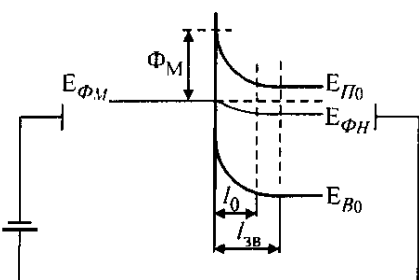
Звідси видно, що з віддаленням від контакту потенціал у напівпровіднику спадає за квадратичним законом. Приймаючи в цьому виразі $x = 0$ і враховуючи, що при цьому $\varphi(x) = \Delta\varphi_0$, одержимо ширину шару просторового заряду:

$$l_0 \approx \left(\frac{2e_0e\Delta\varphi_0}{qN_D} \right)^{1/2} \approx \left[\frac{2e_0e}{q^2n_{n0}} (\Phi_m - \Phi_n) \right]^{1/2}$$

Таким чином, ширина запірного шару тим більша, чим більша різниця робіт виходу металу і напівпровідника і чим менша концентрація основних носіїв у напівпровіднику. Наприклад, при $\Phi_m - \Phi_n = 1\text{eV}$ для кремнію з $N_D = 10^{15}\text{ см}^{-3}$ маємо $l_0 \approx 1,5\text{ мкм}$ і $l_0 \approx 0,1\text{ мкм}$



а)



б)

Рис. 10.12. Пряме (а) і зворотне (б) включення контакту метал-напівпровідник n-типу.

при $\Phi_m - \Phi_n = 0,4 \text{ eV}$ і $N_D = 10^{17} \text{ см}^{-3}$.

З наведеного виразу видно також, що зовнішня прикладена напруга повинна змінити ширину шару. На рис. 10.12, а показано бар'єр після прикладання до контакту напруги в прямому напрямку. "Плюс" джерела приєднаний до металу, "мінус" – до напівпровідника. Електричне поле, що діє в приконтактній ділянці, послаблюється, потенціальний бар'єр знижується: $\Delta\varphi = \Delta\varphi_0 - u$.

Відзначимо, що, оскільки в металі потенціал дорівнює нулю, то величина Φ_{mc} , яка дорівнює енергії, необхідній для переведення електрона з рівня Фермі металу на дно зони провідності напівпровідника у контакті, залишається постійною при будь-якій прикладеній напрузі.

При прикладанні до контакту зовнішньої напруги у зворотному напрямку (рис. 10.12, б), всі рівні напівпровідника опускаються, потенціальний бар'єр підвищується: $\Delta\varphi = \Delta\varphi_0 - u$. Тому:

$$l_0 \approx \left(\frac{2e_0 e (\Delta j_0 \pm u)}{q N_D} \right)^{1/2} \approx \left[\frac{2e_0 e (\Phi_m - \Phi_n \pm qu)}{q^2 n_{n0}} \right]^{1/2}$$

Таким чином, при вмиканні контакту в прямому напрямку ширина запірного шару зменшується, при вмиканні у зворотному – збільшується.

Вольт-амперна характеристика випрямного контакту. В рівноважному стані контакту потоки електронів з металу в напівпровідник і назад однакові, а струми, що визначаються ними, рівні за величиною і протилежні за напрямком – густина сумарного струму через контакт дорівнює нулю: $j_\Sigma = j_s - j_n = 0$.

При вмиканні контакту в прямому напрямку потік електронів з металу зберігається незмінним, а потік електронів з напівпровідника зростає відповідно до закону Больцмана в \exp разів (бар'єр висотою $\Delta\varphi_0 - u$ може подолати в \exp разів більша кількість електронів, ніж бар'єр висотою $\Delta\varphi_0$). Виникає прямий струм через контакт: $j_\Sigma = j_s - j_n > 0$.

При вмиканні контакту в зворотному напрямку потенційний бар'єр для електронів, що йдуть з напівпровідника в метал, зростає на величину прикладеної напруги. Тому потік електронів з напівпровідника в метал зменшується в \exp разів, а потік електронів з металу залишається постійним. Виникає зворотний струм через контакт: $j_\Sigma = j_s - j_n < 0$.

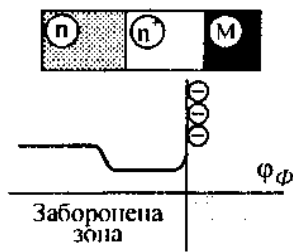
Таким чином, вольт-амперна характеристика випрямного контакту має вигляд:

$$I = I_s \left(\exp \frac{u}{j_T} - 1 \right) \quad (10.19)$$

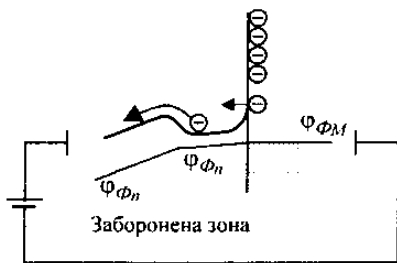
З одержаного виразу видно, що вольт-амперна характеристика розглянутого контакту характеризується

значною нелінійністю. Отже, прямий струм через контакт у багато разів перевищує зворотний струм при одному і тому ж значенні прикладеної напруги. Вольт-амперна характеристика контакту, що визначається з виразу (10.19), на перший погляд збігається з вольт-амперною характеристикою класичного переходу. Однак природа зворотного струму тут інша. Залежно від співвідношення довжини вільного пробігу носіїв заряду та ширини заперного шару змінюється характер пролітання носіїв через цей шар. Більш строгий аналіз руху електронів через запираючий шар дає змогу одержати аналітичний вираз для струму I_s .

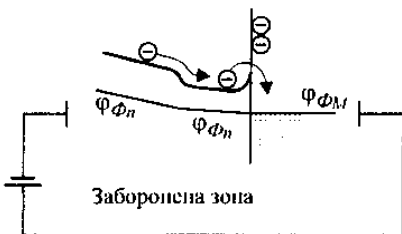
Відзначимо, що у випадку тонкого заперного шару зворотний струм дуже швидко досягає насичення. При однаковій концентрації носіїв заряду n_{n0} тонкий заперний шар буде мати менший опір, ніж товстий, тому при прямому зміщенні через контакт будуть протікати великі струми.



а)



б)



в)

Рис. 10.13. Зонна діаграма омічного контакту в рівноважному стані (а), при прямому (б) і зворотному (в) включенні.

10.5.3. Омічний контакт. Омичним контактом вважається контакт між металом та напівпровідником, який має лінійну вольт-амперну характеристику. Такі властивості має структура метал-сильно легований напівпровідник-напівпровідник. Така структура на основі напівпровідника n-типу показана на рис. 10.13. Структура складається з двох контактів: n-n⁺ і n⁺-n (при цьому робота виходу металу менша від роботи виходу напівпровідника). Обидва контакти неінжектуючі, мають незначний опір та практично лінійну вольт-амперну характеристику.

При прямому вмиканні контакту (рис. 10.13, б) зменшується висота бар'єрів, оскільки зменшуються перешкоди для проходження прямого струму через перехід. При зворотному вмиканні контакту (рис. 10.13, в) потенціальні бар'єри зростають. Однак у цьому випадку потік електронів скерований до позитивного полюса джерела, отже збільшення висоти бар'єрів не має значення.

10.5.4. Контакт метал-метал. Контактна різниця потенціалів. Розглянемо процес, що відбувається при зближенні двох металів на відстань l , при якому внаслідок термоелектронної емісії можливий ефективний обмін електронами. Припустимо, що $\Phi_{M1} > \Phi_{M2}$, а $E_{\phi 1} < E_{\phi 2}$. Енергетична діаграма наведена на рис. 10.14, а. Оскільки $\Phi_{M1} > \Phi_{M2}$, $j_2 > j_1$, де j_1 , j_2 – густини струмів термоелектронної емісії з металу 1 і металу 2. Таким чином, перенесення електронів буде здійснюватися справа наліво. Поверхня металу 1 заряджається негативно, металу 2 – позитивно. В проміжку виникає електричне поле E_K і відповідна різниця потенціалів. Це призводить до встановлення динамічної рівноваги, умовою якої є рівність $E_{\phi 1} = E_{\phi 2}$. При цьому $j_1' = j_2'$, а це значить, що потенціальний бар'єр для електронів, що рухаються справа наліво, дорівнює потенціальному бар'єру для

електронів, що рухаються в зустрічному напрямку: $\Phi_{M2} + qV_k = \Phi_{M1}$ (рис. 10.14, б).

Отже, після встановлення рівноваги створюється контактна різниця потенціалів $V_k = (\Phi_{M1} - \Phi_{M2})/q$, що визначається різницею робіт виходу металів. Вона називається зовнішньою контактною різницею потенціалів.

Оцінимо кількість електронів, перехід яких з одного металу до другого створює V_k . Нехай $\Phi_{M1} - \Phi_{M2} = 1$ еВ, а $l = 10^{-7}$ см, тоді $E_k = 10^7$ В/см. Зобразимо систему у вигляді плоского конденсатора. Густина електронів на його обкладках буде $Q_S = \epsilon_0 E_k$, а кількість електронів, необхідна для створення Q_S , визначається так: $n = Q_S/q \approx 5 \cdot 10^{12}$ см⁻². Оскільки кількість електронів на одиниці поверхні металу $n_S \approx (10^{22})^{2/3} \approx 10^{15}$ см⁻², то $n/n_S \approx 5 \cdot 10^{-3}$. Це значить, що за рахунок мізерно малого збіднення електронами поверхні одного металу та відповідного збагачення поверхні другого створюється значна контактна різниця потенціалів.

У випадку жорсткого контакту цих металів (рис. 10.14, б) внаслідок того, що $E_{\phi 1} < E_{\phi 2}$ і дифузії електронів з металу 2 в метал 1 виникає електричне поле у площині, перпендикулярній до площини контакту і відповідна різниця потенціалів. Вона визначається різницею енергій Фермі ізолюваних металів $V_{ki} = (E_{\phi 2} - E_{\phi 1})/q$ і називається внутрішньою контактною різницею потенціалів. Оскільки енергія Фермі пов'язана з концентрацією електронів та їх ефективною масою, то і внутрішня контактна різниця потенціалів визначається концентрацією електронів у металах та їх ефективними масами. Значення V_{ki} невелике – близько 10^{-3} В. При цьому товщина подвійного електричного шару становить величину близько 10^{-8} см, а для його утворення необхідним є перетікання не більше 2% вільних електронів з поверхні одного металу на контактну поверхню іншого. Це свідчить про те, що провідність шару дуже мало відрізняється від об'ємної провідності контактуючих металів і в ідеальних контактах наявність V_{ki} не перешкоджає проходженню струму. Однак у реальних контактних системах справа виглядає інакше.

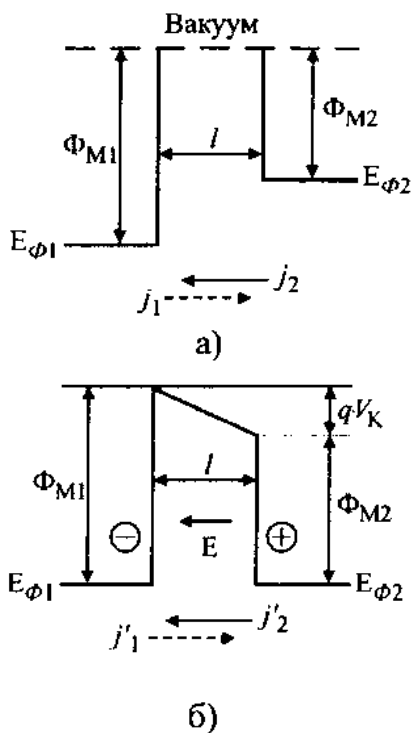


Рис. 10.14. Енергетична діаграма двох металів до (а) і після (б) встановлення термодинамічної рівноваги.

Лекція 11

ДЮДИ І ТРАНЗИСТОРИ.

План

11.1. Напівпровідникові діоди

11.2. Біполярні транзистори

11.3. Модель Еберса-Молла

Література

11.1. Пасынков В. В., Чиркин Л. К. Полупроводниковые приборы. – М.: Высшая школа, 1987. – 479 с.

11.2. Гусев В. Г., Гусев Ю. М. Электроника. – М.: Высшая школа, 1991. – 622 с.

11.3. Зи С. Физика полупроводниковых приборов.: Пер. с. англ. – М.: Мир, 1984. – 637 с.

11.4. Степаненко Н. П. Основы теории транзисторов и транзисторных схем. – М.: Энергия, 1977.-608 с.

11.1. Напівпровідникові діоди.

Структура та основні елементи. Напівпровідниковий діод – це напівпровідниковий прилад з одним випрямним електричним переходом і двома виводами, в якому використовується та чи інша властивість випрямного електричного переходу.

Випрямним електричним переходом у напівпровідникових діодах може бути електронно-дірковий перехід, гетероперехід або випростовувальний перехід, утворений в результаті контакту між металом і напівпровідником (перехід Шоттки).

У діоді з р-п-переходом або з гетеропереходом крім випрямного переходу повинно бути два омичні переходи, через які р- і п-області діода з'єднані з виводами (рис. 11.1, а). Діод з випростовувальним електричним переходом у вигляді контакту між металом і напівпровідником має всього один омичний перехід (рис. 11.1, б).

Зазвичай напівпровідникові діоди мають несиметричні р-п-переходи. Тому при полярності зовнішньої напруги, при якій відбувається зниження потенціального бар'єра в р-п-переході, тобто при прямому напрямку для р-п-переходу, кількість носіїв заряду, інжекттованих із сильнолегованої в слабколеговану область, значно перевищує кількість носіїв, що проходять у протилежному напрямку. За загальним визначенням, область напівпровідникового діода, в яку відбувається інжекція неосновних для цієї області носіїв заряду, називається базою діода. Отже, в діоді базовою областю є слабколегована область. Якщо до діода з несиметричним переходом прикладена напруга, при якій відбувається підвищення потенціального бар'єра в р-п-переході, тобто в зворотному напрямку для р-п-переходу, то екстракція

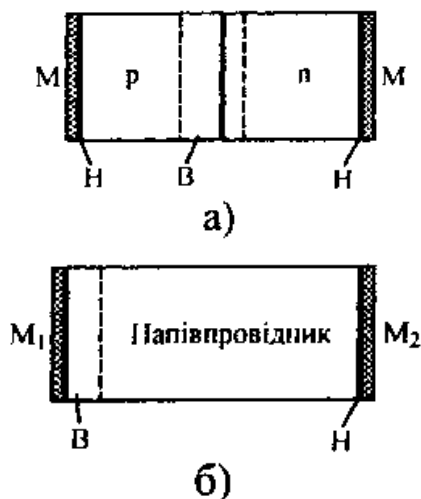


Рис. 11.1. Структури напівпровідникових діодів: а – з випрямним переходом у вигляді р-п-переходу; б – з випрямним електричним переходом на контакті між металом і напівпровідником; В – випрямні електричні переходи; Н – невиврямні (омичні) переходи.

неосновних носіїв заряду буде відбуватися, в основному, з бази діода. Таким чином, база діода може впливати на характеристики та параметри діода.

Залежно від співвідношення лінійних розмірів випрямного електричного переходу та характеристичної довжини розрізняють площинні та точкові діоди. Характеристичною довжиною для діода є найменша за значенням з двох величин, що визначають властивості та характеристики діода: дифузійна довжина неосновних носіїв заряду в базі або товщина бази.

Площинним називають діод, у якого лінійні розміри, що визначають площу випрямного електричного переходу, значно більші, ніж характеристична довжина.

Точковим називають діод, у якого лінійні розміри, що визначають площу випрямного електричного переходу, значно менші, ніж характеристична довжина.

У випрямному електричному переході та прилеглих до нього областях відбуваються різні фізичні процеси, що зумовлюють ефект випрямлення, нелінійне зростання струму зі збільшенням напруги, лавинне розмноження носіїв заряду при ударній іонізації атомів напівпровідника, тунелювання носіїв крізь потенціальний бар'єр випрямного електричного переходу як при зворотній, так, за певних умов, і при прямій напрузі, зміну бар'єрної ємності зі зміною напруги, ефект накопичення і розсмоктування неосновних носіїв заряду в прилеглих до випрямного переходу областях. Усі ці ефекти використовуються для створення різних видів напівпровідникових діодів: випрямних, змішувальних, детекторних і перемикальних, діодів з різким відновленням зворотного опору, стабілітронів, стабісторів, шумових лавинно-пролітних і обернених діодів, варикапів. Деякі з вище згаданих ефектів небажані та шкідливі в одних діодах, але в інших діодах ці ж ефекти можуть бути основою принципу дії.

Пряме вмикання діода. При прямій напрузі на діоді зовнішня напруга частково компенсує контактну різницю потенціалів на р-п-переході, оскільки зовнішнє електричне поле при прямому вмиканні діода спрямоване протилежно дифузійному полю. Тому висота потенціального бар'єра переходу зменшується пропорційно прикладеній до діода напрузі. Нехтуючи падінням напруги на базі діода, розглянемо діод при малих прямих струмах.

Зі зменшенням висоти потенціального бар'єру збільшується кількість носіїв заряду, які долають потенціальний бар'єр і переходять у сусідню область діода, де вони виявляються неосновними носіями. Цей процес називають інжекцією неосновних носіїв заряду через р-п-перехід.

Оскільки висота потенціального бар'єра зменшується пропорційно прикладеній напрузі, а носії

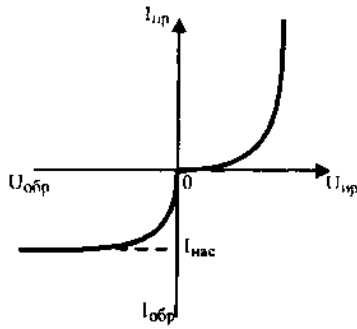


Рис. 11.2. Вольт-ампера характеристика діода при інжекції та екстракції носіїв заряду.

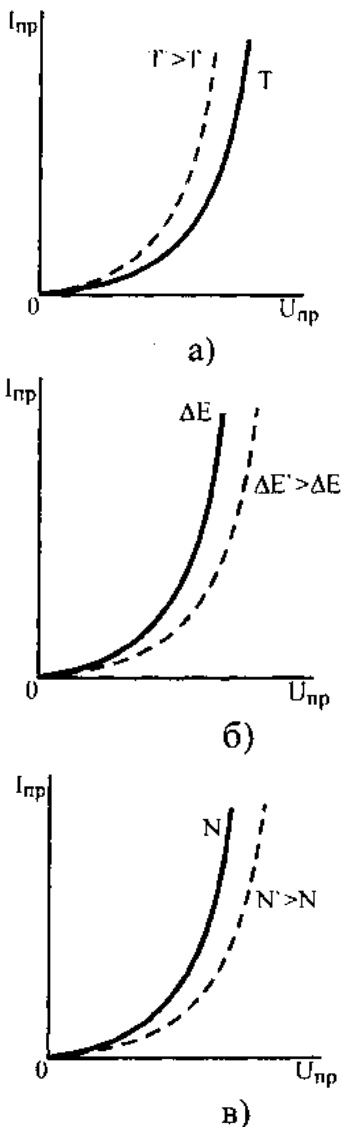


Рис. 11.3. Прямі гілки ВАХ діоду при різних температурах (а), при різній ширині забороненої зони вихідного матеріалу (б) і при різній концентрації домішок у прилеглих до р-п-переходу областях (в).

заряду розподілені за енергіями за експоненційним законом згідно зі статистикою Фермі-Дірака або Максвелла-Больцмана, то пряма гілка ВАХ-діода повинна нагадувати експоненту (рис. 11.2).

Розглянемо вплив деяких факторів на пряму гілку вольт-амперної характеристики діода.

При збільшенні температури діода зменшується висота потенціального бар'єра та змінюється розподіл носіїв заряду за енергіями (електрони, наприклад, займають вищі енергетичні рівні в зоні провідності). В зв'язку з цим прямий струм через діод посилюється зі збільшенням температури при незмінній прямій напрузі (рис. 11.3, а).

При порівнянні прямих гілок двох діодів, виготовлених з різних матеріалів, з різною шириною забороненої зони, видно, що діод з матеріалу з більшою шириною забороненої зони матиме вищий потенціальний бар'єр. Отже, прямий струм через діод з матеріалу з більшою шириною забороненої зони буде меншим при тій самій напрузі (рис. 11.3, б).

Зі збільшенням концентрації домішок у суміжних з р-п-переходом областях збільшується висота потенціального бар'єра переходу, а отже, зменшується прямий струм при тій самій прямій напрузі (рис. 11.3, в).

Зворотне вмикання діода. При зворотному вмиканні діода зовнішнє електричне поле та дифузійне поле в р-п-переході збігаються за напрямком, відбувається екстракція неосновних носіїв заряду з прилеглих до переходу областей. Це зумовлює зменшення граничної концентрації неосновних носіїв заряду біля р-п-переходу та виникнення дифузії неосновних носіїв до переходу – протікає дифузійний струм неосновних носіїв, що з'являються в результаті теплової генерації в об'ємі р- і п-областей діода, а також на омичних переходах.

Протягом життя до р-п-переходу можуть продифундувати неосновні носії, які виникли в п- і р-областях на відстані, що не перевищує відповідної дифузійної хвилі (рис. 11.4, а, б). Решта неосновних носіїв, що не встигли дійти до переходу, рекомбінують в об'ємі. Це справедливо для різних зворотних напруг на діоді, якщо товщини прилеглих до переходу областей перевищують дифузійні довжини неосновних носіїв заряду. Тому зворотний струм, починаючи з дуже малих значень зворотної напруги, не буде змінюватись зі зміною напруги (рис. 11.2). Цей незмінний зі зміною напруги зворотний струм через діод називають струмом насичення.

Розглянувши фізичні процеси в діоді при зворотній напрузі, густину струму насичення можна виразити через параметри напівпровідникового матеріалу. Для цього необхідно згадати загальне співвідношення для густини струму за наявності носіїв заряду двох типів:

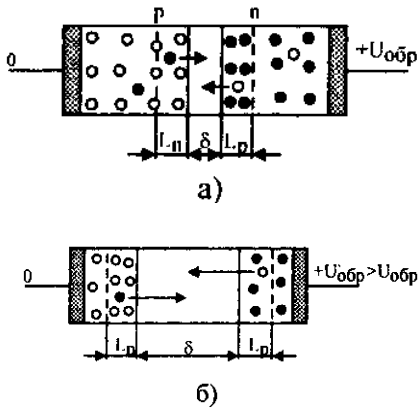


Рис. 11.4. Екстракція неосновних носіїв з прилеглих до р-п-переходу областей при різних зворотних напругах на діоді.

$$J = q(pv_p + nv_n),$$

де v_p і v_n – швидкості або дифузії, або дрейфу дірок і електронів.

У діоді неосновні носії заряду дифундують до переходу, тому їх швидкості можна подати як дифузійні довжини, поділені на відповідний час життя. Замінімо загальні концентрації p і n на концентрації неосновних носіїв, оскільки з рухом саме неосновних носіїв заряду пов'язаний зворотний струм діода. Тоді:

$$J_{нас} = q \left(p_{n0} \frac{L_p}{t_p} + n_{p0} \frac{L_n}{t_n} \right). \quad (11.1)$$

Якщо врахувати співвідношення (10.8), а також практично повну іонізацію домішок при кімнатних температурах, то буде:

$$J_{нас} = qn^2 \left(\frac{L_p}{N_{dn} t_p} + \frac{L_n}{N_{an} t_n} \right). \quad (11.2)$$

Зі збільшенням температури діода густина струму насичення збільшується, оскільки з температурою експоненційно зростає власна концентрація носіїв.

У діодах на основі матеріалу з більшою шириною забороненої зони густина струму насичення повинна бути значно меншою, оскільки власна концентрація експоненційно зменшується зі збільшенням ширини забороненої зони. Порівнюючи германієві та кремнієві діоди та враховуючи різницю у власних концентраціях носіїв у германії та кремнії, яка становить три порядки, робимо висновок, що густина струму насичення для кремнієвих діодів повинна бути меншою на шість порядків.

Зі збільшенням концентрації домішок у прилеглих до переходу областях густина струму насичення відповідно до (11.2) повинна зменшитись.

11.2. Біполярні транзистори.

Біполярним транзистором називається напівпровідниковий прилад з двома взаємодіючими р-п-переходами і трьома виводами. Він має тришарову структуру, яка складається з областей з різними типами електропровідності, що чергуються між собою: р-п-р або п-р-п (рис. 11.5).

Робота біполярного транзистора залежить від носіїв заряду обох полярностей – електронів і дірок – звідси його назва "біполярний".

Основним елементом транзистора є кристал кремнію або германію зі створеними у ньому двома площинними р-п-переходами. Структура такого кристала р-п-р-типу, виготовленого за сплавною технологією, показана на рис. 11.6, а. Пластина напівпровідника п-типу з попередньо введеною в невеликій кількості донорною домішкою є базовою. На неї наплавляються з обох боків таблетки акцепторної домішки: для германію – індій, для

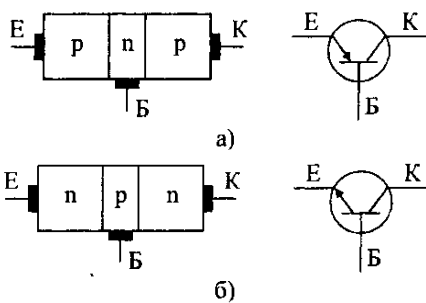


Рис. 11.5. Тришарові структури і умовні графічні позначення транзисторів типу р-п-р (а) і п-р-п (б).

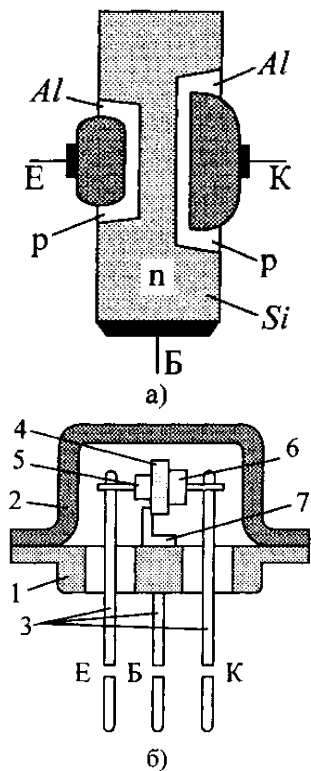


Рис. 11.6. Структура (а) і конструкція (б) сплавленого транзистора: 1 – діод корпусу; 2 – кришка корпусу; 3 – зовнішні виводи; 4 – кристал кремнію n-типу; 5,6 – таблетки алюмінію; 7 – кристалотримач.

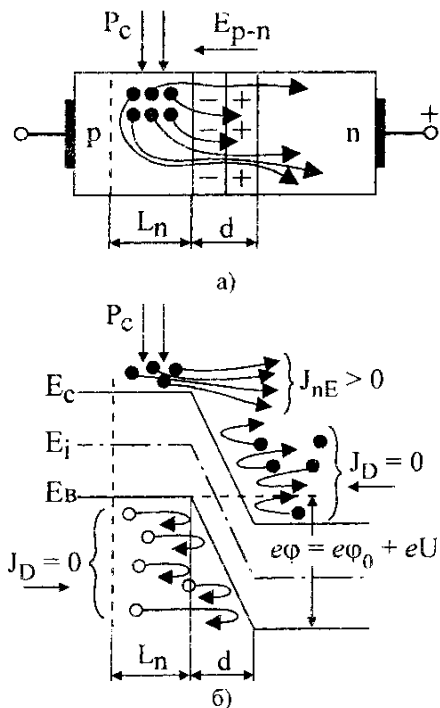


Рис. 11.7. Схема інжекції електронів в р-область на моделі (а) і на енергетичній діаграмі (б).

кремнію – алюміній. Під час термічної обробки атоми акцепторної домішки проникають у кристал, створюючи р-області. Між р-областями та напівпровідником n-типу створюються р-п-переходи. Процес введення домішок контролюється таким чином, щоб в одній р-області їхня концентрація була більшою (на рисунку – в лівій р-області), ніж в іншій. Найменша концентрація домішки залишається в середній області n-типу.

Зовнішня область з найбільшою концентрацією домішки називається емітером, інша зовнішня область – колектором, а внутрішня – базою. Електронно-дірковий перехід між емітером та базою називають емітерним переходом, а між колектором і базою – колекторним переходом.

Відповідно до концентрації основних носіїв заряду база є високоомною областю, колектор – низькоомною, а емітер – найбільш низькоомною. Товщина бази дуже мала і становить одиниці мікрметрів. Площа колекторного переходу в кілька разів перевищує площу емітерного. Приклад конструкції малопотужного германієвого транзистора наведено на рис. 11.6, б.

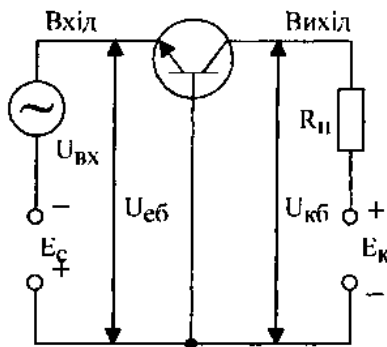
Принцип роботи. Принцип роботи біполярного транзистора полягає в зміні опору зворотозміщеного р-п-переходу за рахунок інжекції носіїв заряду (від англ. transistor-transfer resistor – перетворення резистора).

Опір зворотозміщеного р-п-переходу дуже великий – кілька мегаом і більше. Зворотозміщений р-п-перехід (рис. 11.7, а) має дуже великий опір лише для потоків основних носіїв заряду, а неосновні носії проходять його, "скачуючись з потенціального бар'єра", практично не зустрічаючи опору (рис. 11.7, б). Тому при досить високому рівні інжекції можна значно збільшити струм у зворотозміщеному р-п-переході і тим самим зменшити його опір. Інжекція може здійснюватися з різних джерел, за рахунок різних процесів струмопроходження і за різними схемами. Джерелами інжекттованих носіїв заряду можуть бути: плоский р-п-перехід, точковий контакт, гетероперехід, інверсний шар, потенціальний бар'єр тощо. Процесами струмопроходження при інжекції можуть бути: дифузія, дрейф, тунелювання, емісія Шоттки тощо.

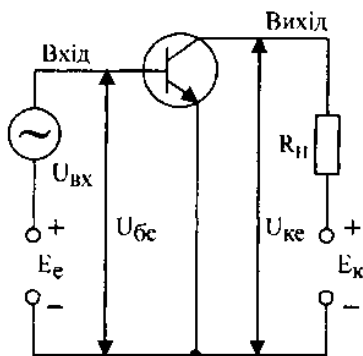
Схема інжекції визначає розміщення області введення інжекттованих носіїв заряду відносно зворотозміщеного р-п-переходу.

Таким чином, принцип підсилення, що полягає в зниженні опору зворотозміщеного р-п-переходу, може реалізовуватися в результаті різних можливих поєднань інжектуючого джерела, процесу струмопроходження та схеми інжекції.

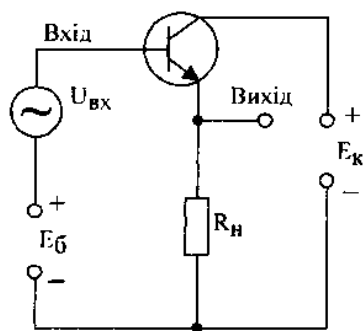
Схеми вмикання та статичні характеристики транзисторів. Транзистор має три електроди, з яких в схемі вмикання один – вхідний, другий – вихідний, а



а)



б)



в)

Рис. 11.8. Схеми включення транзистора типу *n-p-n*: а – зі спільною базою (СБ); б – зі спільним емітером (СЕ); в – зі спільним колектором (СК).

третій – загальний для кіл входу і виходу. Залежності від того, який електрод є спільним, можливі три схеми вмикання транзистора – зі спільною базою СБ, спільним емітером СЕ та спільним колектором СК (рис. 11.8).

При будь-якій схемі вмикання в кожному колі постійний струм проходить від плюса джерела живлення через відповідні області транзистора до мінуса джерела живлення. Стрілка емітера вказує напрямок струму, що проходить крізь нього.

У будь-якій схемі вмикання в кожному з двох кіл діє напруга між двома електродами і протікає струм: у вхідному колі – $U_{вх}$ і $I_{вх}$, у вихідному – $U_{вих}$ і $I_{вих}$. Ці електричні величини визначають режим роботи транзистора і взаємно впливають одна на одну.

Схема СБ характеризується, порівняно зі схемами СЕ та СК, найнижчим вхідним та найвищим вихідним опором. В схемі СБ коефіцієнт підсилення за струмом $K_I = I_{вих}/I_{вх} < 1$, коефіцієнт підсилення за напругою $K_U = U_{вих}/U_{вх} > 1$. Схема ОЕ характеризується відносно високими вхідним та вихідним опором $K_I > 1$, $K_U > 1$, а схема СК – найвищим вхідним та найнижчим вихідним опором $K_I > 1$, $K_U < 1$.

Схема СБ має добрі частотні характеристики, однак вона не забезпечує підсилення за струмом та має малий вхідний опір, тому найчастіше використовується в транзисторній техніці схема СЕ.

Існує чотири режими роботи біполярних транзисторів: нормальний активний, подвійної інжекції (насичення), відсікання та інверсний активний.

В нормальному активному режимі емітерний перехід вмикається в прямому, а колекторний – у зворотному напрямках. В режимі подвійної інжекції (насичення) емітерний та колекторний переходи ввімкнені в прямому напрямку. В режимі відсікання емітерний та колекторний переходи ввімкнені в зворотному напрямку. В інверсному активному режимі колекторний перехід увімкнений у прямому, а емітерний – у зворотному напрямках.

В нормальному активному та інверсному активному режимах транзистор працює як підсилювальний пристрій і принципової різниці між цими двома режимами не існує. В інверсному режимі слабколегований колектор не забезпечує достатньо високого коефіцієнта інжекції, в результаті чого різко знижується підсилення.

В інверсному режимі вираз для коефіцієнта передавання колекторного струму має вигляд:

$$a_I = \frac{I_e}{I_k} \Big|_{U_{еб}} \quad (11.3)$$

Характеристики транзистора являють собою залежність однієї з величин ($U_{вх}$, $I_{вх}$, $U_{вих}$, $I_{вих}$) при незмінній третій величині. Для підтримання однієї з цих електричних величин постійною в схему для зняття характеристик потрібно ввести лише джерела живлення;

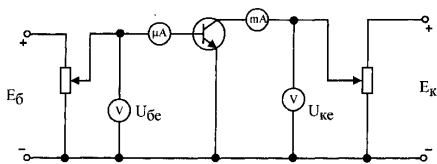


Рис. 11.9. Схема для знімання характеристик транзистора, включеного зі спільним емітером.

навантаження та джерело підсилюваних коливань не вводять. На рис. 11.9 наведено схему для одержання характеристик транзистора n-p-n, увімкненого з загальним емітером. При знятті кожної характеристики спочатку встановлюють за допомогою потенціометрів значення постійної величини, а потім, встановлюючи послідовно різні значення змінної величини, визначають відповідне значення іншої величини, залежної від неї.

Характеристики, зняті без навантаження, коли одна з величин підтримується постійною, називають статичними. Сукупність характеристик, знятих при різних значеннях цієї постійної величини, являє собою сім'ю статичних характеристик.

Найбільше значення при застосуванні транзисторів мають два види характеристик – вхідна та вихідна.

Вхідною характеристикою називають залежність вхідного струму від вхідної напруги при постійній вихідній напрузі:

$$I_{ex} = f(U_{ex}) \text{ при } U_{eux} = const.$$

Вихідною характеристикою називають залежність вихідного струму від вихідної напруги при постійному вхідному струмі:

$$I_{eux} = f(U_{eux}) \text{ при } I_{ex} = const.$$

Вид характеристик транзистора залежить від способу його вмикання, однак для схем СЕ та СК вони практично однакові, тому користуються звичайно вхідними та вихідними характеристиками для схем СБ та СЕ.

Статичні характеристики. Статичні характеристики визначають співвідношення між напругами та струмами в транзисторі.

Статичні характеристики схеми СБ. В схемі СБ керувальним струмом є струм емітера.

Вихідні (колекторні) характеристики записуються у вигляді залежності $I_c(U_{кб})$, вхідні (емітерні) характеристики – у вигляді залежності $I_e(U_{еб})|_{U_{кб}}$.

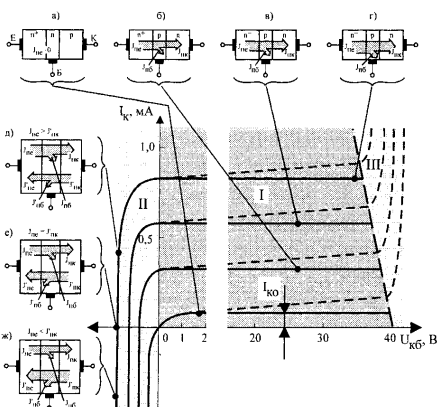


Рис. 11.10. Вихідні характеристики схеми СІВ, побудовані відповідно доз математичної моделі Еберса-Мола (суцільні лінії). Реальні характеристики (пунктирні лінії) (I, II, III-області режимів: нормального активного, насичення і лавинного пробою; $J_{не}$, $J_{нк}$, $J_{нб}$ – потоки електронів, інжектованих з емітера; $J'_{не}$, $J'_{нк}$, $J'_{нб}$ – потоки електронів, інжектованих з колектора).

На рис. 11.10 зображена сім'я вихідних характеристик СБ, побудованих на основі математичної моделі Еберса-Молла. На рисунку видно три режими роботи транзистора: нормальний активний (I), насичення (II), лавинного пробою (III). Активний режим використовується в режимах підсилення, режим подвійної інжекції – в імпульсних ключових схемах, режим лавинного пробою – в схемах на лавинних транзисторах, які не набули широкого розповсюдження внаслідок недостатньої стабільності характеристик. Якщо інжекція з емітера відсутня ($I_e = 0$), то залежність $I_c(U_{кб})$ є звичайною ВАХ p-n-переходу (рис. 11.3). Відмінність полягає в тому, що на рис. 11.10 зворотна гілка лежить в першому квадранті, а пряма – в третьому. В результаті інжекції з емітера при $U_{кб} > 0$, в колекторі виникає потік електронів $J_{нк} = \alpha_N J_{не}$, який створює в колекторному колі струм $I_c = \alpha_N I_e$. Струм колектора I_c (потік $J_{нк}$) зростатиме

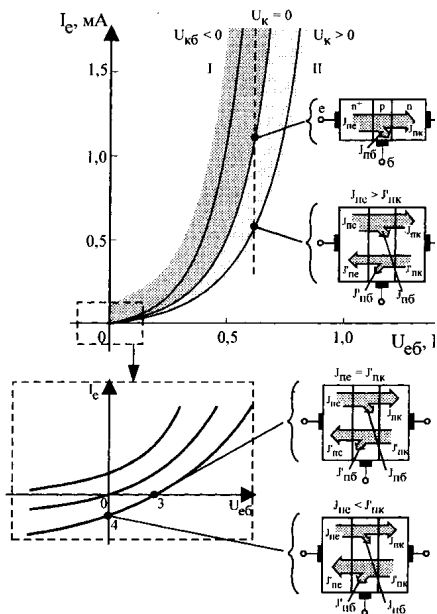


Рис. 11.11. Вхідні характеристики схеми СБ (I, II, III — області режимів: нормального активного, насичення і лавинного пробую: $J_{пк}$, $J_{нк}$, $J_{пб}$ — потоки електронів, інжектованих з емітера; $J'_{пк}$, $J'_{нк}$, $J'_{пб}$ — потоки електронів, інжектованих з колектора).

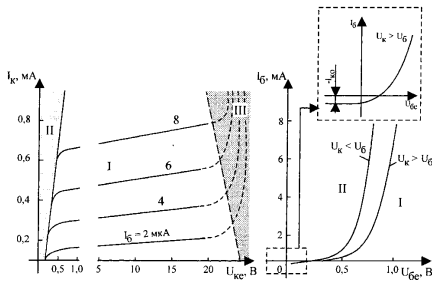


Рис. 11.12. Сім'я вихідних (а) і вхідних (б) характеристик схеми СЕ (I, II, III — області режимів: нормального активного, насичення і пробую колекторного переходу).

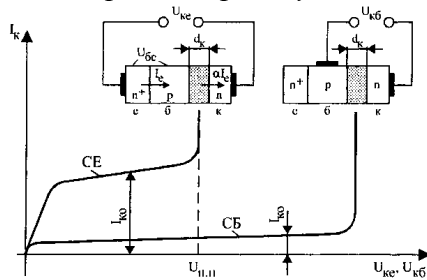


Рис. 11.13. Ілюстрація ходу вихідних характеристик, струмів насичення $I_{к0}$ і пробивних напруг в схемах СБ і СЕ.

пропорційно потоку інжектованих з емітера електронів (рис. 11.10, а-г).

При подаванні на колектор прямої напруги останній сам інжектуватиме зустрічний потік електронів (рис. 11.10, д-ж). Результуючий потік швидко зменшуватиметься до нуля зі зростанням прямої напруги на колекторі (рис. 11.10, д, е) і при подальшому збільшенні прямої напруги набуває зворотного напрямку (рис. 11.10, ж).

З рис. 11.10 видно, що в активному режимі колекторна напруга не впливає на вихідні характеристики схеми СБ. Такий висновок виправданий для більшості практичних випадків. Насправді з підвищенням зворотної колекторної напруги струм колектора в активному режимі буде слабо зростати (пунктирні криві на рис. 11.10) за рахунок збільшення зворотного колекторного струму $I_{ке}$ та коефіцієнта передавання емітерного струму α_N .

Зростання α_N пов'язане з ефектом Ірлі, який полягає в модуляції товщини $W_б$ та опору $R_б$ базової області транзистора при зміні колекторної напруги (рис. 11.11).

На рис. 11.11 наведені вхідні характеристики схеми СБ. При $U_{кб} = 0$ вхідна характеристика є звичайною прямою гілкою емітерного переходу. При подаванні на колектор прямої напруги транзистор переходить в режим насичення, при якому результуючий потік електронів з емітера, а отже, і струм емітера будуть зменшуватися (рис. 11.11, а, б), переходити через нуль (рис. 11.11, в) та змінювати напрямок (рис. 11.11, г.). Зростання струму емітера з підвищенням зворотної емітерної напруги викликане збільшенням градієнту концентрації електронів dn/dx в базі (ефект Ірлі при інверсному вмиканні).

Статичні характеристики схеми СЕ. В схемі СЕ керувальним струмом є струм бази.

Вихідні (колекторні) характеристики мають вигляд залежності $I_к(U_{ке})/I_б$, вхідні (базові) характеристики — у вигляді залежності $I_б(U_{бε})/U_{ке}$.

На рис. 11.12, а, б наведені відповідно вихідні та вхідні характеристики схеми СЕ.

Вихідні ВАХ мають такі особливості. Области нормального активного режиму та режиму насичення розміщені в одному (першому) квадранті, тобто режим подвійної інжекції в схемі СЕ виникає не при негативних потенціалах на колекторі, як у схемі СБ, а при малих позитивних потенціалах $U_к$. Це пояснюється тим, що в схемі з СЕ на базу подається позитивний потенціал, тому колекторний перехід відкривається при $U_к < U_б$. У схемі СЕ, порівняно зі схемою СБ, вихідні характеристики крутіші, струм насичення $I_{к0}$ більший, а пробивна напруга менша (рис. 11.13). Це пояснюється тим, що в схемі СЕ, на відміну від схеми СБ, вихідна напруга $U_{εк}$ частково

прикладена до емітерного переходу, тому з підвищенням $U_{ке}$ зростають $U_{об}$, I_e та $I_k = \alpha_N I_e$.

Вхідні ВАХ мають такі основні особливості. В режимі насичення при $U_k < U_{об}$ базовий струм істотно зростає порівняно з нормальним активним режимом, оскільки в цьому випадку струм бази утворюється в результаті інжекції з двох паралельно ввімкнених переходів – емітерного та колекторного.

В нормальному активному режимі базовий струм зменшується з потенціалом бази, доходить до нуля та змінює свій напрямок (див. вставку на рис. 11.12, б).

Вхідні та вихідні характеристики схеми СК аналогічні до характеристик схеми СЕ.

11.3. Модель Еберса-Молла.

Моделювання ставить своїм завданням встановлення зв'язків між фізичними параметрами та електричними характеристиками приладів. Особливо необхідне моделювання при розробці інтегральних мікросхем, коли за простими та точними моделями приладів вдається визначити поведінку складної схеми.

Модель Еберса-Молла (рис. 11.14) відзначається простотою та добре відображає відсутність принципової відмінності між обома переходами та їх оборотність, яка виявляється, наприклад, при роботі транзистора в інверсному активному режимі. Модель складається з двох діодів та двох джерел струму, ввімкнених зустрічно. Джерела струму керуються діодами. Джерело струму є аналогом джерела ЕРС. Ідеальне джерело ЕРС має нульовий внутрішній опір, джерело струму – нескінченно великий опір і "жорстко" задає струм в колі незалежно від його опору.

Струми інжектованих носіїв заряду позначаються через I_1 та I_2 , а струми носіїв, що збираються, – через $\alpha_N I_1$, та $\alpha_N I_2$.

На основі моделі Еберса-Молла можна скласти рівняння, що зв'язують емітерні, колекторні та базові струми з напругами на переходах. Такі рівняння є математичною моделлю транзистора, на основі якої проводять аналіз його статичних режимів.

Параметри транзисторів. Для оцінки властивостей транзисторів поряд з їх характеристиками використовують параметри. Розрізняють дві групи параметрів: первинні та вторинні.

До первинних належать власні параметри транзисторів, що характеризують його фізичні властивості (рис. 11.15, а) і не залежать від схеми вмикання: r_e – диференційний опір емітерного переходу в прямому напрямку (становить одиниці і десятки Ом); r_b – об'ємний опір бази (становить сотні Ом); r_k – диференційний опір колекторного переходу в зворотному напрямку (становить сотні кОм); C_e – ємність емітерного переходу (становить сотні пФ); C_k – ємність

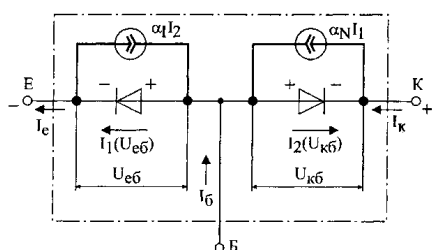
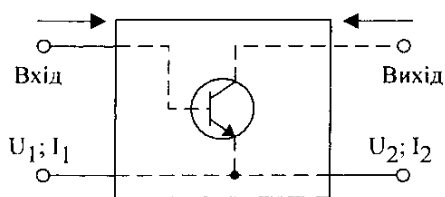
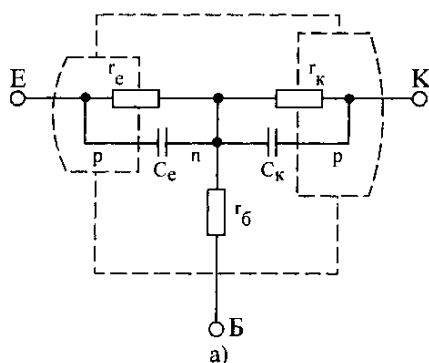


Рис. 11.14. Модель Еберса-Молла.



б)

Рис. 11.15. Структура транзистора, яка демонструє його первинні параметри r_e , r_k , r_b , C_e , C_k (а), і подання транзистора у вигляді чотириполюсника для визначення h -параметрів (б).

колекторного переходу (становить десятки пФ).

Впливом ємностей C_e і C_k в області звукових частот можна знехтувати.

Опори емітерного та колекторного переходів залежать від режиму транзистора і можуть визначатися як диференційні опори для даної робочої точки за статичними характеристиками транзистора в схемі СБ; опір емітерного переходу – за вхідною характеристикою як відношення малого приросту напруги емітера до викликаного ним приросту струму при постійній напрузі на колекторі:

$$r_e = \frac{\Delta U_e}{\Delta I_e} \text{ при } U_k = \text{const},$$

опір колекторного переходу – за вихідною характеристикою як відношення приросту напруги колектора до викликаного ним малого приросту струму колектора при сталому струмі емітера:

$$r_e = \frac{\Delta U_k}{\Delta I_k} \text{ при } I_e = \text{const}.$$

Первинні параметри r_e , r_b і r_k транзистора використовувати незручно через неможливість їх вимірювання за допомогою приладів, бо точки для підключення приладу містяться всередині структури транзистора.

До параметрів транзистора належать також диференційні коефіцієнти підсилення струму в трьох схемах увімкнення. Враховуючи їх залежність від режиму, коефіцієнти підсилення струму визначають як відношення приросту вихідного струму до викликаного ним малого приросту вхідного струму за даної незмінної вихідної напруги.

Для схеми СБ коефіцієнт підсилення струму a :

$$a = \frac{\Delta I_k}{\Delta I_e} \text{ при } U_{kб} = \text{const}; a=0,95-0,99.$$

Для схеми СЕ коефіцієнт підсилення струму β :

$$b = \frac{\Delta I_k}{\Delta I_b} \text{ при } U_{ке} = \text{const}; \beta=20-200.$$

Для схеми СК коефіцієнт підсилення струму γ :

$$g = \frac{\Delta I_e}{\Delta I_b} \text{ при } U_{ек} = \text{const}; \gamma=20-200$$

Коефіцієнти підсилення струму, що називаються також коефіцієнтами передавання струму, в різних схемах вмикання транзистора об'єднуються співвідношеннями:

$$a = \frac{b}{1+b}; a = \frac{g-1}{g};$$

$$b = \frac{a}{1-a}; b = g-1;$$

$$g = 1+b; g = \frac{1}{1-a}.$$

Коефіцієнти підсилення струму a і b можуть визначатися за вихідними характеристиками транзистора в схемах вмикання СБ та СЕ.

Суть вторинних параметрів можна пояснити, уявивши транзистор у вигляді активного чотириполюсника, що має два вхідних і два вихідних виводи та підсилює сигнал (рис. 11.15, б). Вхідні величини позначають індексом "1", а вихідні – індексом "2": I_1 і U_1 – вхідні струм і напруга, I_2 і U_2 – вихідні. Усі міркування справедливі за умови, що сигнали, тобто прирости ΔI_1 , ΔU_1 , ΔI_2 і ΔU_2 , малі.

Ці чотири величини взаємопов'язані і впливають одна на одну. Для розрахунків вибирають дві з них як незалежні, а дві інші величини є залежними змінними. Для них складається система з двох рівнянь, що зв'язують їх з незалежними величинами через коефіцієнти, якими можуть бути або лише опори, або лише провідності, або різні за розмірностями коефіцієнти. Ці коефіцієнти і є вторинними параметрами. Існує кілька варіантів вибору пари незалежних змінних. Відповідними будуть і варіанти вибору системи рівнянь, а значить і сукупності параметрів для цієї системи, названої системою параметрів. Існують різні системи параметрів: система Z -параметрів (Z має розмірність опору), y -параметрів (y має розмірність провідності), h -параметрів та інші.

Найбільшого розповсюдження при розрахунку транзисторних низькочастотних схем набули h -параметри. Їх перевага перед власними параметрами полягає в тому, що їх зручно визначати з допомогою вимірювань у схемі вмикання транзистора, причому для цього легко створити потрібні режими за змінним струмом: коротке замикання на виході, що відповідає умові $\Delta U_2 = 0$ (або $U_2 = const$), і ненавантажений режим на вході, відповідно, $\Delta I_1 = 0$ (або $I_1 = const$).

Для визначення h -параметрів складається система рівнянь, в якій незалежними змінними є ΔI_1 , і ΔU_2 :

$$\Delta U_1 = h_{11}\Delta I_1 + h_{12}\Delta U_2;$$

$$\Delta I_2 = h_{21}\Delta I_1 + h_{22}\Delta U_2.$$

В цій системі є чотири параметри з різною розмірністю: h_{11} , h_{22} , h_{21} , h_{12} .

Індекс параметра є поєднанням двох цифр, що означають відповідне коло: "11" (один-один) стосується вхідного кола; "22" (два-два) – вихідного, "21" (два-один) характеризує залежність вихідної величини від вхідної, а

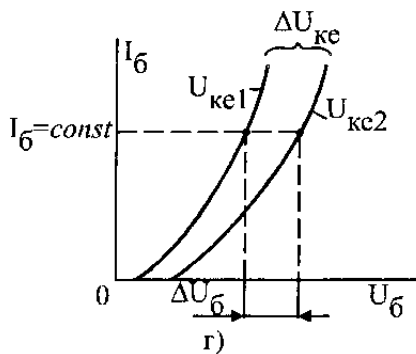
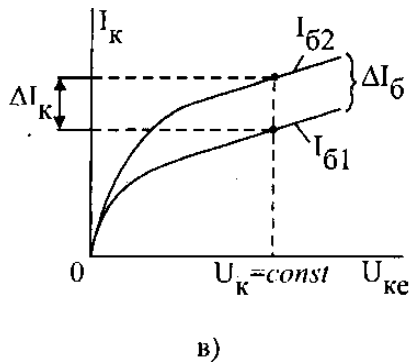
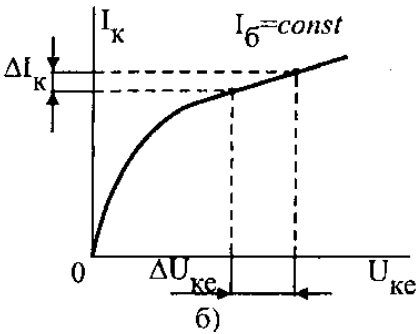
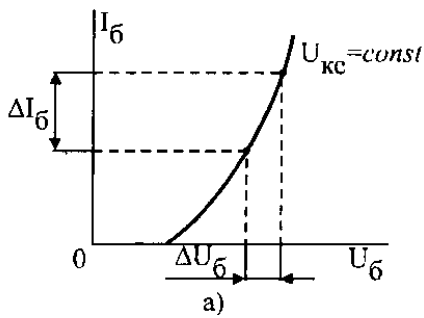


Рис. 11.16. Визначення h -параметрів за статичними характеристиками транзистора в схемі СЕ: а – h_{11e} ; б – h_{22e} ; в – h_{21e} ; г – h_{12e} .

"12" (один-два) – залежність вхідної величини від вихідної. Ці параметри мають такі значення:

h_{11} – вхідний опір транзистора при незмінній вихідній напрузі:

$$h_{11} = \frac{\Delta U_1}{\Delta I_1} \text{ при } U_2 = \text{const},$$

h_{22} – вихідна провідність транзистора при незмінному вхідному струмові:

$$h_{22} = \frac{\Delta I_2}{\Delta U_2} \text{ при } I_1 = \text{const},$$

h_{21} – коефіцієнт підсилення струму при незмінній вихідній напрузі:

$$h_{21} = \frac{\Delta I_2}{\Delta I_1} \text{ при } U_2 = \text{const}$$

h_{12} – коефіцієнт внутрішнього зворотного зв'язку при незмінному вхідному струмові:

$$h_{12} = \frac{\Delta U_1}{\Delta U_2} \text{ при } I_1 = \text{const}$$

Оскільки до системи h -параметрів входять опір, провідність та безрозмірні величини, їх іноді називають змішаними або гібридними параметрами. Ці параметри залежать від схеми вмикання транзистора і у різних схемах мають різні значення. Тому до індекса додають літеру, що означає схему вмикання: для схеми СБ параметри h_{11b} , h_{22b} , h_{21b} , h_{12b} ; для схеми СЕ – h_{11e} , h_{22e} , h_{21e} , h_{12e} ; для схеми СК додається буква "к".

Визначення h -параметрів за статичними характеристиками транзистора для схеми СЕ зображено на рис. 11.16, де h_{11} визначається за одною вхідною характеристикою, h_{22} – за одною вихідною, h_{12} – за двома вхідними, h_{21} – за двома вихідними. Враховуючи, що характеристики транзистора нелінійні і параметри залежать від режиму роботи, їх визначають для робочої точки за малими приростами струмів та напруг:

$$h_{11e} = \frac{\Delta U_{\bar{б}}}{\Delta I_{\bar{б}}} \text{ при } U_{ке} = \text{const} \text{ (рис. 11.16, а);}$$

$$h_{22e} = \frac{\Delta I_{\kappa}}{\Delta U_{ке}} \text{ при } I_{\bar{б}} = \text{const} \text{ (рис. 11.16, б);}$$

$$h_{21e} = \frac{\Delta I_{\kappa}}{\Delta I_{\bar{б}}} \text{ при } U_{ке} = \text{const} \text{ (рис. 11.16, в);}$$

$$h_{12e} = \frac{\Delta U_{\bar{б}}}{\Delta U_{ке}} \text{ при } I_{\bar{б}} = \text{const} \text{ (рис. 11.16, г);}$$

Значення h -параметрів для різних схем вмикання зв'язані співвідношеннями, з яких за h -параметрами однієї схеми можна знайти h -параметри іншої. Наприклад:

$$h_{11e} = h_{11\kappa} = \frac{h_{11\bar{б}}}{1 - h_{21\bar{б}}}; h_{22e} = h_{22\kappa} = \frac{h_{22\bar{б}}}{1 - h_{21\bar{б}}}$$

Крім того, h -параметри можна виразити через первинні параметри транзистора:

$$h_{11\bar{b}} = r_e + r_{\bar{b}}(1-a); h_{11e} = h_{11k} = r_{\bar{b}} + r_e(1-b);$$

$$h_{22\bar{b}} = \frac{1}{r_k}; h_{22e} = h_{22k} = \frac{1+b}{r_k};$$

$$h_{12\bar{b}} = \frac{r_{\bar{b}}}{r_k}; h_{12k} = 1; h_{12e} = \frac{r_e(1+b)}{r_k};$$

$$h_{21\bar{b}} = a; h_{21k} = g = 1+b = \frac{1}{1-a}; h_{21e} = b;$$

Як видно з наведених співвідношень, a і b відповідно дорівнюють $h_{21\bar{b}}$ і h_{21e} ; визначення їх за колекторними характеристиками для схем СЕ та СБ проводиться аналогічно показаному на рис. 11.16, в.

Транзистори характеризуються також величинами, що визначають номінальний і граничний режими роботи та можливості використання транзистора в різних пристроях. До них можна віднести зворотний струм колектора, зворотний струм емітера, граничну частоту підсилення струму, ємність колекторного переходу.

Граничні режими визначаються максимально допустимими значеннями струмів, напруг та розсіяної потужності, які не можна перевищувати за жодних умов експлуатації. До них належать максимально допустимі постійні струми колектора $I_{k \text{ макс}}$, емітера $I_e \text{ макс}$ і бази $I_b \text{ макс}$, максимально допустимі постійні напруги $U_{c\bar{b} \text{ макс}}$, $U_{k\bar{b} \text{ макс}}$, $U_{ke \text{ макс}}$, максимально допустима вхідна потужність $P_{k \text{ макс}}$, яка виділяється на колекторному переході.

Частотні властивості транзистора залежать від впливу ємностей C_e і C_k . Не зважаючи на те, що ємність емітерного переходу на порядок перевищує ємність колекторного, вплив C_k в області високих частот сильніший. Це пояснюється тим, що ємність C_e шунтує дуже малий опір емітерного переходу g_s , а ємність C_k - дуже великий опір r_k . Параметром транзистора, що характеризує його частотні властивості, є гранична частота, при якій коефіцієнт підсилення струму зменшується в разів $\sqrt{2}$. Для схеми СБ це частота f_a , при якій α зменшується в $\sqrt{2}$ разів, а f_β - те ж для β в схемі СЕ.

Оскільки струм бази в $(1 + \beta)$ разів менший від струму емітера, то $f_b = \frac{f_a}{1 + \beta}$.

Звідси видно, що гранична частота в схемі СЕ менша, ніж у схемі СБ, і частотні властивості транзистора гірші.

З розглянутого принципу дії транзистора та його характеристик в схемах з СБ та СЕ видно, що в схемі СБ вихідний струм I_k за величиною практично відповідає вхідному струмові I_c (схема не підсилює струму); у схемі СЕ вихідний струм I_k набагато перевищує вхідний струм $I_{\bar{b}}$ (відбувається підсилення струму). Слід додати, що в схемі

СК струм підсилюється практично так, як і в схемі СЕ, однак вона не дає підсилення за напругою. Очевидно, що найбільше підсилення потужності електричних коливань дає схема СЕ.

Лекція 12

ЕЛЕМЕНТИ ІНТЕГРАЛЬНИХ СХЕМ.

План

12.1. Активні елементи інтегральних схем

12.2. Транзистор системи метал-діелектрик-напівпровідник

12.3. Польові транзистори та прилади з зарядовим зв'язком

Література

12.1. Пасынков В. В., Чиркин Л. К. Полупроводниковые приборы. – М.: Высшая школа, 1987. – 479 с.

12.2 Гусев В. Г., Гусев Ю. М. Электроника. – М.: Высшая школа, 1991. – 622 с.

12.1. Активні елементи інтегральних схем

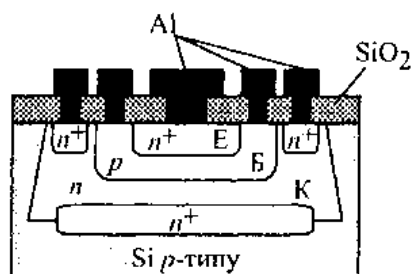
Як активні елементи інтегральних мікросхем використовують різні транзисторні структури, сформовані в кристалах кремнію методами планарної технології. Транзистори інтегральних мікросхем можуть відрізнятися не тільки за структурою, але і принципом дії.

Біполярний транзистор. Біполярний транзистор є поширеним активним елементом сучасних інтегральних мікросхем. Структура біполярного транзистора в інтегральних мікросхемах (інтегрального транзистора) відрізняється від структури дискретного транзистора ізоляцією від підшарку. Друга особливість пов'язана з тим, що вивід від колекторної області інтегрального транзистора здійснюється на верхній поверхні кристала. Тому для зменшення об'ємного опору області колектора перед епітаксійним нарощуванням проводять легування підшарку в тих місцях, де будуть сформовані транзисторні структури, тобто створюється прихований n^+ – шар (рис. 12.1).

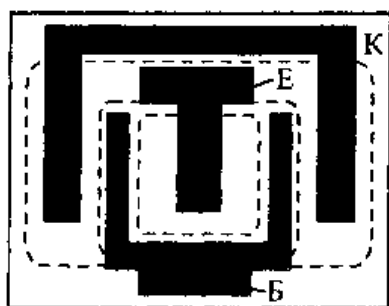
Інтегральний транзистор відрізняється від аналогічного дискретного транзистора, виготовленого за тією ж планарною технологією, значно меншими розмірами.

Основу біполярних інтегральних мікросхем складають транзистори n - p - n -типу, що пояснюється кращими параметрами інтегральних n - p - n -транзисторів порівняно з параметрами інтегральних транзисторів p - n - p -типу.

Це обумовлено тим, що для формування сильнолегованих емітерних областей транзисторів n - p - n -типу зазвичай використовують дифузію фосфору, який має вищу розчинність в кремнії та відносно низький коефіцієнт дифузії. Таким чином, для формування p - n - p -транзистора в інтегральній мікросхемі, яка містить n - p - n -транзистори, необхідне проведення додаткової дифузії якогось акцептора з граничною розчинністю, яка перевищує граничну розчинність фосфору. А такі



а)



б)

Рис. 12.1. Структура біполярного транзистора з прихованим n - шаром (а) і топологія електродів цього транзистора (б).

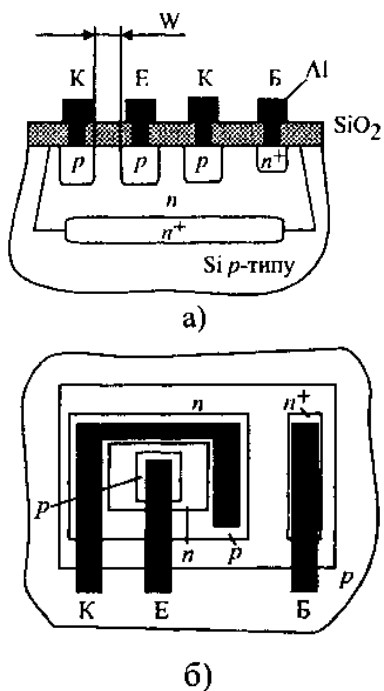


Рис. 12.2. Структура горизонтального транзистора p-n-p-типу (а) і топологія цього транзистора (б).

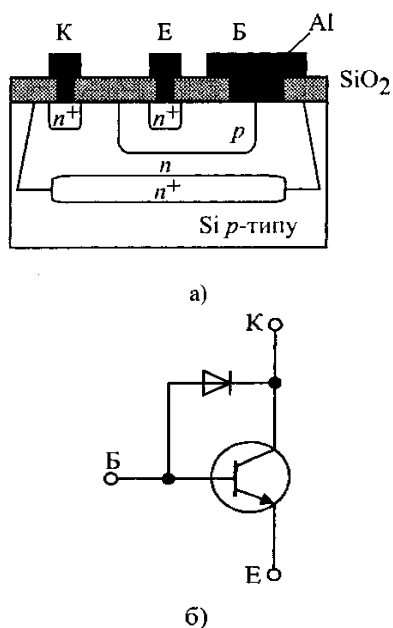


Рис. 12.3. Структура транзистора з діодом Шотткі (а) та його еквівалентна схема (б).

акцептори практично відсутні.

Тому основним прийнятним варіантом інтегрального транзистора p-n-p-типу є так званий горизонтальний або боковий транзистор (рис. 12.2). Горизонтальний транзистор повинен мати однакові пробивні напруги емітерного та колекторного переходів. Близькими повинні бути і коефіцієнти передачі струму емітера при нормальному та інверсному включенні такого транзистора, оскільки області емітера і колектора однакові за властивостями. Горизонтальна структура також дозволяє легко виконати багатоклекторний транзистор. Багатоклекторний транзистор є зручним для деяких цифрових інтегральних мікросхем.

Біполярний транзистор з діодом Шотткі. Біполярний транзистор у цифрових інтегральних мікросхемах зазвичай виконує функцію ключа і постійно працює або в режимі насичення, або в режимі відсікання. В режимі насичення відбувається нагромадження неосновних носіїв заряду в базі транзистора, а також у колекторній області. Процеси нагромадження неосновних носіїв і їх наступного розсмоктування при переведенні транзистора в режим відсікання або у вимкнений стан пов'язані з відносно повільним процесом дифузії неосновних носіїв заряду. Інерційність цих процесів визначає швидкість перемикавання транзистора з увімкненого стану в вимкнений і навпаки, тобто швидкість спрацювання схеми.

Для прискорення процесу перемикавання використовують шунтування колекторного переходу транзистора діодом Шотткі, тобто діодом з випрямним електричним переходом між металом і напівпровідником. Структура такого інтегрального транзистора та його еквівалентна схема показані на рис. 12.3.

12.2. Транзистор системи метал-діелектрик-напівпровідник

Принцип роботи транзистора зі структурою метал-діелектрик-напівпровідник (МДН) полягає на модуляції опору провідного каналу на поверхні напівпровідника під дією ефекту поля. Підсилювальний ефект цього транзистора оснований на тому, що в його збідненій області відбувається зміна густини заряду, за рахунок чого змінюється опір каналу, що з'єднує стік та витік приладу. Керувальним фактором слугує зовнішній потенціал, прикладений до заслону.

Залежно від типу електропровідності каналу розрізняють два типи МДН-транзисторів: з каналами p- і n-типу. Кожний з цих каналів може працювати в двох режимах: збіднення та збагачення. На рис. 12.4 наведено типові структури польових транзисторів, а на рис. 12.5 подані схеми їх включення.

У МДН-транзисторів заслін ізольований від напівпровідника шаром діелектрика і має додатковий

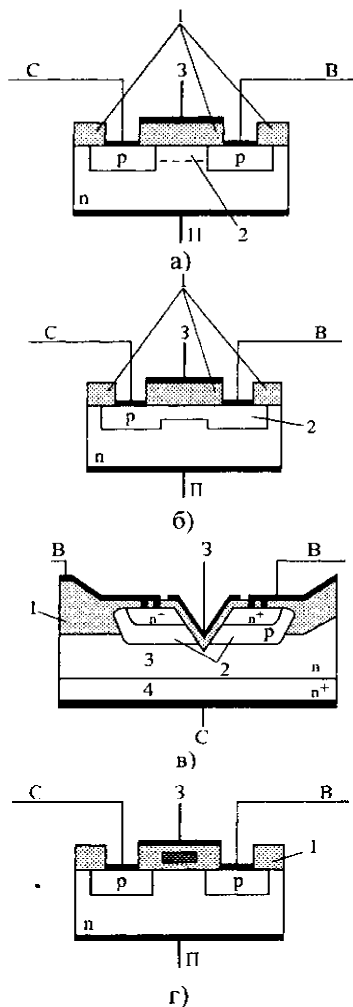


Рис. 12.4. Структури

МДН-транзистора: а – планарний транзистор індукованим каналом; б – планарний транзистор вбудованим каналом; в – транзистор (виготовлений за V-технологією, що забезпечує отримання каналів довжиною до 0,1 мкм), який працює до частот близько 1,5 ГГц; г – транзистор із затвором з полікристалічного кремнію, який має невисоку порогову напругу і високу швидкодію: 1 – діелектрик; 2 – канал; 3 – епітаксійний шар n-типу; 4 – підкладка n⁺-типу.

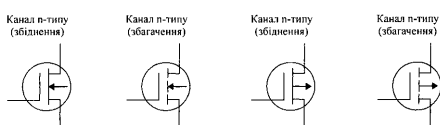


Рис. 12.5. Умовні графічні позначення різних МОН – транзисторів.

вивід від кристала, на якому сформовано прилад (рис. 12.4) – підшарку.

Для спрощення пояснення принципу роботи даного приладу обмежимося лише випадком МДН-транзистора з каналом n-типу, що працює в режимі збагачення. Такий прилад складається з кремнієвої підкладки p-типу, на якій методом іонної імплантації створено дві області n⁺-типу, що називаються виток (В) і сток (С). На ділянку поверхні напівпровідника завдовжки 1 нанесено шар оксиду SiO₂. Зовнішня поверхня цього шару покрита алюмінієвою плівкою, яка відіграє роль заслону (3). Четвертим виводом приладу є вивід підшарку (П).

МДН-транзистор працює так. При U_{ЗВ} = 0 поверхневий канал відсутній і провідність ділянки між витком і стоком дорівнює нулю. Транзистор залишається закритим незалежно від величини і полярності напруги стік-витік.

Це пояснюється тим, що при будь-якій полярності U_{СВ} обидва р-n-переходи (витік – підшарок та стік – підшарок) перебувають під зворотною напругою.

Позитивний потенціал заслону індукує в напівпровідник питомий заряд тим більший, чим більша питома ємність існує між металом та заслоном. Отже, питома ємність заслін-напівпровідник визначає керувальну здатність заслону і є його важливим параметром:

$$C_0 = \frac{\epsilon_H \epsilon_0}{d}$$

Зменшення товщини підзаслонного діелектрика обмежується його пробоем. Типові значення її становлять d = 0,7-012 мкм.

У вихідному стані в підзаслонному діелектрику звичайно присутній рівноважний заряд, що спричиняє вигин зон (рис. 12.6, а). Тому порогову напругу можна розділити на дві складові:

$$U_0 = U_{OF} + U_{OB}$$

Величина U_{OF} – напруга випрямлення зон, тобто напруга, яку потрібно прикласти до заслону для ліквідації вихідного викривлення зон (рис. 12.6, б):

$$U_{OF} = \frac{j_0 + Q_{OS}}{C_0}$$

де φ₀ – контактна різниця потенціалів між металом заслону та діелектриком; Q_{OS} – рівноважний питомий заряд поверхні.

На рис. 12.6, б зображено вихідне викривлення зон, протилежне тому, при якому утворюється канал.

Величина U_{OB} – напруга вигину зон, тобто напруга, яку необхідно прикласти до заслону для вигину зон у бік, необхідний для утворення провідного каналу (рис. 12.6, в):

$$U_{OB} = j_{Sm} + \frac{a}{C_0} \sqrt{j_{Sm}}$$

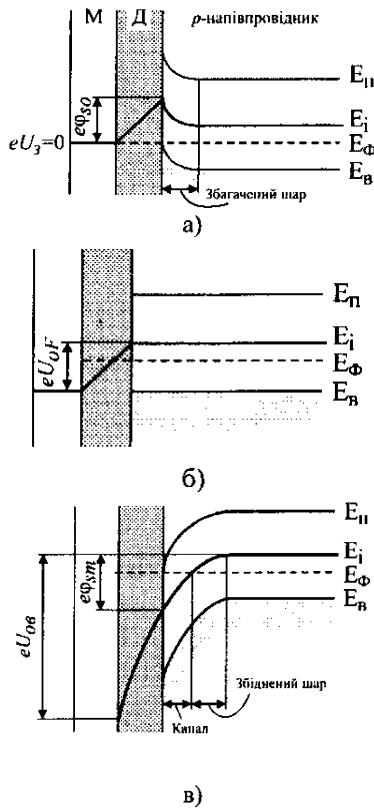


Рис. 12.6. Енергетичні діаграми МДН-транзистора: а — вихідний стан; б — стан після подачі напруги випрямлення зон U_{0F} ; в — стан після подачі напруги згину зон U_{0B} .

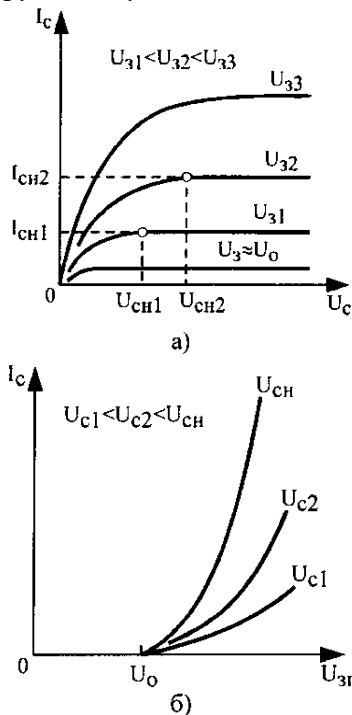


Рис. 12.7. Статичні характеристики МДН-транзистора: а — вхідні; б — передавальні.

де a – параметр, що характеризує вплив об'ємного заряду в підкладці.

З рис. 12.6, в видно, що для зменшення значення величини U_{0B} , а отже і порогової напруги U_0 , потрібно зменшувати рівень легування р-підшарку. Дійсно, зі зменшенням концентрації акцепторів у підшарку рівень E_F буде підніматися вгору, а відстань між рівнями E_i-E_F зменшуватися. Тому буде зменшуватися також напруга, яку потрібно прикласти для утворення інверсного каналу U_{0B} .

Статичні характеристики. Співвідношення між напругами та струмами в МДН-транзисторах визначаються за допомогою вихідних та передавальних характеристик (рис. 12.7, а, б).

На вихідні ВАХ істотно впливають зміни в структурі каналу, що виникають в результаті струмопроходження. Якщо напруги $U_B = U_C = 0$, то поле в діелектрику однорідне і товщина утвореного каналу буде однаковою по всій довжині l (рис. 12.7, а). Якщо напруга $U_{CB} > 0$ не дуже велика, то канал поводить себе як звичайний опір, при цьому струм буде збільшуватися пропорційно напрузі стоку. Ця область на вихідних ВАХ називається лінійною областю роботи транзистора.

Зі збільшенням напруги U_{CB} буде зростати струм стоку і потенціал поверхні напівпровідника в напрямку від витоку до стоку. Тому різниця потенціалів між заслоном та поверхнею напівпровідника буде зменшуватися в напрямку до стоку. Відповідно зменшується напруженість поля в діелектрику і питомий заряд електронів у каналі. В результаті переріз каналу починає звужуватися в напрямку до стоку. Після досягнення напругою стоку значення, рівного напрузі насичення U_{CH} в точці $x = 1$, стають рівними нулю різниця потенціалів між заслоном та поверхнею напівпровідника, напруженість поля в діелектрику та питомий заряд електронів. Тому довжина каналу виявляється рівною нулю (рис. 12.8, б). Ці умови відповідають початку режиму відсікання каналу (утворення "горловини" каналу).

Напруга насичення визначається виразом:

$$U_{CH} = U_{3B} - U_0 \quad (12.1)$$

При $U_{CB} > U_{CH}$ точка відсікання зсувається до витоку, і канал скорочується на величину Δl . При цьому збіднений шар зворотноміщеного переходу сток – підшарок, який при $U_{CB} < U_{CH}$ відокремлювався від поверхні каналом, виходить на поверхню напівпровідника на ділянці Δl (рис. 12.8, в). Потенціал у точці $x = 1$ зберігає значення U_{CH} яке було на початку насичення. Після відсікання каналу струм стоку практично перестає залежати від потенціалу стоку. Ця область на вихідних ВАХ називається областю насичення струму стоку.

Насичення струму стоку пояснюється так. В точці $x = l$ на вістрі каналу концентрується електричне поле,

напруженість якого $E > E_{кр}$. В результаті настає режим насичення швидкості дрейфу електронів, інжекттованих з вістря каналу в шар об'ємного заряду. Тому струм через канал $I_n = e\mu_n n E_{кр}$ при $v_{др} = \mu E_{кр} = const$ і $n = const$ перестає залежати від напруги.

Аналітичний вираз для вихідних ВАХ у першому наближенні має вигляд:

$$I_c = S_0 \left[(U_{3B} - U) U_{CB} - \frac{U_{CB}^2}{2} \right] \quad (12.2)$$

тут параметр

$$S_0 = m_n C_0 \frac{b}{l} = \frac{e_0 e_H b}{d l}$$

називається питомою крутизною МДН-транзистора.

Вираз (12.2) визначає струм стоку за умови $U_{CB} < U_{CH}$, тобто лише на початкових крутих ділянках вихідних характеристик (рис. 12.8, а).

При $U_{CB} > U_{CH}$ струм стоку практично не змінюється (рис. 12.8, а).

Підставляючи (12.1) в (12.2), одержимо вираз для вихідних ВАХ в області насичення:

$$I_{CH} = \frac{1}{2} \frac{m_n C_3}{l} (U_{3B} - U_0)^2 \quad (12.3)$$

Малосигнальні параметри. При роботі МДН-транзисторів у режимі підсилення використовуються ділянки вихідних ВАХ в області насичення. В цій області при оптимальних значеннях малосигнальних параметрів можливе одержання мінімальних нелінійних спотворень сигналів, що зазнають підсилення.

МДН-транзистори характеризуються такими малосигнальними параметрами: крутизною

$$S = \left. \frac{dI_c}{dU_{3B}} \right|_{U_{CB}} \quad (12.4)$$

внутрішнім опором

$$R_C = \left. \frac{dU_{CB}}{dI_c} \right|_{U_{3B}} \quad (12.5)$$

коефіцієнтом підсилення

$$K = \left. \frac{dU_{CB}}{dU_{3B}} \right|_{I_c} \quad (12.6)$$

Малосигнальні параметри зв'язані співвідношенням:

$$K = S R_C \quad (12.7)$$

Крутизна в області насичення визначається з виразу (12.3):

$$S = S_0 (U_{3B} - U_0) \quad (12.8)$$

При $(U_{3B} - U_0) = I_B$, $S = S_0$, тому параметр S_0 називають питомою крутизною. З (12.8) і (12.3) одержуємо залежність $S(I_c)$ у вигляді:

$$S = \sqrt{2S_0 I_c} \quad (12.9)$$

Внутрішній опір:

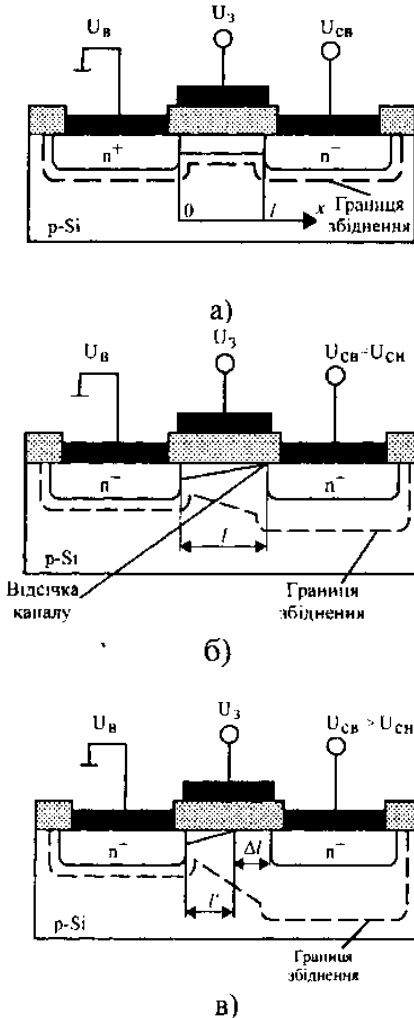


Рис. 12.8. Структура каналу і області об'ємного заряду МДН-транзистора: а - в лінійному режимі (низькі напруги на стоці); б - на початку насичення (відсікання каналу на границі зі стоком); в - у режимі насичення.

$$R_C = \frac{l}{I_C} \sqrt{\frac{2eN_A U_{CB}}{e_0 e_H}} \quad (12.10)$$

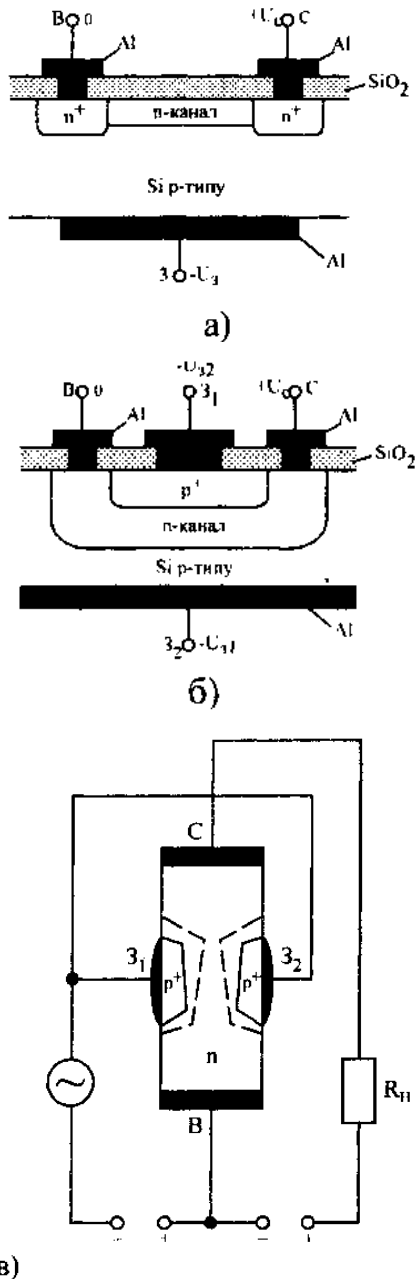
12.3. Польові транзистори та прилади з зарядовим зв'язком

Польовий транзистор – це напівпровідниковий прилад, підсилювальні властивості якого зумовлені потоком основних носіїв, що протікають через провідний канал і керовані електричним полем.

Польовий транзистор з керувальним переходом – це польовий транзистор, керування потоком основних носіїв у якому відбувається за допомогою випрямного електричного переходу, зміщеного у зворотному напрямку.

Випрямним електричним переходом, за допомогою якого здійснюється керування потоком основних носіїв заряду в польовому транзисторі, може бути р-п-перехід, гетероперехід або випрямний перехід Шотткі. Польові транзистори з керувальним гетеропереходом виготовляють, в основному, методом епітаксії сполук типу $A^{III} B^V$ на кристали арсеніду галію. Випрямні переходи Шотткі створюють нанесенням металу на кристали кремнію, арсеніду галію або різні гетероструктури. Однак найпоширенішими є польові транзистори з керувальним р-п-переходом у кристалах кремнію. Як приклад розглянемо польові транзистори з керувальним переходом, оскільки принцип дії польових транзисторів з різними керувальними переходами ідентичний.

Польовий транзистор з керувальним р-п-переходом має два омичні переходи до області напівпровідника, якою проходить керований або регульований потік основних носіїв заряду, і один або два керувальні р-п-переходи, зміщені у зворотному напрямку (рис. 12.9). При зміні зворотної напруги на керувальному р-п-переході змінюється його товщина, а отже, товщина області, якою проходить керований потік основних носіїв заряду. Область напівпровідника, в якій регулюється потік основних носіїв заряду, називають провідним каналом. Електрод польового транзистора, через який у провідний канал надходять носії заряду, називають витоком. Електрод польового транзистора, через який з каналу виходять носії заряду, називають стоком. Електрод польового транзистора, на який подають сигнал, називають заслоном. Провідний канал може мати електропровідність як n-, так і р-типу. Відповідно розрізняють польові транзистори з n-каналом і р-каналом. Усі полярності напруг зміщення, що подаються на електроди польових транзисторів з n- і р-каналом, протилежні. Польові транзистори з провідними каналами з різними типами електропровідності, називають комплементарними, тобто доповнювальними один одного за типом електропровідності провідних каналів. Комплементарні транзистори є зручними елементами для



в)
Рис. 12.9. Структури польових транзисторів з одним керувальним р-п-переходом (а), з двома керувальними р-п-переходами (б) і принципова схема вмикання зі спільним витоком (в).

побудови низки схем.

Керування струмом стоку, тобто струмом від зовнішнього відносно потужного джерела живлення у ланцюзі навантаження, відбувається при зміні зворотної напруги на р-n-переході заслону (або на двох р-n-переходах одночасно). У зв'язку з малими зворотними струмами потужність, необхідна для керування струмом стоку і використана від джерела сигналу в колі заслону, є надто малою. Тому польовий транзистор може забезпечити підсилення електричних сигналів як за потужністю, так і за струмом та напругою.

Отже, польовий транзистор за принципом дії аналогічний вакуумному триоду. Витік у польовому транзисторі подібний до катода вакуумного триода, заслін – до сітки, стік – до анода. В той же час польовий транзистор суттєво відрізняється від вакуумного триода. По-перше, для роботи польового транзистора не потрібне підігрівання катода. По-друге, будь-яку з функцій витоку і стоку може виконувати кожний з цих електродів. По-третє, польові транзистори можуть бути виготовлені як з n-каналом, так і з р-каналом, що дозволяє успішно поєднувати ці два типи польових транзисторів у схемах.

Від біполярного транзистора і, в тому числі, від одноперехідного транзистора польовий транзистор відрізняється, по-перше, принципом дії: у біполярному транзисторі керування вихідним сигналом здійснюється вхідним струмом, а в польовому транзисторі – вхідною напругою або електричним полем. По-друге, польові транзистори мають більші вхідні опори, зумовлені зворотним зміщенням р-n-переходу заслону в даному типі польових транзисторів. По-третє, польові транзистори можуть мати низький рівень шуму (особливо на низьких частотах), оскільки в польових транзисторах не використовується явище інжекції неосновних носіїв заряду, і канал польового транзистора відокремлений від поверхні напівпровідникового кристала (рис. 12.9, б, в). Процеси рекомбінації носіїв у р-n-переході і в базі біполярного транзистора, а також генераційно-рекомбінаційні процеси на поверхні кристала напівпровідника супроводжуються виникненням низькочастотних шумів.

Польові транзистори з ізолюваним заслоном. Польовий транзистор з ізолюваним заслоном, в якому як ізоляційний шар між металічним заслоном і провідним каналом використано діелектрик, називають польовим транзистором типу метал-діелектрик-напівпровідник або МДН-транзистор.

Випрямні електричні переходи під стоком і витоком можуть бути виконані не тільки в вигляді р-n-переходів, але й у вигляді випрямних переходів Шотткі.

Існує два різновиди МДН-транзисторів: з індукованим каналом і з вбудованим каналом (рис. 12.10).

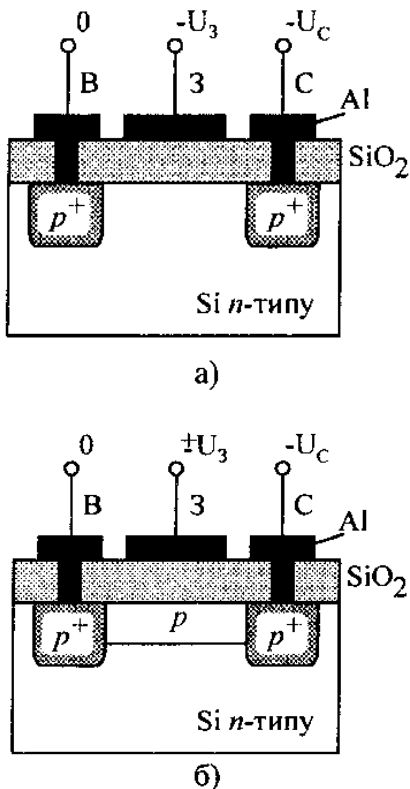


Рис. 12.10. Структура польових транзисторів з ізолюваним заслоном (з р-каналом): а – з індукованим каналом; б – з вбудованим каналом.

В МДН-транзисторах з індукованим каналом (рис. 12.10, а) провідний канал між сильнолегованими областями витоку і стоку і, відповідно, помітний струм стоку з'являються тільки при певній полярності і при певному значенні напруги на заслоні відносно витоку, яку називають пороговою напругою ($U_{3пор}$).

В МДН-транзисторах з вбудованим каналом (рис. 12.10, б) поблизу поверхні напівпровідника під заслоном при нульовій напрузі на заслоні відносно стоку існує інверсний шар – канал, який з'єднує витік зі стоком.

Зображені на рис. 12.10 структури польових транзисторів з ізолюваним заслоном мають підшарок з електропровідністю n-типу. Тому сильнолеговані області під витоком і стоком, а також індукований і вбудований канали мають електропровідність р-типу. Якщо ж аналогічні транзистори створені на підшарку з електропровідністю р-типу, то канал у них буде мати електропровідність n-типу.

МДН-транзистори з індукованим каналом. При нульовій напрузі на заслоні відносно витоку і при наявності напруги на стоці струм стоку буде безмежно малим. Це буде зворотний струм між підшарком і сильнолегованою областю стоку. При від'ємному потенціалі на заслоні (для структури на рис. 12.10) в результаті проникнення електричного поля через діелектричний шар до напівпровідника при малих напругах на заслоні (менших порогової напруги) біля поверхні під заслоном виникають збіднений основними носіями заряду шар і область об'ємного заряду, яка складається з іонізованих некомпенсованих домішкових атомів. При напругах на заслоні, вищих за порогову, поблизу поверхні напівпровідника під заслоном виникає інверсний шар, який і є провідним каналом між витоком і стоком. При зміні напруги на заслоні змінюється концентрація носіїв заряду в провідному каналі, а також товщина або поперечний переріз провідного каналу. Основною причиною модуляції опору провідного каналу в МДН-транзисторах з індукованим каналом є зміна концентрації носіїв заряду в провідному каналі; в польових транзисторах з керувальним переходом – зміна товщини або поперечного перерізу каналу.

МДН-транзистори з вбудованим каналом. Провідний канал під заслоном МДН-транзистора може створюватися за допомогою локальної дифузії або іонної імплантації відповідної домішки в приповерхневий шар підшарку. Він може виникнути через перерозподіл домішок поблизу поверхні напівпровідникового підшарку в процесі термічного окислення її поверхні. Також провідний канал може виникнути під заслоном внаслідок фіксованого заряду в підзаслонному шарі діоксиду кремнію, на поверхневих енергетичних рівнях, а також внаслідок

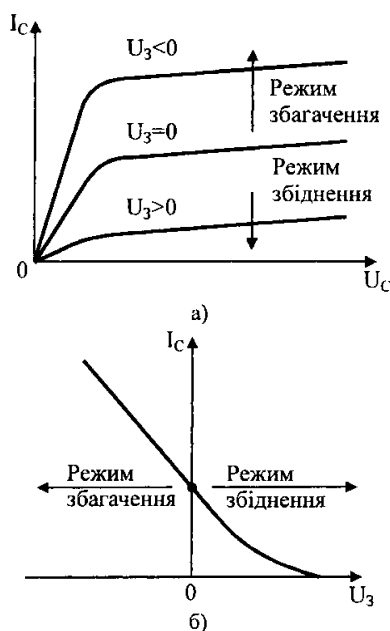


Рис. 12.11. Вихідні статичні характеристики (а) і статичні характеристики передачі (б) МДН-транзистора з вбудованим р-каналом.

контактної різниці потенціалів між металом заслону і напівпровідником підшарку.

Модуляція опору провідного каналу МДН-транзистора може відбуватися при зміні напруги на заслоні як додатної так і від'ємної полярності. Таким чином, МДН-транзистор з вбудованим каналом може працювати в двох режимах: у режимі збагачення і в режимі збіднення каналу носіями заряду. Ця особливість МДН-транзисторів з вбудованим каналом відображається і на зміщенні вихідних статичних характеристик при зміні напруги на затворі і його полярності (рис. 12.11, а).

Статичні характеристики передачі (рис. 12.11, б) виходять з точки на осі абсцис, яка відповідає напрузі відсікання $U_{звідс}$, тобто напрузі між заслоном і витокком МДН-транзистора з вбудованим каналом, який працює в режимі збіднення, коли струм стоку досягає заданого низького значення.

Лекція 13

МАГНІТОРЕЗИСТОРИ.

План

- 13.1. Загальна характеристика магніторезисторів
- 13.2. Матеріали для магніторезисторів
- 13.3. Магніторезистори на основі магнітоконцентраційного ефекту
- 13.4. GMR-магніторезистори
- 13.5. Параметри магніторезисторів

Література

- 13.1. Датчики. Справочник \ под ред. З.Ю.Готры и О.И.Чайковского. – Львов: Каменяр, 1995. -312с.
- 13.2.Хомерики О.К. Полупроводниковые преобразователи магнитного поля. – М.: Энергоатомиздат, 1986.-136с
- 13.3.Вайсс Г. Физика гальваномагнитных приборов и их применение: Пер. с нем. – М.: Энергия, 1974.-384с.

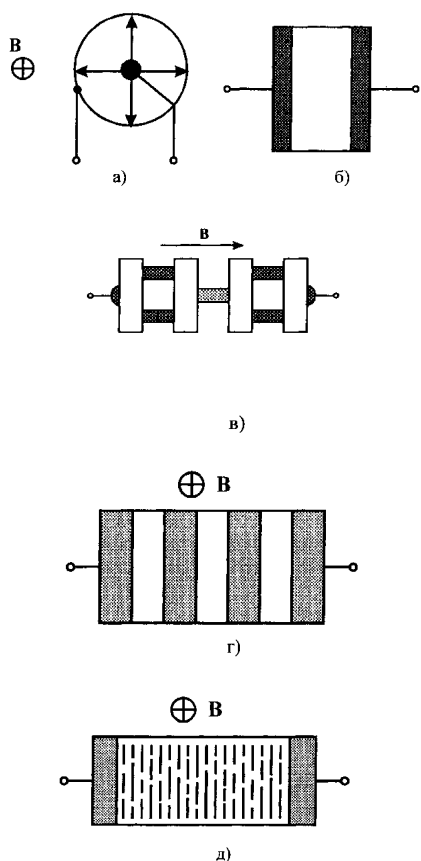


Рис. 13.1. Конструкції магніторезисторів.

13.1. Загальна характеристика магніторезисторів

Магніторезисторами називають напівпровідникові резистори, опір яких змінюється в магнітному полі. Оскільки ефект магнітоопору максимальний у напівпровіднику, не обмеженому в напрямку, перпендикулярному до струму, то в реальних магніторезисторах прагнуть максимально наблизитися до цієї умови. Крім цього, напруга холлівського електричного поля зменшує магніторезистивний ефект, тому конструкція магніторезистора повинна бути такою, щоб максимально зменшити або повністю усунути ефект Холла.

Найкращим наближенням до необмеженого за шириною зразка є диск Корбіно (рис. 13.1, а). За відсутності магнітного поля струм у такому зразку проходить у радіальному напрямку від центра диска до другого електрода, який знаходиться на периметрі диска. Відхилення носіїв під дією магнітного поля відбувається в напрямку, перпендикулярному до радіуса, а оскільки не існує граней, на яких може відбуватися накопичення заряду, то розділення носіїв заряду і утворення електричного поля Холла в цьому зразку не відбувається. Іншою конструкцією магніторезистора, хоча і з меншим проявленням магніторезистивного ефекту, ніж у диску Корбіно, є пластина напівпровідника з шириною, набагато більшою від її довжини (рис. 13.1, б).

Ці дві структури мають найбільшу відносну зміну опору в магнітному полі (рис. 13.2). Однак, їх суттєвим недоліком є низьке абсолютне значення вихідного опору при $B = 0$, що зумовлено їх конфігурацією. Для збільшення вихідного опору використовують послідовне

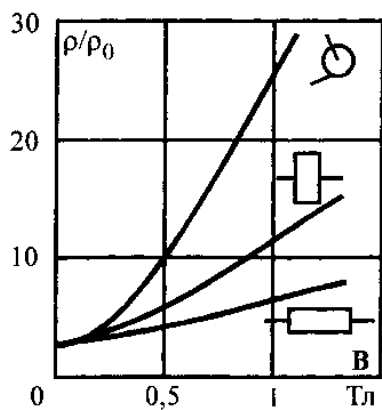


Рис. 13.2. Відносна зміна опору в магнітному полі.

з'єднання кількох магніторезисторів. Послідовне з'єднання дисків Корбіно наведене на рис. 13.1, в. Недоліком такої структури є велика довжина в напрямку магнітного поля, що вимагає використання великого зазору, якщо помістити магніторезистор між полюсами магніту і, відповідно, до послаблення поля.

Цього недоліку не має структура, наведена на рис. 13.1, г, де замість послідовного з'єднання магніторезисторів використовується одна довга напівпровідникова пластина, на поверхню якої нанесені металеві провідні смужки, що розділяють пластину на області, довжина яких менша від їх ширини. Отже, кожна область між смужками є окремим магніторезистором. Можна також вважати, що металеві смужки зменшують вплив ефекту Холла, і, відповідно, напруги Холла, що збільшує магнітоопір. Замість пластини з металевими смужками можна використовувати матеріал, у якому області з високою електропровідністю створені при вирощуванні кристала. В цьому випадку відпадає необхідність у нанесенні металевих смужок.

Магніторезистори такого типу створені на основі кристалів InSb з добавкою 1,8 % NiSb. Включення NiSb утворюють в кристалі InSb голки з питомим опором, майже вдвічі меншим, ніж питомий опір самого кристала. Магнітоопір такого матеріалу не залежить від форми зразка, необхідно тільки, щоб напрям голок був перпендикулярний до напрямку осі і магнітного поля (рис. 13.1, д). Крім того, антимонід нікелю не розчиняється у затверділому розчині антимоніду індію і не має легуючої дії на властивості останнього. Завдяки цьому з'являється можливість вводити домішки в основний матеріал InSb і досягати необхідної рухливості електронів. Як легуючий матеріал використовують телур.

13.2. Матеріали для магніторезисторів

Для створення магніторезисторів необхідно використовувати напівпровідниковий матеріал з високою рухливістю носіїв заряду. Матеріалом з високою провідністю можна вважати InSb ($\mu_n = 7 - 8 \text{ м}^2/\text{В}\cdot\text{с}$). Виготовлені з такого матеріалу магніторезистори мають коефіцієнт відносної зміни опору до 100 %/Тл, температурний коефіцієнт опору 1 %/К. При індукції магнітного поля до 0,3 – 0,5 Тл залежність квадратична, а при більших B – лінійна. Магніторезистори на основі InSb-NiSb мають чутливість вдвічі більшу, ніж магніторезистори з InSb при аналогічній залежності $R(B)$. Максимальну магніточутливість мають магніторезистори з InSb з власною провідністю. Однак вони мають також найбільший температурний коефіцієнт, оскільки концентрація носіїв заряду у власному напівпровіднику найсильніше змінюється з температурою. Залежність рухливості від температури має максимум при температурі $T = 77 \text{ К}$, тому магніточутливість при цій

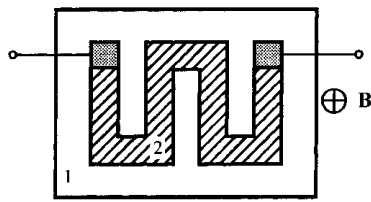


Рис. 13.3. Магнітрезистор у вигляді меандра.

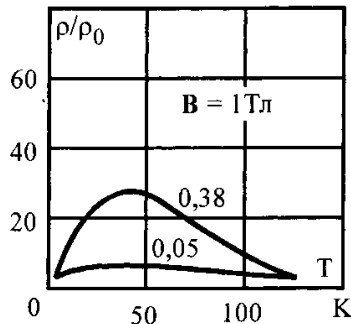


Рис. 13.4. Температурні залежності опору магнітрезистора на основі сполуки кадмій-ртуть-телур у різних магнітних полях.

температурі максимальна і зменшується в обидва боки так, що при гелієвій і кімнатній температурі вона однакова і в 4 – 5 разів менша від максимальної.

Для створення магнітрезисторів також використовувався твердий розчин $(\text{HgTe})_{1-x} + (\text{CdTe})_x$. Залежно від складу і температури цього матеріалу змінюється ширина забороненої зони і рухливість носіїв. Найбільшу рухливість мають розчини з шириною забороненої зони, близькою до нуля. Відповідно, такі магнітрезистори мають вищу магніточутливість, однак вихідний опір у них виявляється меншим. Для збільшення вихідного опору магнітрезистори виготовляються у вигляді меандра (рис. 13.3). З рис. 13.4 видно, що чутливість таких резисторів ($x = 0,1$) має максимум при температурах 20 ...100 К і залежить від магнітної індукції.

13.3. Магнітрезистори на основі магнітоконцентраційного ефекту

Магнітне поле, спрямоване перпендикулярно до струму через зразок, приводить до перерозподілу неосновних і основних носіїв заряду в об'ємі напівпровідника і утворення нерівноважної електронно-діркової плазми. Якщо швидкість рекомбінації нерівноважних носіїв у місці накопичення плазми більша, ніж в середньому по зразку, то це приводить до посилення рекомбінації і зменшення середньої концентрації носіїв у зразку, тобто до збільшення опору зразка. Область з високою швидкістю рекомбінації переважно створюється на поверхні напівпровідника. Відносна зміна опору (магніточутливість) зростає при наближенні провідності напівпровідника до власної внаслідок збільшення концентрації електронно-діркової плазми порівняно з концентрацією носіїв, яка зумовлена іонізацією домішки. Область відхилення концентрації носіїв від рівноважної простягається вглиб напівпровідника на відстань, трохи більшу від довжини дифузійного зміщення $L_{n,p}$. Тому чутливість максимальна при $d \approx 1,4 \cdot L_{n,p}$, де d – розмір зразка в напрямку дії сили Лоренца.

Зі всіх матеріалів германій найкраще підходить для створення резисторів на основі магнітоконцентраційного ефекту. При провідності, близькій до власної, він поєднує достатньо високу концентрацію електронно-діркової плазми з високим значенням $L_{n,p}$ (до 3 мм). Германієві резистори промислового виготовлення мають розмір $8 \times 0,6 \times 0,2$ мм. На кінцях напівпровідника створюються струмові омичні контакти і на одній з бокових граней ($8 \times 0,2$) – область з високою швидкістю рекомбінації. Магнітне поле спрямоване перпендикулярно до грані $8 \times 0,6$. Магнітрезистор має такі параметри при номінальному струмі живлення 3 мА: початковий опір 24 кОм, вольтова магніточутливість 50 мВ/мТл, температурний коефіцієнт чутливості в інтервалі

температур $+10...(-35) \text{ }^\circ\text{C}$ не більше ніж $0,5 \text{ } \%/^\circ\text{C}$, частотний діапазон $0... 10^3 \text{ Гц}$.

Функціональні можливості такого магнітрезистора можна суттєво покращити створенням польового електрода (МДН-типу) на боковій поверхні. При зміні напруги на польовому електроді змінюється швидкість рекомбінації на прилеглій поверхні напівпровідника і, відповідно, створюється можливість керування величиною магніточутливості.

13.4. GMR-магнітрезистори

Наприкінці 80-х років був виявлений ефект зміни опору майже на 50% під дією магнітного поля, який виникає в блоках, складених з тонких шарів заліза і хрому при низьких температурах. Оскільки ці блоки шарів, виготовлених із заліза та інших металів, взаємодіють з магнітним полем, то вони були названі Giant Magneto Resistor (гігантські магнітрезистори).

Товщина окремих шарів знаходиться в межах кількох нанометрів. Немагнітна мідь розділяє між собою окремі шари заліза або кобальту (рис. 13.5). Це розділення є настільки тонким, що шари кобальту зчеплюються і в результаті утворюється штучний антиферромагнетик. Зовнішні шари виготовлені з магнітом'яких матеріалів, намагнічених у напрямку дії зовнішнього магнітного поля, при цьому шари кобальту не змінюють напрямку намагніченості під дією зовнішнього поля.

Спіни електронів у зовнішніх шарах також впорядковуються в напрямку дії зовнішнього поля. Середнє значення довжини вільного пробігу електронів зі спінами, паралельними до напрямку намагнічування штучних антиферромагнетиків, значно перевищує товщину шару в блоці. Тому ці електрони практично не розсіюються. З іншого боку, електрони з антипаралельним спіном практично повністю розсіюються всередині блока і суттєво збільшують його електричний опір. Найменшого розсіювання електрони зазнають тоді, коли зовнішні і внутрішні шари намагнічені в одному напрямку, в цьому випадку опір зразка мінімальний. Максимальний опір спостерігається, коли намагніченості зовнішніх і внутрішніх шарів знаходяться під кутом 180° . Отже, зміна опору зразка залежить не від значення магнітного поля, а від зміни напрямку дії поля, оскільки тільки кут між напрямками намагніченості визначає сумарний опір системи (рис. 13.6).

Ефективність GMR-резисторів $\Delta R/R_0$ зменшується при збільшенні температури, оскільки термічне збудження коливань ґратки і спінових хвиль приводить до збільшення опору R_0 .

Надтонкі шари, в яких відбувається антиферромагнітна взаємодія, вимагають сучасних технологій створення цих шарів. Так, фірмою Siemens випускаються GMR-

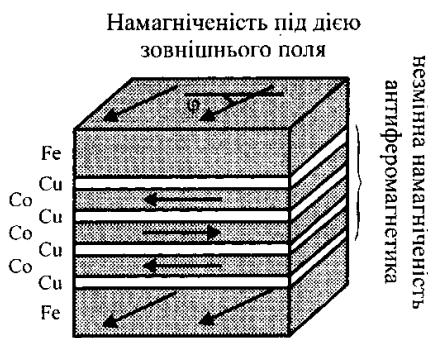


Рис. 13.5. Структура GMR-резистора.

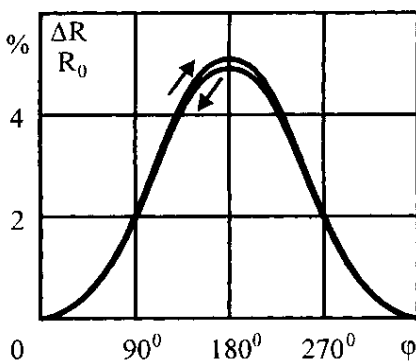


Рис. 13.6. Зміна опору при зміні напрямку зовнішнього магнітного поля.

резистори з 11 шарами завтовшки 25 нм. Магнітні кобальтові і немагнітні мідні шари формують штучний антиферромагнетик. Як зовнішнє покриття використовуються магнітом'які шари заліза. Початковий опір становить не менше ніж 700 Ом і відносна зміна опору становить не менше ніж 4%.

13.5. Параметри магніторезисторів

Основними параметрами магніторезистора є:

- початковий опір R_0 при $B = 0$;
- магніторезистивне відношення R_B/R_0 ;
- магнітна чутливість S_R ;
- навантажувальна здатність;
- залежність опору від температури.

Так, значення початкового опору залежить від провідності основного матеріалу y , яка, в свою чергу, визначається ступенем легування, шириною магніторезистора (для меандрової конструкції – шириною смуги, оптимальне значення 80 мкм), товщини магніторезистора – оптимальне значення 25 мкм, довжини магніторезистора, яка залежить від необхідного значення опору. Залежно від провідності основного матеріалу y для меандрової форми отримують такі значення опору на 1 мм^2 площі магніторезистора:

- для $y = 2,0 \cdot 10^4 \text{ Ом}^{-1} \cdot \text{м}^{-1}$ – приблизно 100 Ом;
- для $y = 5,5 \cdot 10^4 \text{ Ом}^{-1} \cdot \text{м}^{-1}$ – приблизно 40 Ом;
- для $y = 8,0 \cdot 10^4 \text{ Ом}^{-1} \cdot \text{м}^{-1}$ – приблизно 25 Ом.

Відхилення R_0 від номінального значення залежить від однорідності основного матеріалу та відтворюваності геометричних розмірів, зокрема меандра. При серійному виробництві відхилення від номіналу становить $\pm 20\%$. Значення R_0 для магніторезисторів, виготовених із евтектичного сплаву InSb-NiSb, коливається від 10 Ом до 5 кОм.

Магніторезистивне співвідношення – це співвідношення опору магніторезистора при певних значеннях магнітної індукції до початкового опору. Магніторезистивне співвідношення сильно залежить від температури, тому необхідно наводити значення температури при якому наводиться R_B/R_0 (переважно це 25°C). Магніторезистивне співвідношення найвище для нелегованого евтектичного сплаву (близько 15 при $B = 1 \text{ Тл}$), однак у нього найгірші температурні показники. Зі збільшенням легування евтектичного сплаву InSb-NiSb магніторезистивне співвідношення зменшується, оскільки зменшується початковий опір. На рис. 13.7 наведені залежності магніторезистивного співвідношення від індукції магнітного поля для власного (1) та легованого сплаву InSb-NiSb (2,3). Для (2) розмір напівпровідникової пластини, паралельної до напрямку голок NiSb становить 100 мкм, а для (3) цей розмір становить 60 мкм, що викликає послаблення замикаючого ефекту. В табл. 13.1. наведені параметри цих трьох типів магніторезисторів.

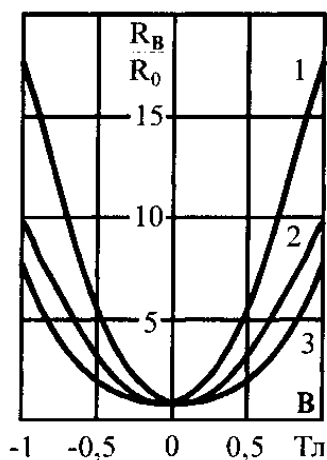


Рис. 13.7. Залежність магніторезистивного співвідношення від індукції магнітного поля.

Таблиця 13.1.

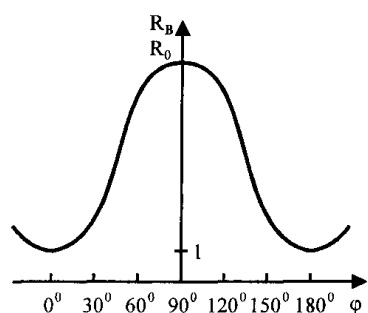


Рис. 13.8. Залежність магніторезистивного співвідношення від кута φ між напрямом струму і магнітної індукції.

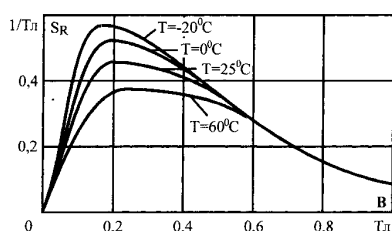


Рис. 13.9. Залежності магніточутливості від індукції магнітного поля сплаву InSb-NiSb .

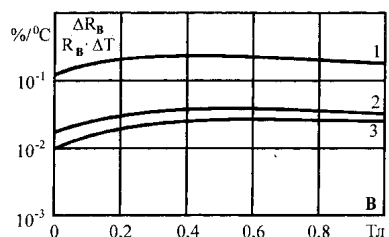


Рис. 13.10. Залежність температурного коефіцієнта опору $(\Delta R_B/R_B)/\Delta T$ від індукції магнітного поля B (позначення 1, 2, 3-відповідають рис. 13.7).

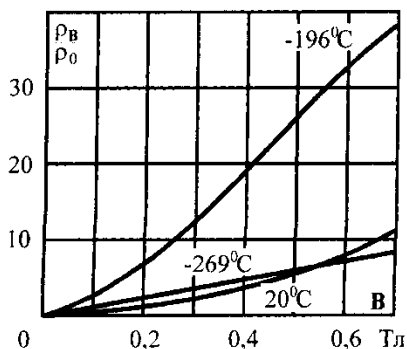


Рис. 13.11. Залежність відносного питомого опору власного сплаву InSb-NiSb від індукції магнітного поля.

Параметри магніто резисторів

| Параметр | Індукція B , Тл | Тип напівпровідникового матеріалу | | |
|--------------------------------------|-------------------|-----------------------------------|-------|---------|
| | | 1 | 2 | 3 |
| $\frac{R_B}{R_0}$ | 0,3 | 3,0-3,2 | 1,9 | 1,6-1,9 |
| $\frac{R_0}{R_B}$ | 1,0 | 13-18 | 7-10 | 6-8 |
| $\frac{1}{B} \frac{dR_B}{dT}$, %/°C | 0 | -1,8 | -0,12 | -0,16 |
| | 0,3 | -2,7 | -0,35 | -0,49 |
| | 1,0 | -2,9 | -0,50 | -0,55 |

Магніторезистивне співвідношення залежить також від кута між напрямом струму і магнітною індукцією, причому напрям голок повинен бути перпендикулярним до магнітного поля (рис. 13.8.)

Магнітна чутливість визначається як відносний приріст опору магніторезистора, поділений на відповідний приріст індукції магнітного поля:

$$S_R = \frac{1}{B} \frac{\Delta R}{R}$$

Оскільки опір магніторезистора збільшується зі збільшенням магнітного поля, а при великих полях крутизна наростання опору постійно зменшується, то максимум магнітної чутливості досягається при малих та середніх значеннях індукції магнітного поля. На рис. 13.9 наведені залежності магніточутливості від індукції при різних температурах для нелегованого евтектичного сплаву InSb-NiSb . Ці залежності дають змогу вибрати попереднє намагнічування магніторезистора з метою отримання максимального значення магніточутливості.

Навантажувальна здатність магніторезистора визначається максимальною температурою перегрівання, який допускається для магніторезистора за умови, що він залишиться робочим. Ця температура дорівнює приблизно 150°C , однак вона істотно залежить від умов тепловіддачі.

Як відомо, електронна рухливість в антимоніді індію зменшується при зростанні температури як для власних, так і для слабологованих зразків. Оскільки відносна зміна опору квадратично залежить від рухливості електронів, то збільшення температури приводить до зменшення відносного опору порівняно з його значеннями при кімнатній температурі. На рис. 13.10 наведена залежність температурного коефіцієнта опору $(\Delta R_B/R_B)/\Delta T$ від магнітної індукції при 22°C .

Іноді, особливо при дослідженнях надпровідності виникає необхідність застосовувати магніторезистори при низьких температурах. На рис. 13.11 наведена залежність відносного питомого опору власного сплаву InSb-NiSb від індукції магнітного поля при кімнатній температурі, при температурах рідкого азоту і рідкого гелію. Максимальна чутливість отримана при температурі рідкого азоту,

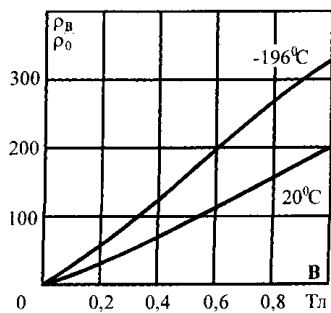


Рис. 13.12. Залежність відносного питомого опору власного сплаву InSb-NiSb від індукції магнітного поля.

оскільки подальше охолодження до температури рідкого гелію викликає падіння рухливості. Отже, магнітоопір при температурі -269°C і кімнатній температурі приблизно однаковий. На рис. 13.12 наведена залежність відносного питомого опору легованого сплаву InSb-NiSb від індукції магнітного поля при кімнатній температурі і при температурі рідкого азоту. Температурний коефіцієнт у даному діапазоні індукції магнітного поля досить малий, тому вхідна потужність на зразку повинна бути достатньо високою. Відхилення відносного питомого опору від лінійності незначне.

Лекція 14

МАГНІТОДІОДИ

План

- 14.1. Загальна характеристика магнітодіодів
- 14.2. Матеріали для магнітодіодів
- 14.3. Торцеві магнітодіоди
- 14.4. Планарні магнітодіоди
- 14.5. Планарні магнітодіоди з металевим контактом
- 14.6. Полярні магнітодіоди
- 14.7. Магнітодіоди з ефектами порогового перемикання і "пам'яті"
- 14.8. Інтегральні магнітодіоди

Література

- 14.1. Датчики. Справочник \ под ред. З.Ю.Готры и О.И.Чайковского. – Львов: Каменяр, 1995. – 312с.
- 14.2. Хомерики О.К. Полупроводниковые преобразователи магнитного поля. – М.: Энергоатомиздат, 1986.-136с
- 14.3. Вайсс Г. Физика гальваномагнитных приборов и их применение: Пер. с нем. – М.: Энергия, 1974.-384с.

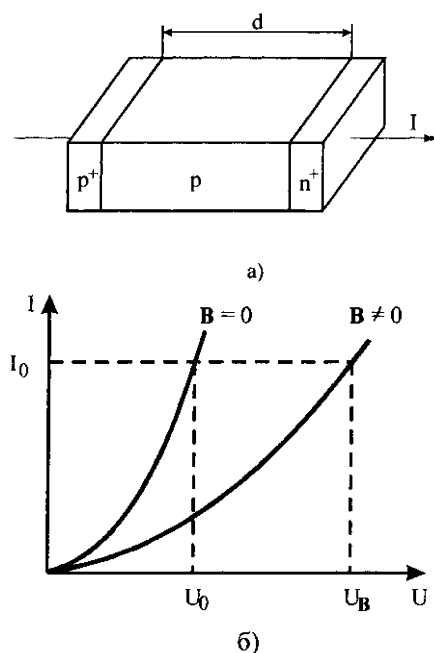


Рис. 14.1. Конструкція (а) і типова ВАХ (б) магнітодіода.

14.1. Загальна характеристика магнітодіодів

На основі магнітодіодного ефекту, був запропонований новий сенсор магнітного поля – магнітодіод. Магнітодіод є напівпровідниковим приладом з р-n-переходом і невідпрямним (омічним або антизамикаючим) контактом, між якими знаходиться область високоомного напівпровідника (рис. 14.1, а). Відмінність від звичайних діодів полягає в тому, що магнітодіод виготовляється з високоомного напівпровідника з провідністю, близькою до власної, довжина бази якого d в кілька раз більша від довжини дифузійного зміщення носіїв $L_{n,p}$, тоді як у звичайних діодах $d < L_{n,p}$. У "довгих" діодах під час проходження електричного струму визначальними стають процеси, які залежать від рекомбінації та руху нерівноважних носіїв у базі та на поверхні.

Магнітодіоди працюють при високому рівні інжекції, тобто опір їхньої базової області визначається концентрацією неосновних носіїв заряду, а магніточутливість – магнітодіодним і магнітоконцентраційним ефектами при подвійній інжекції неосновних і основних носіїв заряду, відповідно з переходів p^+-n^- і n^+-n^- (рис. 14.1, а). Сила Лоренца відхиляє носії заряду в базі у поперечному напрямку, тим самим збільшуючи опір бази і спадання напруги на ній, що приводить до зменшення спадання напруги на переходах і, відповідно, зменшення інжекції та подальшого збільшення опору бази. Наявність такого

додатного зв'язку дає змогу в ідеальному варіанті отримувати достатньо високі значення (до 104 В/А) магніточутливості. Опір бази підвищиться і за рахунок підвищення ролі поверхневої рекомбінації носіїв, відхилених до поверхні. Еквівалентну схему магніодіода можна зобразити у вигляді магніторезистора з послідовно увімкненим підсилювачем. Типова вольт-амперна характеристика магніодіода зображена на рис. 14.1, б.

Однак при високих значеннях магніточутливості нижня межа напруженості вимірюваного поля для магніодіодів не менша ніж 30 мТл. Останнє зумовлено високими значеннями температурного коефіцієнта струму (ТКС) та температурного коефіцієнта напруги (ТКН) магніодіодів. ТКН залежно від режиму роботи становить від -0,2 – +0,05 В/К, що при вольтовій магніточутливості, наприклад, $S_A = 0,02$ В/Тл, визначає температурну похибку $\Delta T = (-10 - (+2.5))$ мТл/К. Враховуючи, що вольтова магніточутливість при зменшенні значення індукції поля зменшується, реальний поріг магніточутливості магніодіодів не менший ніж 30 мТл. Для кожного значення струму, що протікає через магніодіод, існують значення магнітної індукції, при яких температурна залежність струму в магніодіодах зменшується із зростанням напруги зміщення. Поблизу цих значень магнітної індукції забезпечується слабка залежність від температури як вольт-амперної характеристики, так і магніточутливості. Тому для кожного типу магніодіодів існує рекомендована область магнітних полів та зміщень, при яких забезпечується максимальні значення експлуатаційних параметрів.

Однак для промислових магніодіодів оптимальні значення індукції магнітного поля лежать у межах 80...150 мТл. Принципово зменшення оптимального значення індукції вимірювального поля можливе тільки при збільшенні робочого струму. При цьому безпосереднє збільшення струму обмежується тепловою потужністю, виділеною на магніодіоді. Так, при $I = 5$ мА перегрівання кристала становитиме 10 – 30 К, що безумовно призводить до значної нестабільності роботи.

Існує ще один суттєвий чинник, що обмежує чутливість магніодіодів – велике значення початкового зміщення на діоді. Для забезпечення високої чутливості пристрою необхідно сформувати сигнал зміни вихідної напруги, що характеризується незалежністю від абсолютного значення напруги на діоді. Вимірювальний пристрій повинен містити високостабільні та прецизійні джерело еталонної напруги, диференціальний підсилювач тощо. Отже, високі значення абсолютної магніточутливості магніодіодів на практиці використати практично неможливо.

Основними параметрами магніодіодів є вольтова та струмова магніточутливості.

Вольтова магніточутливість S_U визначається зміною напруги на магнітодіоді при зміні магнітного поля на 1 Тл і постійному значенні струму через магнітодіод

$$S_U = \left(\frac{\Delta U}{\Delta B} \right)_{I=const} = \left(\frac{U_B - U_0}{\Delta B} \right)_{I=const} \approx \approx \frac{d}{L_{n,p}} \cdot \left(Y + \frac{L_{n,p}}{d} - 1 \right) \cdot \frac{c \cdot k \cdot T}{q} \cdot \frac{\Delta L_E}{L_E \cdot \Delta B}, \quad (14.1)$$

де L_E – усереднений шлях, який проходять носії за час життя в напрямку зовнішнього електричного поля;

$Y = \frac{q \cdot U}{c \cdot k \cdot T}$ - коефіцієнт нерівноважності за високих рівнів інжекції.

Струмова магніточутливість S_I визначається зміною струму через магнітодіод при зміні магнітного поля на 1 Тл і постійній напрузі на магнітодіоді

$$S_I = \left(\frac{\Delta I}{\Delta B} \right)_{U=const} = \left(\frac{I_B - I_0}{\Delta B} \right)_{U=const} \approx \approx j \cdot \frac{d}{L_{n,p}} \cdot \left(Y + \frac{L_{n,p}}{d} - 1 \right) \cdot \frac{c \cdot k \cdot T}{q} \cdot \frac{\Delta L_E}{L_E \cdot \Delta B}. \quad (14.2)$$

У більшості випадків при великих швидкостях поверхневої рекомбінації спостерігається залежність часу життя носіїв від магнітного поля. Якщо швидкість поверхневої рекомбінації на обох гранях однакова, то час життя носіїв із збільшенням магнітного поля зменшується. Якщо на одній грані швидкість рекомбінації більша, ніж на іншій, то при відхиленні носіїв до першої грані спостерігається зменшення, а при відхиленні до другої грані – збільшення часу життя носіїв. Згідно з цим змінюється і довжина дифузійного зміщення. Якщо зміна ефективного часу життя достатньо велика, то вона може перевищити вплив зміни рухливості та викривлення ліній струму і визначатиме магніточутливість магнітодіода.

14.2. Матеріали для магнітодіодів

Для розроблення і виготовлення магнітодіодів необхідно використовувати матеріали з низькою концентрацією власних носіїв зарядів, тобто з великою шириною забороненої зони. Крім цього, бажано використовувати напівпровідники з високою рухливістю носіїв заряду, оскільки ефект магнітоопору пропорційний до квадрата рухливості. Час життя носіїв також повинен бути високим, а концентрація дефектів структури – мінімальною.

Для виготовлення магнітодіодів в основному використовується високоомний кремній. Ширина його забороненої зони $E_g = 1,12 \text{eV}$, тому при кімнатній температурі питомий опір становить приблизно $10^2 - 10^3 \text{ Ом}\cdot\text{м}$. Це забезпечує отримання достатньо високих рівнів інжекції при високих значеннях $d/L_{n,p}$ і високій магніточутливості. Однак при великих значеннях

$d/L_{n,p}$ отримання високих значень інжекції по всій базі стає неможливим, і поблизу другого контакту з'являється немодульована область з великим опором, що різко зменшує магніточутливість. Оптимальне значення $d/L_{n,p}$ визначається з умов, що концентрація нерівноважних носіїв по всій довжині бази повинна в кілька разів перевищувати концентрацію рівноважних основних носіїв заряду. Тому, чим більший питомий опір напівпровідника, тим більші значення $d/L_{n,p}$ можуть бути використані в магнітодіодах.

Оптимальне значення співвідношення $d/L_{n,p}$ для матеріалів з різними питомими опорами виведене емпірично і становить $(d/L_{n,p})_{opt} \approx 1,2 + 0,5 \cdot \ln(P \cdot \rho_0)$, де $P = p/S$ – питома потужність, яка розсіюється, S – площа поперечного перетину магнітодіода, ρ_0 – питомий опір базової області.

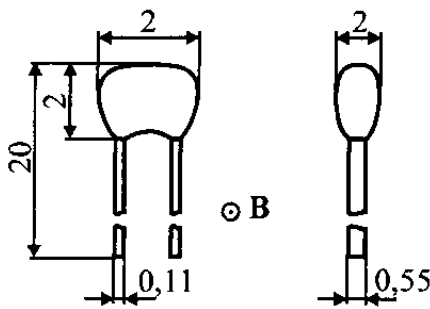
Недоліком кремнію можна вважати порівняно малі значення рухливості носіїв заряду, що зменшує магніточутливість магнітодіодів, виготовлених на його основі. Однак кремній має високі значення питомого опору, що забезпечує отримання високих рівнів інжекції при великих значеннях співвідношення $d/L_{n,p}$. Оскільки збільшення цього співвідношення суттєво підвищує магніточутливість, кремнієві магнітодіоди виявляються ефективнішими, ніж германієві.

В наш час для створення магнітодіодів використовують антимонід індію, германій, кремній, арсенід галію. Крім того, інтерес викликають арсенід індію, тверді розчини деяких напівпровідників. Для покращання інжекційних властивостей стає можливим використання гетеропереходів.

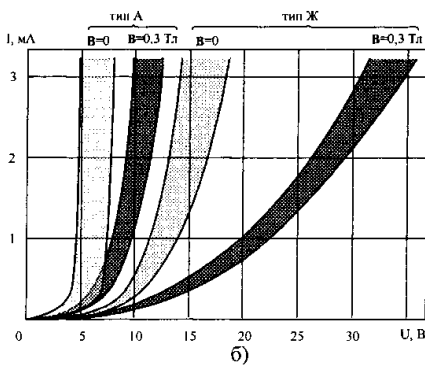
14.3. Торцеві магнітодіоди

Для створення магнітодіодів було випробувано кремній різних марок. Найкращі параметри мають магнітодіоди на основі високоомного кремнію марки КМД-20Б^a р-типу провідності з $\rho \geq 200$ Ом·м і часом життя носіїв заряду більше ніж 600 мкс. Оскільки під час високотемпературної обробки (більше ніж 1000 К) напівізолюючого кремнію КМД-20Б^a спостерігається зменшення його питомого опору і часу життя носіїв заряду, що різко зменшує магніточутливість магнітодіода, виготовленого на його основі, то для отримання магнітодіодів необхідно користуватися низькотемпературними процесами. Тому найперспективнішою для виготовлення кремнієвих магнітодіодів є технологія іонного легування.

Торцеві магнітодіоди виготовлялися методами іонного легування і сплавної технології. Використовувалися пластини кремнію КМД-20Б^a з кристалографічною орієнтацією по площині [111] і завтовшки 1 мм. Для створення n-області використовували іони фосфору, а для створення p⁺-області – іони бору. Області n⁻ і p⁺-типу



а)



б)

Рис. 14.2. Конструкція (а) і типова ВАХ (б) торцевого магніодіода.

формувався на протилежних гранях пластини. Для активації впроваджених іонів проводився відпал при температурі 850-900 К в середовищі аргону протягом 30-60 хв. Як контакти до n^+ - і p^+ -областей використовувалися V, Al, Ni.

Для виготовлення торцевих магніодіодів за сплавною технологією використовувалися такі самі напівпровідникові пластини. Переходи $p-n$ і p^+-p створювалися на протилежних гранях пластини. Перехід $p-n$ формувався вакуумним напиленням Au+Sb (1 %) з наступним сплавленням при температурі 735 К. Потім методами вакуумного напилення осаджувався Ni. Перехід p^+-p створювався вплавленням алюмінієвої шайби при температурі 945 К. Омичні контакти до переходів здійснювалися хімічним осадженням Zn з наступним гальванічним осадженням Ni і Sn. Далі пластини розрізаються на елементи розміром $1 \times 0,5 \times 0,5$ мм. Електричні виводи реалізовувалися термопаянням срібних виводів до контактних площинок. Елементи захищалися епоксидною емаллю. Конструкція і ВАХ магніодіодів наведені на рис. 14.2. У торцевих магніодіодах конструкція виключає вплив поверхні напівпровідника на їх характеристики, тому вони мають однакову чутливість до магнітної індукції незалежно від її напрямку.

14.4. Планарні магніодіоди

Такі магніодіоди виготовляються за планарною технологією. Це дає змогу поряд з іншими перевагами планарної технології отримувати магніодіоди із захищеними SiO_2 поверхнями кремнію. Планарна технологія дає можливість різними методами обробки протилежних граней кремнієвої пластини створювати на них ділянки з різними швидкостями поверхневої рекомбінації. Наявність на протилежних гранях кремнієвої пластини областей з різними швидкостями рекомбінації зумовлює залежність магніточутливості від напрямку магнітної індукції.

На асиметрію магніточутливості значною мірою впливають час життя неосновних носіїв заряду τ , довжина d і товщина h бази магніодіода, швидкості поверхневої рекомбінації s_1 і s_2 на протилежних гранях бази.

Якщо використовувати кремній з відомим часом життя носіїв заряду і технологію, яка забезпечує отримання певних значень s_1 і s_2 , розміри d і h стають основними параметрами, які визначають характеристики магніодіода.

Для виготовлення магніодіодів, чутливих до напрямку магнітного поля, основним технологічним процесом є отримання зон зі зниженою швидкістю поверхневої рекомбінації. Зони з підвищеною швидкістю рекомбінації створюються за допомогою порушення структури напівпровідника в локальних ділянках і-шару

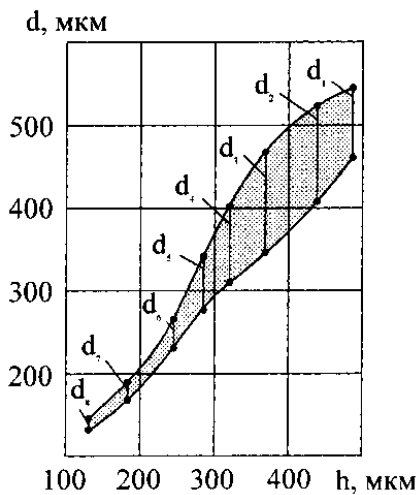


Рис. 14.3. Залежності відстані між контактами d від товщини пластини h (заштрихована ділянка відповідає значенням з напругою 4- 15 В і струмом 3 мА).

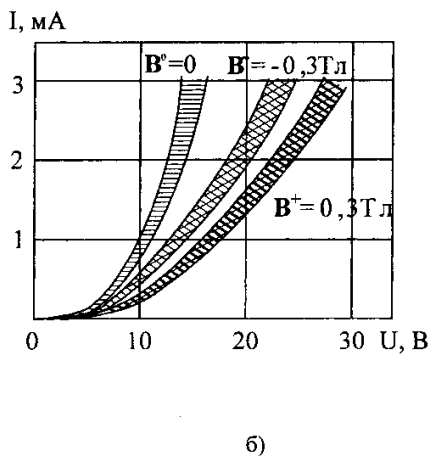
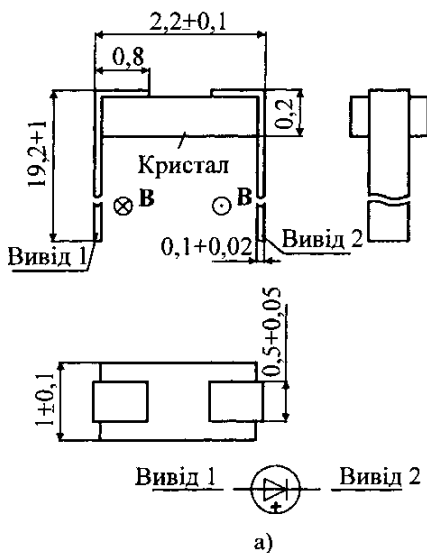


Рис. 14.4. Конструкція (а) і типова ВАХ (б) планарного магнітодіода.

або на самій поверхні, або біля її межі. Такі порушення створюються піскоструминною обробкою, іонним бомбардуванням, введенням у приповерхневий шар домішок з глибокими рівнями. Для забезпечення малої швидкості поверхневої рекомбінації (менше ніж 0,1-1,0 м/с) необхідно створювати межі розділу напівпровідника з діелектриком з малою концентрацією рекомбінаційних центрів і вигином зон, які перешкоджають рухові неосновних носіїв до поверхні. Вигин зон визначається типом і густиною поверхневих станів і зарядів у діелектрику.

Поверхня з малою швидкістю поверхневої рекомбінації створюється низькотемпературною обробкою межі розділу кремнію з оксидом кремнію у відновлювальному середовищі, нанесенням на поверхню кремнію фосфоросилікатного скла, низькотемпературним окисленням, формуванням на межі розділу негативного заряду.

Для виготовлення магнітодіодів використовується кремній КМД-20Б^а. Пластини мають кристалографічну орієнтацію по площині [111]. Для пасивації поверхні кремнію і захисту елементів конструкції магнітодіода наноситься методом низькотемпературного розкладання (до 775 К) плівка піролітичного тетраетоксисилану $\text{Si}(\text{OC}_2\text{H}_5)_4$ (два шари загальною товщиною 0,3 – 0,4 мкм). У технології використовуються виключно низькотемпературні процеси, тому що при високих температурах властивості високоомного кремнію можуть дуже змінюватися через появу термоакцепторів.

Зона з підвищеною швидкістю поверхневої рекомбінації формується грубим шліфуванням грані, протилежної до грані з контактами. Таким методом забезпечується швидкість рекомбінації, більша ніж 20 м/с. На планарній грані швидкість рекомбінації суттєво менша.

Інжектуючий і антизамикаючий контакти створюються за планарно-сплавною технологією. Вплавлення при температурі 875 К напиленого шару алюмінію забезпечує отримання задовільного р⁺-р-переходу, а прогрівання при температурі 740 К сплаву золота з 1% сурми формує р-n-перехід. На сьогодні однозначного визначення довжини базової області при планарній технології немає. У таких магнітодіодах мінімальна відстань між краями контактних ділянок становить 0,8 мм, а між їхніми центрами – 1,3 мм.

Кожному значенню довжини бази відповідає оптимальне значення товщини магнітодіода, яке забезпечує максимальну магніточутливість. Підбір значень d і h для отримання магнітодіодів з прямим спаданням напруги 4 – 15 В і, відповідно, струмом 3 мА можна провести за допомогою рис. 14.3.

Конструкції і типові характеристики магнітодіодів

наведені на рис. 14.4. У магнітодіодах описаних типів p^+ - p - і p - n -переходи можуть також виготовлятися методом іонного легування. Вольт-амперні характеристики магнітодіодів, виготовлених іонним легуванням, практично не відрізняються від ВАХ магнітодіодів, виготовлених методами планарно-сплавної технології.

14.5. Планарні магнітодіоди з металевим контактом

У розглянутих вище магнітодіодах нерівноважна провідність створювалася інжекцією нерівноважних носіїв прямо зміщеним p - n -переходом. Інжектувати нерівноважні носії може і контакт метал-напівпровідник при відповідному співвідношенні різниці робіт виходу. Сам метал може інжектувати як дірки, так і електрони, оскільки вище від рівня Фермі рівні заповнені дірками, а нижче – електронами. Однак, ефективніша інжекція забезпечується протягом виникнення у напівпровіднику інверсного шару під дією контактної різниці потенціалів. В цей час основний спад напруг існуватиме на цьому індукваному контактною різницею потенціалів p - n -переході, який і забезпечить інжекцію носіїв при прямому зміщенні. У напівпровіднику з вихідною провідністю, близькою до власної, легко отримати в інверсному шарі набагато більшу концентрацію основних носіїв, порівняно з концентрацією в базовій області, тобто створити несиметричний p - n -перехід і забезпечити умови односторонньої інжекції з інверсного шару в базову область. Отже, на основі таких структур можна створити високоефективні магнітодіоди.

Одним з металів, найбільш подібних на кремній як за різницею робіт виходу, так і за технологією виготовлення, є алюміній. Робота виходу електронів з алюмінію становить 3,74 eV, тоді як для кремнію p -типу вона становить 4,6 eV. З цієї причини на поверхні кремнію p -типу, яка межує з алюмінієм, повинен створюватися інверсний шар з високою концентрацією електронів. Це дасть можливість отримати малі зворотні струми і високу ефективність інжекції електронів при прикладанні від'ємної напруги зміщення до алюмінію. В симетричних структурах, у яких обидва контакти створені напиленням алюмінію на кремній p -типу, прикладена зовнішня напруга практично повністю падає на контакті, зміщеному в зворотному напрямку, як і в будь-якій транзисторній структурі. Тому створити на ньому помітну нерівноважну провідність практично неможливо. Однак добрі випрямні властивості контакт алюмінію з p -кремнієм має тільки за малої концентрації акцепторів у вихідному матеріалі. Якщо рівень легування високий, то інверсний шар виявляється дуже тонким і вже при невеликих зворотних зміщеннях протікання через нього струму визначається тунельним ефектом, його опір зменшується, тому основне зміщення відбувається на

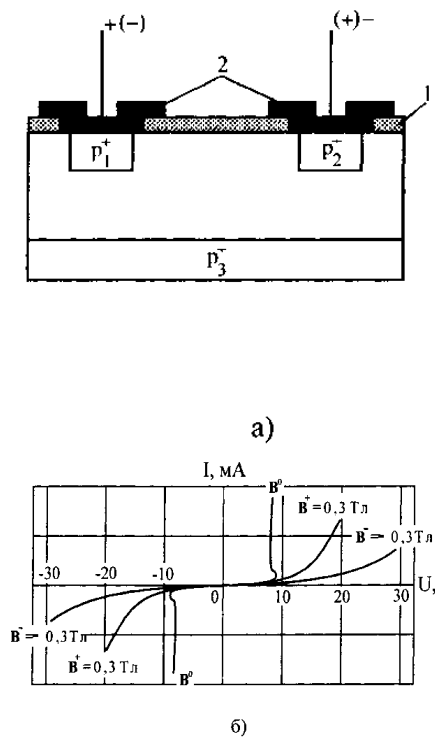


Рис. 14.5. Конструкція (а) і ВАХ (б) симетричних магніодіодів при $B = \pm 0,3 \text{ Тл}$: 1 – шар SiO_2 , 2 – електрод Al .

прямозмщеному р-п-переході. При високому рівні початкового легування неможливо створити високі рівні інжекції в базовій області і, відповідно, забезпечити високу магніточутливість.

Отже, для симетричного магніодіода з металевим контактом була вибрана конструкція, зображена на рис. 14.5. На кремнієвій р-пластині типу КМД-20Б^а легуванням алюмінієм створюються сильнолеговані області p_1^+ і p_2^+ площею $0,32 \text{ мм}^2$. Потім на ці області у вакуумі напилюється алюміній так, щоб повністю закрити p_1^+ і p_2^+ -області і невеликі ділянки вихідного напівпровідника, які межують з цими областями. Площа напиленого алюмінію становить $0,56 \text{ мм}^2$, відстань між краями напиленого алюмінію $470 - 500 \text{ мкм}$. Вибір параметрів d і h здійснюється аналогічно, як і для попередніх типів магніодіодів. У такій структурі в прямому напрямку основний струм протікає через контакт металу з вихідним р-напівпровідником, а в зворотному – через контакт металу з р⁺-областю. Цим досягається можливість пропускання великих інжекційних струмів і, відповідно, високих рівнів інжекції і значної магніточутливості. Для зменшення швидкості поверхневої рекомбінації на протилежній грані кремнієвої пластини створюється збагачений p_3^+ -шар, який забезпечує значну різницю в магніточутливості залежно від напрямку магнітного поля.

Вказана технологія порівняно з технологією виготовлення торцевих і звичайних планарних магніодіодів, значно спрощує технологічний процес виготовлення, забезпечуючи високий відсоток виходу придатних приладів. Оскільки структура даних магніодіодів симетрична, при протилежних напрямках струму вольт-амперних характеристик (ВАХ) і магніточутливість приблизно однакові (рис. 14.5, б).

Випускається також модифікація планарних симетричних магніодіодів, конструкція і ВАХ яких наведені на рис. 14.6.

14.6. Полярні магніодіоди

У магніодіодах планарної топології магніточутливість залежить від напрямку магнітного поля. У магнітному полі, спрямованому так, що носії відхиляються до грані, на якій розміщені контакти, відбувається збільшення рекомбінації носіїв на цій грані. Одночасно їх шлях від р-п-переходу до другого контакту скорочується. При протилежній полярності магнітного поля рекомбінація на планарній грані суттєво зменшується. Залежно від топології магніодіода, швидкості поверхневої рекомбінації, товщини напівпровідникової пластини, питомого опору і часу життя носіїв в базовій області магніточутливість буде більшою при тій чи іншій полярності магнітного поля.

Якщо створити на планарній грані високу швидкість

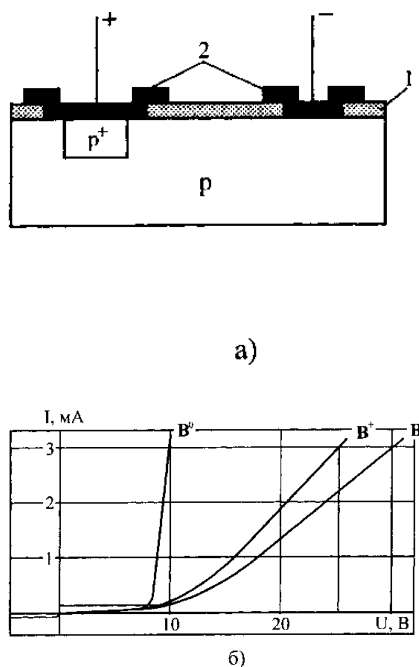


Рис. 14.6. Конструкція (а) і ВАХ (б) модифікованих планарних магніодіодів.

поверхневої рекомбінації, то при певному підборі геометричних і електрофізичних параметрів можна виготовити магніодіоди в яких при відхиленні носіїв у протилежний від планарної грані бік вплив зростання ефективного часу життя носіїв внаслідок зменшення ролі поверхневої рекомбінації буде більшим, ніж вплив видовження ліній струму. Магнітне поле такої полярності буде не зменшувати струм магніодіода, а збільшувати його. Магнітне поле протилежної полярності, в свою чергу, буде зменшувати струм магніодіода. Магніодіоди, в яких знак зміни струму залежить від напрямку магнітного поля, називаються полярними. Одним із найзручніших методів забезпечення полярної магніточутливості є використання для магніодіодів тонких ($h < L_{n,p}$) напівпровідникових пластин і забезпечення на грані, протилежній планарній, високої швидкості поверхневої рекомбінації. Зі збільшенням магнітного поля ефект полярної магніточутливості поступово зникає, тобто в сильних полях струм магніодіода зменшується при будь-якій полярності, що пов'язано із закручуванням ліній струму під дією сильних магнітних полів. Однак збільшення співвідношення $d/L_{n,p}$ розширює діапазон значень індукції магнітного поля, при яких спостерігається полярна магніточутливість.

Полярні магніодіоди виготовляються на основі кремнію р-типу ($\rho \approx 5$ Ом·м, $\tau < 30$ мкс). Для отримання області з підвищеною швидкістю поверхневої рекомбінації поверхня пластини піддається піскоструминній обробці. Протилежна поверхня полірується і травиться розчином CP4 або термічно окислюється для отримання області з пониженою швидкістю рекомбінації.

Полярні магніодіоди виготовляються також і за технологією іонного легування з кремнію р-типу з такими параметрами: питомий опір ≥ 200 Ом·м, довжина дифузійного зміщення $(5-9) \cdot 10^{-4}$ м. Вони мають структуру р⁺-р-n типу. Область із зниженою швидкістю рекомбінації на планарній грані створюється після хімікодинамічного полірування обробкою в біхроматі калію до створення окислу або нанесенням піролітичним методом плівки оксиду алюмінію. Висока швидкість на протилежній грані створюється нанесенням шару металу.

Магніодіоди із захисною плівкою з оксиду алюмінію на планарній грані мають вищу магніточутливість при обох полярностях магнітного поля, ніж магніодіоди із захисною плівкою з оксиду кремнію.

Типова вольт-амперна характеристика кремнієвого магніодіода з полярною магніточутливістю зображена на рис. 14.7, а. Залежність струму від напруги при малих зміщеннях близька до лінійної. У робочій області струмів ВАХ можна апроксимувати степеневим законом $I \sim U^n$, показник степеня $n > 2$. Показник степеня n залежить від

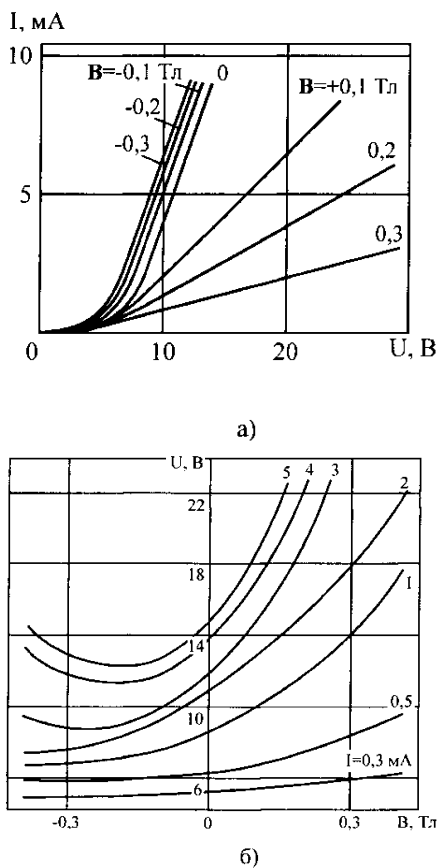


Рис. 14.7. Вольт-амперна (а) і вольт-теслова (б) характеристики полярного магніодіода.

топології магнітодіода, його товщини і відстані між контактами, від напрямку і значення індукції магнітного поля.

Полярна магніточутливість спостерігається в широкому діапазоні струмів і значень індукції магнітного поля. Зі збільшенням індукції магнітного поля діапазон струмів з полярною магніточутливістю зменшується, аналогічно зростання струму магнітодіода зменшує діапазон значень індукції магнітного поля з таким ефектом (рис. 14.7, б). При малих значеннях індукції магнітного поля $B < 0,1$ Тл від'ємна і додатна магніточутливості приблизно однакові. Мінімум вольт-теслової характеристики полярних магнітодіодів зміщений у діапазон від'ємних значень індукції магнітного поля.

14.7. Магнітодіоди з ефектами порогового перемикавання і "пам'яті"

Шари халькогенідних напівпровідникових стекел між металевими електродами мають симетричні вольт-амперні характеристики з ділянками від'ємного диференційного опору, які розділяють два можливі стійкі стани – високоомний і низькоомний. При певних складах таких халькогенідних стекел ці структури після переходу в низькоомний стан залишаються в ньому і після зняття напруги, тобто в них спостерігається ефект "пам'яті". Перепад опорів у високоомному і низькоомному станах може перевищувати 3 – 4 порядки при відношенні порогової напруги до залишкової до 50 – 100 разів. Ефекти перемикавання і "пам'яті" визначаються об'ємними процесами і не пов'язані з існуванням р-п- і гетеропереходів.

Використання халькогенідних стекел дає змогу створювати магнітодіоди з ефектами порогового перемикавання і "пам'яті". Базовою основою для таких магнітодіодів є планарні магнітодіоди. На р-п-перехід напилюється шар халькогенідного скла необхідного складу. Потім на цей шар наносять шари молібдену і алюмінію, а на р⁺-шар другого контакту – алюміній (рис. 14.8).

Як халькогенідні напівпровідникові стекла в магнітодіодах використовують $Si_{12}-Ge_{10}-As_{30}-Te_{48}$ або $Ga_{12}-Ge_{6,5}-Si_{6,5}-As_{25}-Te_{50}$. Порогова напруженість електричного поля, яке забезпечує перемикавання у провідний стан, $10^7 - 10^8$ В/м, час перемикавання – 10^{-9} с при часі затримки – 10^{-6} с. Час затримки зменшується зі збільшенням напруги зміщення. Напруга вмикання, залишкова напруга і опір у низькоомному стані сильно залежать від індукції магнітного поля. Вольт-амперні характеристики таких магнітодіодів симетричні (рис. 14.9).

На рис. 14.10 зображені вольт-амперні характеристики магнітодіодів, створених на основі планарних

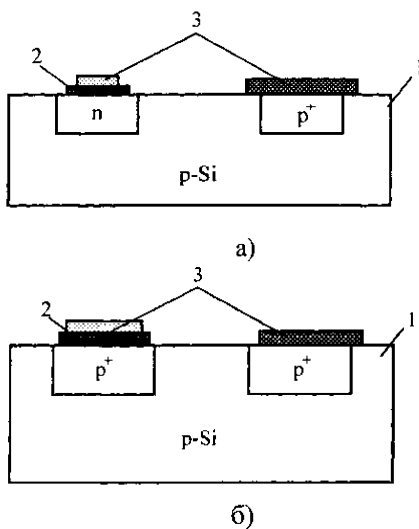


Рис. 14.8. Структура магнітодіода з аморфно-кристалічним гетеропереходом: а) на базі звичайного планарного магнітодіода; б) на базі планарного магнітодіода з металевим контактом: 1 – база кремнію р-типу, 2 – шар халькогенідного скла, 3 – металеві контакти.

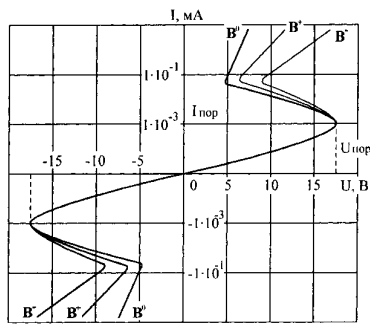
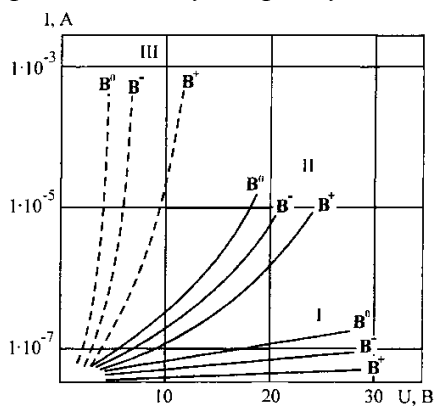
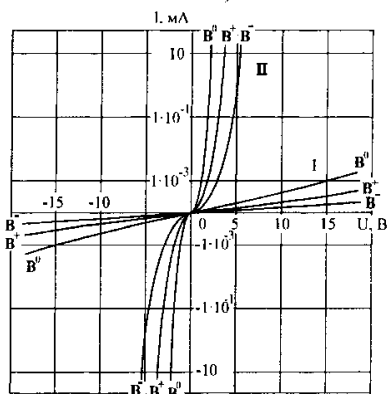


Рис. 14.9. Симетричні ВАХ магніодіодів з пороговим перемиканням при відсутності магнітного поля B^0 , в полі, коли носії відхиляються до планарної грані B^- і в протилежному напрямку B^+ .



а)



б)

Рис. 14.10. Типові ВАХ магніодіодів з пам'яттю за відсутності магнітного поля і в полі з різними напрямками: а) на базі планарного магніодіода; б) на базі модифікованого симетричного магніодіода: I – високоомний стан халькогенідного скла, II – низькоомний стан халькогенідного скла, III- без шару халькогенідного скла.

магнітодіодів, за відсутності магнітного поля і в магнітному полі 0,3 Тл. Порогова напруга 28 – 36 В. Вольт-амперна характеристика до перемикання лінійна. Час перемикання магнітодіода менший ніж 1 мкс. Стан з низьким опором після перемикання залишається і після зняття напруги. Напруга перемикання і параметри після перемикання залежать не тільки від індукції магнітного поля, але й від напрямку.

Оскільки магнітодіод з шаром халькогенідного скла зберігає інформацію про провідний стан і при нульовому зміщенні, він може бути використаний як елемент пам'яті в магнітокерованих схемах. Повний опір магнітодіода з халькогенідним склом у провідному стані на 1-2 порядки перевищує опір магнітодіода без цього шару.

На базі планарного магнітодіода напиленням халькогенідного скла системи Te-Se на всю поверхню створено фоточутливий магнітодіод з пам'яттю.

Ефект "пам'яті" в шарах халькогенідних стекел пояснюється проростанням кристалізуючого "шнура" від негативно зарядженого контакту до позитивного. Провідність цих з'єднань у кристалічному стані набагато вища, ніж в аморфному. Коли "шнур" досягає протилежного металевго контакту, встановлюється низькоомний стан. "Шнур" зберігається і без прикладання напруги зміщення, що й забезпечує ефект "пам'яті". Отже, наявність двох метастабільних станів магнітодіодів з шаром халькогенідних стекел пов'язана з фазовим переходом в об'ємі шару. Стирання пам'яті, тобто перехід з низькоомного у високоомний стан, забезпечується подаванням короткого імпульсу струму тривалістю близько 10^{-5} с. Такий струм проходить по провідному каналу, розігріває і розплавляє його.

Протягом швидкого подальшого охолодження він не встигає закристалізуватися і речовина залишається в аморфному стані з низькою провідністю.

14.8. Інтегральні магнітодіоди

Як зазначалося вище, магнітодіоди (МД) виготовлялися, в основному, на високоомному кремнії або германії, мали великі розміри і відрізнялися неможливістю інтеграції в інтегральні схеми з технологічної і конструктивної точки зору. Сучасні магнітодіоди працюють на магнітоконцентраційному ефекті і створюються за стандартними КМОН- (комплементарна метал-оксид-напівпровідник) та КНС- ("кремній на сапфірі") технологіями і, відповідно, можуть входити до складу сенсорних пристроїв, виготовлених за даними технологічними циклами.

Швидкості поверхневої рекомбінації s_1 і s_2 повинні сильно відрізнятися одна від одної і задовольняти співвідношення:

$$s_1 \gg \frac{L_p}{t_p} \gg s_2, \quad (14.3)$$

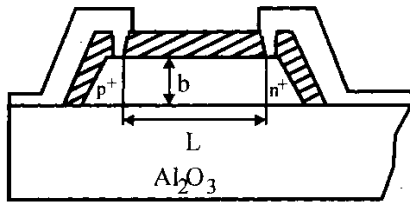


Рис. 14.11. Конструкція магнітодіода на основі КНС-технології.

де L_p і τ_p – дифузійна довжина і час життя дірок в об'ємі бази, відповідно. При цьому питома магніточутливість магнітодіода дорівнюватиме:

$$S_s = \frac{m_n + m_p}{2 \cdot j_T} \cdot E_x \cdot \frac{L_p}{b} \cdot \frac{1 - ch^{-1}(b/L_p)}{1 - b/L_p}, \quad (14.4)$$

де μ_n і μ_p – рухливості електронів і дірок, відповідно, φ_T – тепловий потенціал, E_x – напруженість електричного поля в базі, b – товщина бази магнітодіода (рис. 14.11). Максимальна чутливість досягатиме 100 В/Тл при товщинах бази $b = 1,5 \cdot L_p$. У випадку відхилення електронно-діркової плазми за рахунок дії сили Лоренца до однієї з граней кількість носіїв змінюється через зміну ефективної швидкості рекомбінації, впливаючи при цьому на опір бази.

У магнітодіодах на КНС межі розділу Si-SiO₂ і Si-Al₂O₃ (рис. 14.11) мають значення меж з малою і високою швидкістю рекомбінації, відповідно. Ці величини відрізняються більш як на два порядки. Такі МД працюють в експоненціальному, омичному, напівпровідниковому, діелектричному та дифузійному режимах. Найбільша чутливість у слабких магнітних полях спостерігається в напівпровідниковому режимі при високих рівнях інжекції. ВАХ такого режиму має вигляд

$$I = \frac{9}{8} \cdot q \cdot n_0 \cdot m_n \cdot t_{eff} \cdot \frac{U^2}{d^3} \quad (14.5)$$

де n_0 – вихідна концентрація основних носіїв заряду в базі, t_{eff} – ефективний час життя носіїв в базі:

$$t_{eff} : (s_2 \cdot B \cdot U)^{-1}, \quad (14.6)$$

де s_2 – швидкість рекомбінації на межі Si-Al₂O₃. Експериментальні значення t_{eff} не перевищують 1 не за відсутності магнітного поля, тому магнітодіоди на КНС в принципі можуть працювати в широкому діапазоні частот – від 0 до 10⁸ Гц.

У реальних магнітодіодах, виконаних на плівці Si (100) завтовшки 0,7 мкм, досягнута максимальна чутливість 14 В/Тл та відносна 13 %/Тл при струмах живлення 1 мА, довжині і ширині бази 50 мкм, вихідній концентрації домішок $5 \cdot 10^{21} \text{ м}^{-3}$ і дифузійній довжині дірок у базі 1 мкм. Висока чутливість спостерігається при концентрації електронів в базі $2 \cdot 10^{22} \text{ м}^{-3}$. Теоретичне значення порога чутливості магнітодіода на КНС при відношенні сигнал/шум, який дорівнює 1 досягає 10⁻⁹ Тл, коли частоти 0,1-10⁵ Гц, а еквівалентний температурний дрейф – 10⁻³ Тл/К. Експериментальне дослідження шумів у магнітодіоді в діапазоні частот від 0 до 10⁵ Гц показало, що співвідношення сигнал/шум та температурний дрейф суттєво залежать від довжини бази, значення початкової концентрації носіїв, температури, індукції магнітного поля, що дуже обмежує використання такого магнітодіода. Поріг чутливості досягав 10⁻⁷ Тл при

довжині і ширині бази, відповідно, 10 і 50 мкм і абсолютній чутливості 10 В/Тл. При цьому шуми зростали зі збільшенням довжини бази і напруги на ній та зі зменшенням концентрації домішки в базі і її ширини. Зменшення температури та індукції магнітного поля збільшувало спектральну густину шумів і магніточутливість.

Різновидом магнітодіода на КНС є магнітодіод з бар'єром Шотткі, робота якого базується на модуляції зворотного струму переходу Шотткі, сформованого в області бази магнітодіода при відхиленні до нього електронно-діркової плазми прикладеним магнітним полем. Такий МД має чутливість 10 В/Тл при чутливості магнітодіода на КНС 1 В/Тл.

Основним недоліком магнітодіодів на КНС є їхня висока собівартість внаслідок використання дорогих вихідних матеріалів. Дешевшою і більш поширеною є КМОН-технологія, тому КМОН магнітодіоди (рис. 14.12) перспективніші. Робочою областю таких магнітодіодів є n^- -кишеня, сформована методом іонної імплантації в p -підкладку, а межею з великою швидкістю рекомбінації – межа p - n переходу робоча область-підкладка, який увімкнено у зворотному напрямку. На цій межі відбувається збирання носіїв, швидкість рекомбінації поблизу неї безмежно велика. Магнітодіод з площею емітера 18×28 мкм, довжиною бази 14-136 мкм, шириною 50 мкм і початковою концентрацією носіїв 10^{21} м^{-3} має чутливість до 25 В/Тл. Він працює в режимі сильної інжекції, інформаційним параметром служать вхідні ВАХ переходу емітер-база.

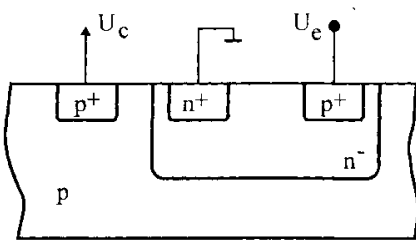


Рис. 14.12. Конструкція магнітодіода на основі КМОН-технології.

Лекція 15

ПЛІВКОВІ НАПІВПРОВІДНИКОВІ СЕНСОРИ ВОЛОГОСТІ.

План

- 15.1. Принцип роботи сенсорів вологості
- 15.2. Тонкоплівкові алюмінієво-оксидні сенсорні вологості
- 15.3. Плівкові сенсорні для вимірювання точки роси

Література

- 2.1. Фізичні основи електронної техніки: підручник / З.Ю. Готра, І.Є. Лопатинський, Б.А. Лук'янець, З.М. Микитюк, І.В. За ред. Готри З.Ю. Львів: Бескид Біт, 2004. – 880 с.
- 2.2. Мікроелектронні сенсорні фізичних величин: Науково-навчальне видання. В 3 томах. / Вуйцік В., Готра З.Ю., Григор'єв В.В., Каліта В., Мельник О.М., Потенцік Є.; за ред. З.Ю. Готри. – Львів: Ліга-Прес, 2002. – 475 с.

15.1. Плівкові напівпровідникові сенсорні вологості

Дія сенсорні такого типу ґрунтується на зміні провідності тонкого шару напівпровідника під впливом заряду адсорбованих на його поверхні або поверхні розміщеного на ньому діелектричного шару полярних молекул, в тому числі молекул води. Характеристика таких сенсорні задовільно описується ізотермою адсорбції Фрейндліха:

$$\Delta d = K \cdot C^n \quad (15.1)$$

де Δd – зміна провідності, K – коефіцієнт, C – концентрація парів води в газі, n – стала ($n < 0,5$). Чутливість напівпровідникових сенсорні відносно парів води різко зменшується при збільшенні його робочої температури.

Одним із перших сенсорні такого типу є діод Шоткі, утворений переходом кремній (германій) – двоокис олова. Контакти наносять на протилежні поверхні структури. Недоліками сенсора є достатньо велика стала часу і низька чутливість при малих рівнях вологості. Для збільшення чутливості на поверхню тонкого (50-80 нм) шару окису цинку, розміщеного на діелектричній підкладці з двома електродами, наносять шар діелектрика товщиною 10-40 нм. Діелектриком служить тефлон (при визначенні вологості понад 80%), що має гідрофобні властивості або полімер, отриманий поліконденсацією в плазмі тетраетоксисилану (при визначенні вологості в діапазоні від 0 до 100%). Вимірювання проводять на постійному струмі, що є перевагою сенсорні такого типу. Опір сенсора в діапазоні від 0 до 100% змінюється приблизно від 100 МОм до кількох сотень кілоом.

Зменшення діапазону вимірювання опору сенсора і, відповідно, підвищення точності вимірювання досягнуте за рахунок введення додаткового керуючого електрода, який наносять на поверхню діелектричного шару. Подаючи на цей електрод додатний або від'ємний

потенціал, можна зміщувати характеристики сенсора в заданий діапазон і забезпечувати в ньому оптимальні чутливість і точність вимірювання. Площа, яку займає керуючий електрод, не повинна перевищувати 45-50% від площі діелектричного шару. Керуючий електрод виготовляють з алюмінію, що має найкращу адгезію до діелектрика, отриманого поліконденсацією тетраетоксисилану. В якості напівпровідника використовували двоокис олова. Можливе також застосування окису цинку, окислів марганцю, міді, нікелю, титану, кадмію, заліза, срібла, вісмуту, свинцю, кальцію, магнію, ванадію, молібдену або вольфраму.

Як сенсор використовують кремнієвий МОН транзистор з відкритим затвором. Для збільшення чутливості на відкриту частину підзатворного окислу товщиною 0,1 мкм наносять полімерну плівку товщиною 0,3 мкм. Канальна провідність такого транзистора залежить від кількості адсорбованих молекул води і прикладеної до затвору керуючої напруги. Вологість середовища визначають за часом затримки відкриття каналу стік-витік, який залежить від опору полімерної плівки.

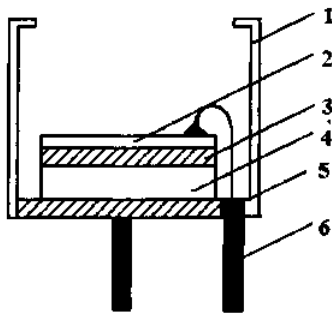


Рис. 15.1. Структура сенсора вологості на основі плівки Al_2O_3 : 1 – корпус; 2 – тонка плівка; 3 – пориста плівка Al_2O_3 ; 4 – шар Al; 5 – металева підкладка; 6 – виводи.

15.2. Тонкоплівкові алюмінієво-оксидні сенсори вологості

Тонкоплівкові сенсори вологості на основі Al_2O_3 – один із перших видів твердотільних сенсорів. Вони набули широкого розповсюдження у зв'язку з простотою їх виготовлення. За принципом дії вони дуже подібні до керамічних сенсорів. Головним процесом, який визначає чутливість до парів води, є фізична адсорбція.

Такий сенсор (рис. 15.1) представляє собою плоский конденсатор, однією із обкладинок якого є плівка алюмінію, діелектриком – пориста плівка Al_2O_3 , отримана анодуванням алюмінію; верхньою обкладинкою є напилена тонка плівка золота, проникна для парів води (рис. 15.2, а). Модель утворення пор у плівках Al_2O_3 показана на рис. 15.2, б.

Товщина бар'єрного шару визначається напругою анодування: типова його величина становить 1 нм/В. Діаметр пор та загальна пористість структури залежить від природи електроліту та режиму вирощування плівки. Процеси вирощування таких анодних плівок Al_2O_3 досить добре вивчені. Механізми заповнення пор можна пояснити за допомогою теорії капілярної конденсації. Згідно з даною теорією, пори заповнюються в два етапи. Спочатку при невеликій вологості стінки пор заповнюються мономолекулярним шаром води. При підвищенні вологості після насичення стінок вода починає конденсуватися, причому спочатку заповнюються пори з меншим діаметром. В зв'язку з цим існує три стани заповнення пор (рис. 15.2 б).

Цей процес моделюється електричною еквівалентною

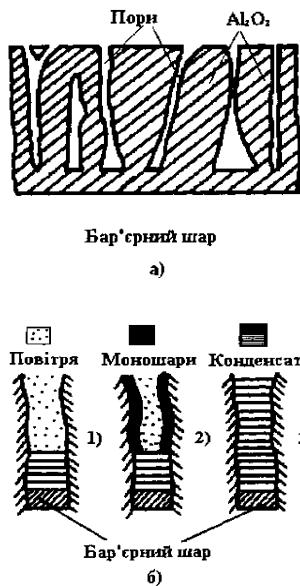


Рис. 15.2. Розріз структури сенсора (а) та етапи заповнення пор водою (б): 1 – напівпуста пора; 2 – утворення моношарів на стінках пор; 3 – заповнена пора.

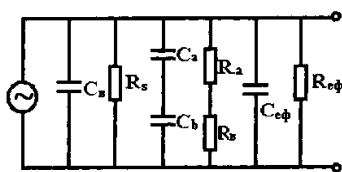


Рис. 15.3. Електрична еквівалентна схема моделі Al_2O_3 – сенсора вологості.

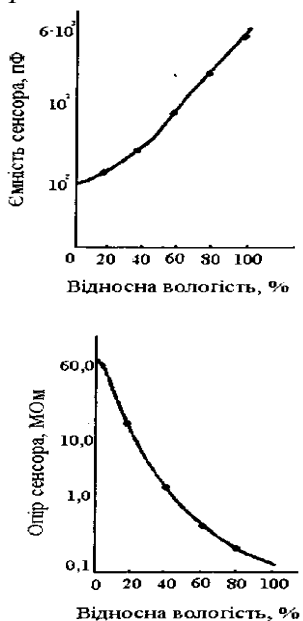


Рис. 15.4. Графіки залежності ємності (а) і паралельного опору (б) Al_2O_3 – сенсора від відносної вологості.

схемою (рис. 15.3).

На схемі елементи C_s і R_s представляють собою ємність та електричний опір стінок пор, відповідно; C_a і R_a – ємність і опір бар'єрного шару; C_{ef} – ємність багатошарової структури, в яку входять бар'єрний та хемосорбційний шари, а також шар води, яка сконденсована в порах; R_{ef} – паралельний до цієї ємності опір багатошарової структури. Використовуючи цю еквівалентну схему, можна отримати:

$$C = C_s + C_a C_b (C_a + C_b)^{-1} + C_{ef} \quad (15.2)$$

$$R^{-1} = R_s^{-1} + (R_a + R_b)^{-1} + R_{ef}^{-1} \quad (15.3)$$

При відсутності вологості $C_{ef} = 0$, тоді

$$C = \epsilon_0 \epsilon_s A \left[\frac{l-a}{d} + \frac{e_a a}{(e_a d + e_s l)} \right] \quad (15.4)$$

де ϵ_0 – діелектрична проникність вакууму; ϵ_s – діелектрична стала Al_2O_3 ($\epsilon_s = 8$); ϵ_a – відносна проникність повітря ($\epsilon_a = 1$); A – площа приладу; a – пористість плівки; d – товщина бар'єрного шару; l – довжина пор на якій може відбуватися конденсація води (важливий параметр, який визначає метрологічні характеристики сенсора).

Ємність при неконтрольованій вологості може бути оцінена так:

$$C = \frac{\epsilon_0 \epsilon_s (l-a) A}{d} + \epsilon_0 a A \left[\frac{x e_1}{d} + \frac{(l-x) e_a e_b}{e_a B + e_s l} \right], \quad (15.5)$$

де x – частина пор заповнених водою ($x = \alpha A$). Аналогічно для електричного опору:

$$R = d (w \epsilon_1 \epsilon_2 x a A)^{-1}, \quad (15.6)$$

де ϵ_1 та ϵ_2 – складові діелектричної сталої тришарової структури, бар'єрний шар – хемосорбційний шар – вода; ω – циклічна частота.

На рисунку 15.4 наведені залежності ємності і електричного опору сенсорів відносної вологості. Сенсор має такі конструктивні характеристики: площа $A = 0,07 \text{ см}^2$, товщина плівки $d = 10 \text{ мкм}$, товщина бар'єрного шару $b = 9 \text{ нм}$, пористість $a = 16\%$. Плівки Al_2O_3 можна використовувати в якості сенсорних елементів сенсорів вологості в режимі вимірювання як ємності, так і опору, але в останньому випадку чутливість сенсора майже на 30% менша. Зменшення товщини шару Al_2O_3 до $0,2 \text{ мкм}$ і менше, дає змогу значно збільшити швидкість сенсора, його чутливість у області низьких концентрацій вологи ($< 1\%$) і одночасно зменшити розміри. Подальший крок у цьому напрямку – перехід на інтегральні технології. Сенсор (рис. 15.5) формується на хімічно окислений кремнієвій підкладці. Після напилення алюмінію електрохімічним окисленням отримують плівку пористого Al_2O_3 товщиною менше $0,25 \text{ мкм}$. Після цього наносять верхній вологопроникний золотий електрод

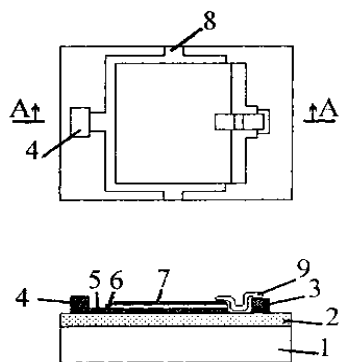


Рис. 15.5. Конструкція алюмінієвооксидного сенсора вологості: 1 – кремнієва підкладка; 2 – окис кремнію; 3, 4 – контактні площадки; 5 – нижній алюмінієвий електрод; 6 – шар Al_2O_3 ; 7 – вологопроникний золотий електрод, 8 – доріжка для з'єднання з іншими сенсорами; 9 – додаткова металізація контактної золоті площадки.

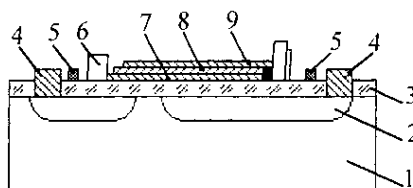


Рис. 15.6. Алюмінієвооксидний сенсор з підігрівачем та сенсором температури: 1 – кремнієва підкладка n-типу; 2 – дифузійні області p-типу; 3 – шар SiO_2 ; 4 – контакти області p-типу; 5 – нагрівач (ніхром); 6 – контактні площадки; 7 – нижній алюмінієвий електрод; 8 – анодна плівка окису алюмінію; 9 – верхній золотий електрод.

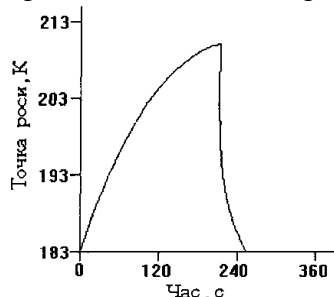


Рис. 15.7. Перехідна характеристика алюмінієвооксидного сенсора вологості.

товщиною 10-50 нм і додатковий шар золота на поверхню контактної площадки для забезпечення більш надійного контакту.

В іншому виконанні сенсора, крім вологочутливої структури на основі окису алюмінію (або двоокису кремнію) товщиною менше 0,25 мкм, є термостабілізуючий нагрівач і сенсор температури (рис. 15.6). Нагрівач виконаний з ніхрому. Сенсор температури можна сформувати на периферії кристала розміщенням додаткового резистора з нікелю чи платини або, як показано на рис. 15.6, р-п-перехід, сформований в матеріалі підкладки. Температура поверхні дорівнює 323 ± 1 К і підтримується постійною, що дає змогу проводити вимірювання вологості при температурі аналізованого середовища від 263 до 343 К, а також виключити деградацію характеристик сенсора в середовищі з вологістю 95-100%. Потужність, що витрачається на нагрівання чутливого елемента, не перевищує 500 мВт. Діапазон вологості, що вимірюється за точкою роси знаходиться в межах від 263 до 313 К. Сенсор дає змогу проводити вимірювання як у газах, так і в органічних рідинах. Перехідна характеристика сенсора представлена на рис. 15.7. При зміні вологості від 0,1 до 1 ppm стала часу не перевищує 60 с.

Результати дослідження сенсорів вологості на основі Al_2O_3 показали, що розкид чутливості від сенсора до сенсора може дорівнювати 100%, тобто необхідно провести індивідуальне калібрування кожного приладу. ТКО чутливості сенсора становить 4-5% /град. Чутливість сенсора зменшується з часом, а також після відпалу при 373 К. Витримування сенсора в вакуумі збільшує його чутливість, але навіть зберігання приладу в атмосфері азоту не оберігає від часового дрейфу чутливості. До недоліків сенсора відноситься і гістерезис його характеристик, що пов'язано з розбіжностями в процесах осадження і виведення вологи в порах. Ці самі процеси призводять до того, що час відгуку сенсора при малій відносній вологості (менше 20%) майже на порядок менший часу відгуку при високій вологості (більше 40%).

Технологія виготовлення сенсорів на основі Al_2O_3 має істотний вплив на характеристики приладу, оскільки визначає структуру чутливого елемента плівки. Тому хороші параметри мають сенсори, виготовлені електролітичним анодуванням у 10%-ому розчині H_2SO_4 чистої (99,9%) алюмінієвої фольги при температурі електроліту 279 К (густина струму – 10 mA/cm^2 , початкова напруга формування плівки – 13 В) протягом 40 хв. Золоту плівку товщиною 300 мкм наносять вакуумним напиленням. При цьому отримані такі параметри окисної плівки: загальна товщина – 10 мкм; діаметр пор – 13 нм; пористість – 16%; товщина бар'єрного шару – 13 нм; кількість пор в сенсорі діаметром 3 мм – $8,4 \cdot 10^9$.

Сенсори вологості на основі Al_2O_3 знайшли широке використання. Найчастіше вони використовуються в якості аналізатора вологи в корпусах сучасних інтегральних схем. Водяна пара, як було встановлено, відіграє важливу роль в утворенні дефектів ІС, викликаючи порушення металізації та міграцію металів на поверхні схеми, пасивуючих шарів і т.д. Виготовлення таких сенсорів вологості на одній підкладці з мікросхемою дає змогу оцінити рівень вологи в корпусі навіть в процесі експлуатації, встановити кореляцію між вологістю, режимом роботи і причинами відмовлень схеми, прогнозувати збої і дати рекомендації для покращення якості корпусів, а також вибору оптимальних матеріалів схеми, які також виділяють вологу в корпус, а отже – підвищити надійність мікросхем. Сенсори, повністю виготовлені за тонкоплівковою технологією, мають високі значення температурної і часової стабільності чутливості. Циклічний відпал у сухому азоті протягом 500 год. при температурі 443 К практично не змінює калібрувальну характеристику сенсора, а циклічні дослідження при зміні вологості від 2 до 10000 ppm не показали суттєвого гістерезису. Часовий відгук сенсора неоднозначний при збільшенні і зменшенні вологості: в інтервалі від 1 до 1000 ppm він становить 17 с, а в зворотному напрямку – 10 с.

15.3. Плівкові сенсори для вимірювання точки роси

Робота мікроелектронних сенсорів для визначення точки роси базується на виявленні конденсованої води на охолоджуваній до температури випадання роси поверхні сенсора. При цьому, як правило, вихідний електричний сигнал пропорційний до маси вологи, яка випала в конденсат. Наявність конденсованої вологи може визначатись оптоелектронним, електричним або кольорометричним методами, а температура точки роси реєструється в такому сенсорі за допомогою температурного сенсора. Сенсор можна виконати за гібридною або напівпровідниковою технологією.

Структура сенсора для вимірювання поверхневого опору або ємності між двома гребінчастими електродами показана на рис. 15.8.

Для непористих матеріалів, особливо скла, електричний опір між двома поверхневими електродами істотно залежить від кількості вологи на поверхні. Адсорбція водяних парів використовується в такому випадку для вимірювання відносної вологості, а їх конденсація – для вимірювання точки роси. Спосіб визначення точки роси за результатами вимірювання електричної провідності характеризується високою чутливістю, але поганою відтворюваністю, і тому менше використовується, ніж метод, який базується на визначенні ємності між електродами та характеризується

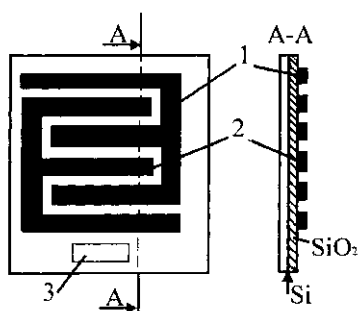


Рис. 15.8. Структура двох-електродного сенсора точки роси резистивно-ємнісного типу: 1, 2 – електроди; 3 – температурний сенсор.

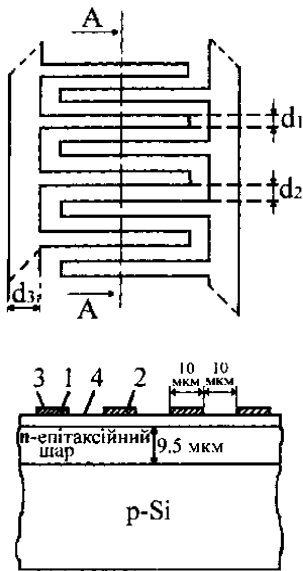


Рис. 15.9. Структура напівпровідникового сенсора точки роси ємнісного типу (без сенсора температури): 1, 2 – електроди; 3- шар алюмінію; 4- шар SiO_2 .

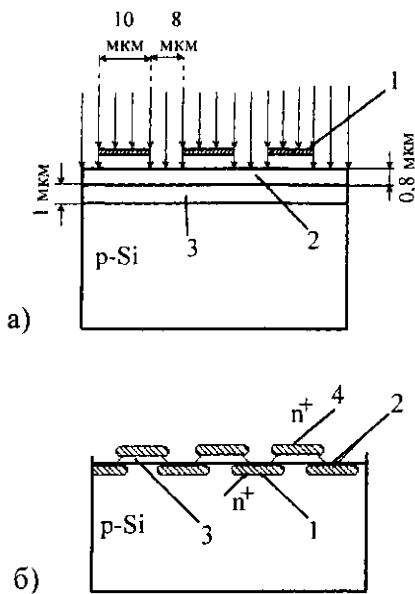


Рис. 15.10. Технологічний етап виготовлення (а) та структура модернізованого напівпровідникового сенсора точки роси ємнісного типу (б): 1-маска; 2-полікристалічний кремній; 3-шар SiO_2 ; 4-монокристалічний кремній.

високою відтворюваністю результатів при меншій чутливості.

Структура мікроелектронного ємнісного сенсора точки роси, виготовленого на кремнієвій підкладці, показана на рис. 15.9. Відстань d_3 в гребінчастих електродах у декілька раз більша за d_1 для зведення до мінімуму послідовного опору електродної смужки, крім цього, величини d_1 та d_2 намагаються зробити якнайменшими для збільшення міжелектродної ємності. Даний сенсор виготовляється за стандартною біполярною технологією; як сенсор температури використовують біполярний транзистор n-p-n-типу, сформований дифузією в епітаксійному n-шарі, вирощений на високоомній підкладці p-типу.

Матеріалом електродів є алюміній, або легований полікристалічний кремній. Міжелектродна ємність даної конструкції становить $17,2 \pm 0,3$ пФ, причому, близько 85% її величини зумовлено ємнісним зв'язком через підкладку і лише 15% – зв'язком через надповерхневий простір, який залежить від вмісту вологи, і тому є корисною частиною ємності.

Для збільшення ємності між електродами розроблена структура сенсорів МДН-типу зі зменшеною відстанню між електродами і збільшеною густиною компонування (рис. 15.10).

Роль першого електрода виконують смужки легованого полікристалічного кремнію n⁺-типу, гребінчаста структура якого сформована травленням тонкої полікристалічної кремнієвої плівки, а другим електродом служить дифузійний n⁺-шар, створений у високоомній підкладці p-типу дифузією, в процесі якої перший електрод служить маскою. При одній і тій самій площі цей сенсор має в три рази більшу початкову ємність та вищу чутливість на початковому етапі випадання роси, порівняно з попередньою конструкцією.

Підвищення чутливості пов'язане з тим, що конденсація починається в першу чергу у невеликому проміжку між електродами, утвореному підтравленням окислу SiO_2 . Ємність вимірюється шляхом подачі синусоїдального 1-В сигналу на частоті 100 кГц з одночасною реєстрацією ємнісної складової струму. Резистивна складова струму при частотах більших від 100 кГц стає набагато меншою ємнісною, але для послаблення впливу цієї складової та підвищення точності реєстрації точки роси в схемі визначення використовують фоточутливу схему. В майбутньому точність вимірювання буде обмежуватись помилками калібрування, нелінійністю та нестабільністю температурного сенсора, який входить до складу сенсора точки роси. Для описаних вище сенсорів у ході багатьох експериментів було показано, що відтворюваність результатів вимірювання становить 0,1 К.

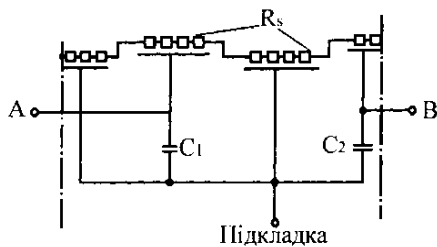


Рис. 15.11. Еквівалентна електрична схема сенсора вологості меандрового типу: R_s – розподілений поверхневий опір; C_s – розподілена ємність; $C_{1,2}$ – ємності зв'язку електродів з підкладкою через термічний окисел.

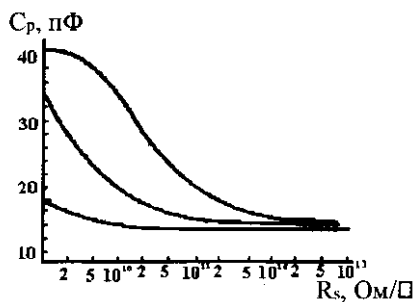


Рис. 15.12. Залежність міжелектродної ємності структури від поверхневого опору.

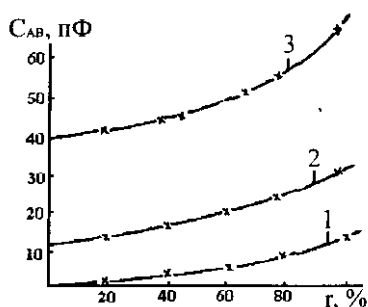


Рис. 15.13. Залежність міжелектродної ємності структури від відносної вологості для трьох режимів ввімкнення.

Відомий сенсор, призначений як для вимірювання відносної вологості, так і точки роси, в якому є тонке діелектричне покриття, яке охороняє електроди "меандрового" типу. З одного боку діелектричне покриття служить активною поверхнею, адсорбуючої молекули водяної пари, з другого боку – блокує постійну складову струму через структуру, тому пристрій, на відміну від сенсора, показаного на рис. 15.9, працює в НЧ (1-100 кГц) діапазоні.

У "меандровій" конструкції сенсора виводи 1 і 3 належать зовнішньому електроду А, а вивід 2 – внутрішньому електроду В. Вимірювання ємності проводиться між електродами А і В, при цьому електрод А сформований із металу з великим ТКО, що дає змогу використовувати його в якості терморезистора для вимірювання температури в сенсорі точки роси. Діелектричним покриттям є тонка двохшарова плівка з високою адсорбцією молекул води, отримана хімічним осадженням із парової фази, яка забезпечує велику ємність зв'язку між електродами і має поверхневий опір, суттєво залежний від вологи. Електричну схему сенсора можна представити у вигляді розподіленої RC-лінії (рис. 15.11).

Результати розрахунку еквівалентної схеми показані на рис. 15.12.

Міжелектродну ємність вимірювали при різних увімкненнях підкладки структури: "плаваюча" підкладка – ємність вимірюється між електродами при відсутності контакту в підкладці; заземлена підкладка – один із електродів і підкладка заземлені; активна підкладка – вимірюється ємність між електрично з'єднаними електродами та заземленою підкладкою.

На рис. 15.13 показані експериментальні залежності для сенсора з розмірами: $d_1 = 8$ мкм; $d_2 = 8$ мкм; $d_3 = 4$ мкм при використанні першого режиму увімкнення. Якщо максимальна ємність спостерігається при мінімальному поверхневому опорі, наприклад у точці роси, то мінімальна її величина – при послідовному з'єднанні C_1 і C_2 .

Максимальна чутливість приладу відповідає великому поверхневому опорі (приблизно 10^{11} Ом/□), і одночасно діапазон цього опору досить вузький – два порядки. Аналіз результатів дослідження еквівалентної схеми показав, що найбільший вклад у чутливість приладу вносить поверхня діелектрика безпосередньо над електродами, а не між ними.

Сенсори точки роси потребують використання пристроїв охолодження активної поверхні приладу. В інтегральних сенсорах з цією метою доцільно використовувати термоелектричні ефекти Пельтьє і Томпсона, які полягають у тому, що електричний струм у замкнутому колі, складеному з матеріалів різнорізною

або неоднорідною структурою, викликає локальне видалення та відбір тепла в різних ділянках кола. Сучасні термомодулі, які виготовляються серійно для охолодження ІС, працюють за принципом охолодження обох переходів, що одержується протіканням струму з напівпровідника р-типу в метал та з металу в напівпровідник n-типу. Серійні термомодулі мають охолоджуючу здатність біля 0,3 Вт. Сам процес охолодження сенсора точки роси може бути або циклічним, або неперервним. В циклічному режимі температуру зменшують зі сталою швидкістю, коли вихідний сигнал сенсора, як показує дисплей, є функцією температури. Інформацію про температуру точки роси отримуємо, обробляючи отриману залежність. В неперервному режимі сенсор точки роси використовують у колі зворотного зв'язку (рис. 15.14).

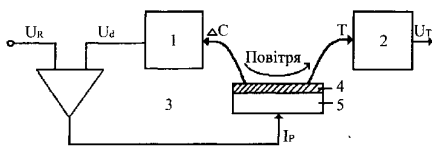


Рис. 15.14. Блок-схема виміральної системи точки роси неперервної дії: 1 – схема детектування роси; 2 – схема вимірювання температури; 3 – керуючий підсилювач; 4 – активна поверхня; 5 – термоелемент.

Температуру сенсора в цьому випадку підтримують відповідною рівністю між опорною напругою U_R та вихідною напругою U_d схеми детектування роси. Кількість вологи на чутливій поверхні сенсора визначається опорною напругою. При постійній кількості роси на поверхні сенсора випаровування та конденсація знаходяться в рівновазі, а покази температурного сенсора відповідають точці роси.

Лекція 16

НАПІВПРОВІДНИКОВІ ТЕНЗОСЕНСОРИ

План

16.1. Напівпровідникові тензосенсори

16.1.1. Конструкція інтегральних тензосенсорів

16.1.2. Інтегральний сенсор тиску на основі біполярних тензотранзисторів

16.1.3. Тензорезистивні плоскі сенсори відносного і абсолютного тиску

16.1.4. Інтегральний балочний акселерометр

16.2. Напівпровідникові сенсори газу резистивного типу

16.2.1. Матеріали чутливих елементів сенсорів

16.2.2. Конструкції сенсорів

16.2.3. Технологія виготовлення сенсорів

16.2.4. Тонкоплівкові сенсори

16.2.5. Товстоплівкові сенсори

Література

16.1. Елементи теорії мікроелектронних сенсорів / За ред. З.Ю. Готри – Львів: Ліга-Прес, 2001. – 636 с.

16.2. Полупроводниковые и твердоэлектролитные сенсоры / П.М. Таланчук, Б.П. Шматко, Л.С. Заика, О.Ю. Цветаева. – К.: Техніка, 1992. – 224 с.

16.3. Горбачев В.В., Спицана Л.Г. Физика полупроводников и металлов. – Мир: Металлургия, 1982.-336 с.

16.4. Газочувствительные датчики на основе металлоокисных полупроводников / Бутурлин А.И., Габузян Т.А., Голованов Н.А. и др. // Зарубежная электронная техника, 1985, N 3, с. 3 – 39.

16.1. Напівпровідникові тензосенсори

16.1.1. Конструкція інтегральних тензосенсорів.

Гібридні інтегральні перетворювачі тиску. На основі одного базового модуля – напівпровідникового кристала з кремнієвою мембраною – можна створити ряд типоміналів перетворювачів (абсолютного, відносного, диференційного) з діапазоном вимірювальних тисків від 10 кПа до 35 МПа.

Гібридна схема інтегрального перетворювача має всі базові елементи для здійснення повної функції перетворення: вакуумовану з одного боку для створення еталонного тиску, мембрану з основою, повну мостову тензорезисторну схему, електронну схему на інтегральних операційних підсилювачах для одержання стандартного вихідного сигналу, температурної компенсації, балансування тощо. На рис. 16.1 показана схема, виконана на трьох кристалах на спільній керамічній підкладці.

Чутливий елемент сенсора тиску являє собою кремнієвий кристал, в якому виконана прямокутна кремнієва мембрана; на решті площі розташована діодна схема термокомпенсації, що балансує міст резисторів, та

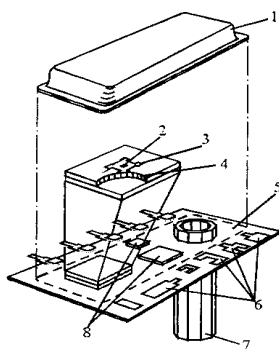


Рис.16.1. Конструкція гібридної інтегральної схеми перетворювача: 1 – кришка; 2 – термочутливий елемент; 3 – тензорезисторна мостова схема; 4 – кремнієва мембрана; 5 – керамічна плата; 6 – товстоплівкові резистори, що підстроюються; 7 – трубка для подачі тиску; 8 – підсилювачі.

стабілізатори напруги на опорному діоді.

На кристалі розміщені також буферні та вихідні підсилювачі. В центрі мостової тензорезисторної схеми, що являє собою квадратну рамку, розташований потужний транзистор, який використовується для нагрівання та підтримання температури перетворювача в процесі настроювання та випробувань. Емітерно-базовий перехід цього транзистора є стабільним термочутливим елементом. В результаті забезпечуються одночасні вимірювання температури і тиску.

Інтегральний тензорезисторний сенсор абсолютного тиску має структуру, показану на рис. 16.2.

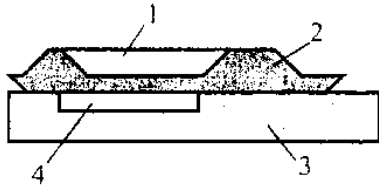


Рис. 16.2. Структура інтегрального перетворювача абсолютного тиску: 1 – кремнієва пластина; 2 – основа; 3 – скляна пластина; 4 – герметична камера.

Конструкція складається зі скляної плоскої кришки, в якій витравлено заглиблення, і кремнієвого кристала з тонкою мембраною, в якій виготовлено чотири тензорезистори, з'єднані в мостову схему. Мембрана оточена товстою кремнієвою основою у формі кільця. Заглиблення в скляній пластині розташоване над мембраною, а сама пластинка скріплена з кремнієвим кристалом за допомогою анодної посадки. Анодне з'єднання здійснюється нагріванням скляної і кремнієвої пластинок до температури 400°C з подальшою подачею напруги 600 В на кремній відносно заземленої скляної пластинки. Анодна посадка забезпечує абсолютно герметичне з'єднання, завдяки чому заглиблення в скляній пластинці разом з мембраною утворює замкнуту герметичну камеру.

Перетворювач з розмірами $1,5 \times 2,0 \times 0,2$ мм забезпечує чутливість $0,2 \text{ мВ}/(\text{кПа}\cdot\text{В})$. Температурний дрейф еквівалентний 250 Па на 1°C . Вихідна напруга лінійно змінюється до тиску 40 кПа .

16.1.2. Інтегральний сенсор тиску на основі біполярних тензотранзисторів. Конструкція інтегрального сенсору тиску з тензочутливими n-p-n-планарними та p-n-p-повздожніми тензотранзисторами має таку будову. Посередині кожної з чотирьох сторін квадратної мембрани, де вона закріплена з основою, розміщено по одному транзистору (два n-p-n та два p-n-p) з різноманітною орієнтацією відносно краю мембрани. В центрі мембрани розташовані два транзистори: один n-p-n, другий p-n-p. Такий набір компонентів дозволяє використовувати їх як незалежно, так і включати у складніші схеми для підвищення чутливості. Крім того, поза мембраною на основі кристала виготовлений планарний транзистор, який може використовуватися як для вимірювання температури, так і для компенсації температурних похибок.

Інтегральна схема сенсору тиску на тензочутливих біполярних транзисторах. Новим напрямком в інтегральних тензосенсорах є створення схем, що складаються з великої кількості компонентів, чутливих до дії тиску. В результаті на пружному елементі

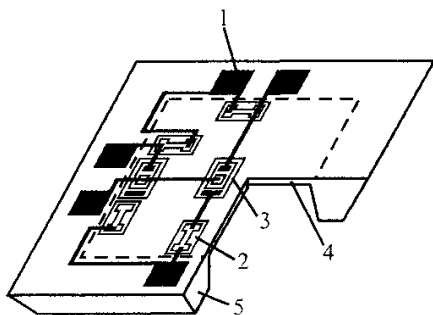


Рис. 16.3. Чутливий елемент інтегрального перетворювача тиску на тензорезисторах та тензо-транзисторах:

- 1 – контактна площадка;
- 2 – тензорезистор;
- 3 – тензотранзистор;
- 4 – мембрана;
- 5 – основа

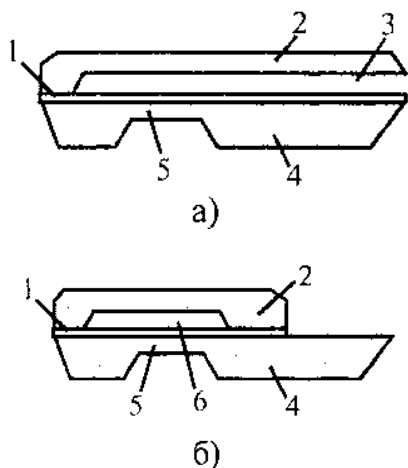


Рис. 16.4. Повздожняя мікро-конструкція перетворювачів відносного (а) та абсолютного і надлишкового тисків (б): 1 – з'єднувальний шар; 2 – кришка; 3 – атмосферний тиск; 4 – основа; 5 – мембрана; б – порожнина атмосферного тиску.

створюються інтегральні мікросхеми, кожний компонент яких залежить від механічного напруження, що виникає в пружному елементі при зовнішній дії. Знак і співвідношення чутливостей окремих компонентів такі, що чутливість усієї схеми загалом на 1-2 порядки вища, ніж у випадках мостових тензорезисторних схем. Таким чином вирішується проблема підвищення чутливості перетворювача при збереженні розмірів та зменшення розмірів при збереженні чутливості. Крім того, можна підвищити потужність вихідного сигналу за рахунок більшого розподілення компонентів на поверхні пружного елемента, покращити температурну стабільність перетворювача та підвищити порогову чутливість.

На рис. 16.3 наведено приклад інтегральної схеми подібного сенсора тиску. Сенсор складається з кремнієвої пластинки, зорієнтованої в площині, в якій є тонка мембрана, на якій за допомогою інтегральної технології виготовлена інтегральна схема. Одне плече схеми складається з центрального транзистора і двох резисторів, один з яких з'єднується струмопровідною доріжкою з колектором транзистора, а другий – з базою. Друге плече схеми утворене транзистором і резисторами, розташованими в центральній частині однієї з сторін квадратної мембрани. Знаки тензочутливих відповідних компонентів (транзисторів, колекторних та базових резисторів) в різних плечах схеми протилежні. За рахунок цього вдається вдвічі підвищити чутливість.

Інтегральний сенсор тиску на основі МОП-транзисторів. Фактично МОП-транзистор можна розглядати як достатньо високоомний резистор, опір якого керується потенціалом затвора. Тому поряд з тензорезистивним ефектом в МОП-транзисторі має місце залежність щільності поверхневого заряду на межі Si-SiO₂ від механічних напружень. Малі розміри МОП-транзисторів та їх самоізоляція на загальній підкладці роблять їх перспективними для тензоперетворювачів.

16.1.3. Тензорезистивні плоскі сенсори відносного і абсолютного тиску. В зонд: для вимірювання внутрішньосерцевого кров'яного тиску кристал чутливого мембранного елемента кріпиться в точці катетера перпендикулярно до його осі.

Головні конструктивні частини перетворювача – кремнієва основа, на якій розташований пружний елемент (тонка квадратна мембрана) зі сформованими на ній тензочутливими елементами, і кришка, що забезпечує подачу еталонного тиску. Це може бути атмосферний тиск, фіксований тиск або вакуум. Відповідно перетворювачі можуть бути відносного, абсолютного та надлишкового тисків.

У першому випадку кришка перетворювача має отвір збоку для зв'язку з атмосферою (рис. 16.4, а).

В перетворювачах надлишкового і абсолютного тиску між кришкою та основою утворюється замкнута порожнина з еталонним тиском (рис. 16.4, б). Обидва конструктивні елементи – кришка і основа виготовляються з кремнію на основі планарної технології.

16.1.4. Інтегральний балочний акселерометр. Даний прилад являє собою кристал кремнію, на якому як одне ціле виготовлена тонка балка, що деформується при дії прискорення. Балка з'єднана одним кінцем з основою, а з трьох інших боків відокремлена від основи тонким щілинним отвором. У місці закріплення балки розміщено концентратор механічного напруження. Чутливим елементом є повна мостова тензорезисторна схема, виготовлена за планарною технологією інтегральних мікросхем, розташована в балці близько до поверхні і до місця закріплення. Тензорезистори виготовлені методом іонної імплантації домішок, що дозволило суттєво спростити технологію перетворювачів.

Подібний балочний перетворювач можна використовувати як лінійний акселерометр, перетворювач вібрацій та положення в просторі.

16.2. Напівпровідникові сенсори газу резистивного типу

16.2.1. Матеріали чутливих елементів сенсорів. Всі основні метрологічні характеристики сенсорів, до яких відносяться чутливість, швидкодія, точність вимірювання, вхідний і вихідний опір, потужність споживання, варіація показів, передавальна функція, в найбільшій мірі визначаються вибором ЧЕ. Матеріал, що використовується, крім свого основного призначення – селективно реагувати на газ, що вимірюється, – повинен мати стабільні термодинамічні, хімічні і електрофізичні характеристиками.

Серед найчастіше застосовуваних сполук, які використовують активний матеріал для визначення газів, є окиси металів з напівпровідниковими властивостями. Слід виділити напівпровідникові матеріали типу n і p. До напівпровідникових матеріалів типу p, що найчастіше використовуються в газових сенсорах, належать прості окисли, такі як SnO_2 , ZnO_2 , Fe_2O_3 , WO_3 , TiO_2 і складні комплекси окислів. Вони зменшують свій опір в оточенні редуційних газів, таких як: H_2 , CH_4 , CO , $\text{C}_2\text{H}_5\text{OH}$ (пари), а збільшуються під впливом окислювальної атмосфери. Приклад часових відгуків найчастіше застосовуваних напівпровідникових газочутливих окислів до деяких активних (агресивних) газів представлено на рис. 16.5 і 16.6.

Промислове значення отримали сенсори на основі SnO_2 , ZnO , Fe_2O_3 , при цьому сенсори на двоокисі Sn використовуються найчастіше.

В табл. 16.1. представлені складові газу, виявлені за

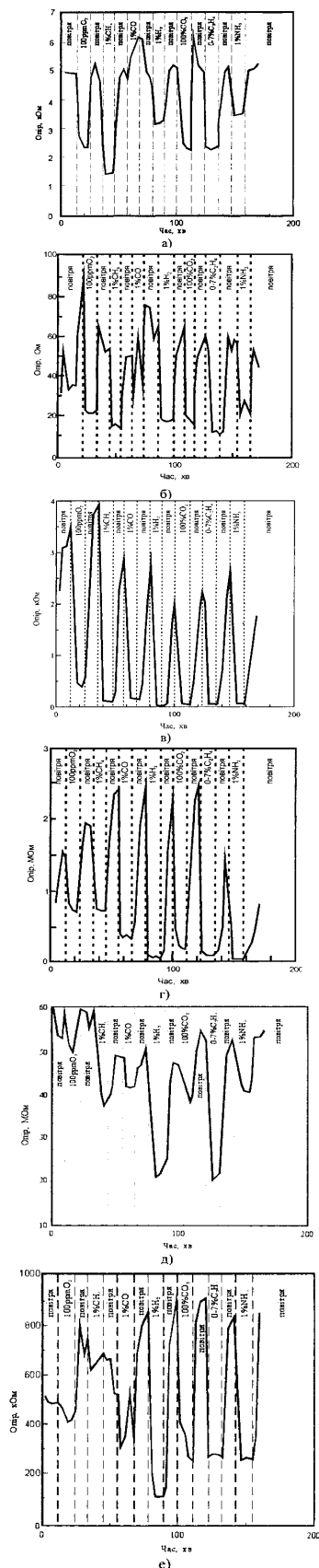


Рис. 16.5. Часові відгуки деяких окислів на різні активні гази (при робочій температурі 500° C): а) SnO₂; б) In₂O₃; в) ZnO; г) MoO₃; д) Ga₂O₃; е) Nb₂O₃.

допомогою сенсорів на основі SnO₂.

Чутливий елемент напівпровідникових сенсорів для вимірювання парціального тиску кисню в газах (до 10⁻¹⁵ Па) виготовляється з оксиду титану TiO₂, оксиду цинку ZnO, оксиду ніобію Nb₂O₅ або кобальту CoO тощо.

Для вимірювання парціального тиску водню використовується ЧЕ з оксидів вольфраму WO₃, паладію PdO платини PtO, титану TiO₂ і титанату барію BaTiO₂. Плівку з паладію і платини застосовують для збільшення чутливості сенсора. З цією метою можна використати також керамічні пористі плівки, які дозволяють водню проникати до ЧЕ через пори і затримувати інші гази з молекулами більшого діаметра. При вимірюванні парціального тиску оксиду вуглецю CO застосовують ЧЕ з оксидів різних металів. Найчастіше для вимірювання тиску CO використовується ЧЕ з оксиду олова SnO₂. У цьому випадку опір ЧЕ при контакті з газом зменшується в 150 разів, а при наявності водню – в 1,5 рази.

У сенсорів з добавкою до оксиду олова оксиду торію ThO₂, що характеризується сильною дегідратаційною активністю, чутливість до оксиду вуглецю зростає, а до водню залишається практично такою ж, як і в елемента без добавок. Характеристики матеріалів ЧЕ напівпровідникових плівкових сенсорів для аналізу різних газів наведені в табл. 16.2.

Чутливі елементи з оксиду срібла Ag₂O також реагують на зміну концентрації оксиду вуглецю в повітрі, що відбувається з швидкістю 1-10 %/с на рівні 100 – 800 ppm. Час встановлення адсорбційної рівноваги можна істотно збільшити, підвищивши робочу температуру, а також додавши оксиди інших металів (оксид кобальту CoO₂ або оксид марганцю MnO₂ (0,1 – 0,5%). При відносно низьких концентраціях вуглецю в повітрі (-30 ppm) реакція є оборотною.

Таблиця 16.1.

Газові складові для повітря і діапазони їх концентрації виявлені сенсорами на основі SnO₂

| Вид газу | Концентрація газу |
|----------------------------------------------------------------------------------|-------------------|
| H ₂ | 500... 10000 ppm |
| CH ₄ , C ₃ H ₈ , C ₄ H ₁₀ | 500... 10000 ppm |
| CO | 50... 1000 ppm |
| C ₂ H ₅ OH | 1... 300 ppm |
| CH ₃ COOH | більше 50 ppm |
| NH ₃ | більше 30 ppm |
| NO ₂ | більше 100 ppm |
| AsH ₂ | більше 0,6 ppm |
| H ₂ S | 5... 100 ppm |

Таблиця 16.2.

Характеристики матеріалів ЧЕ напівпровідникових
плівкових сенсорів

| Матеріал чутливого елемента | Газ, що вимірюється | Робоча температура, К | Примітка |
|-------------------------------------------|-------------------------|-----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| ZnO | Кисень -O ₂ | 523 - 1203 | Вимірювання атомарного кисню O в N ₂ , H ₂ та інертних газах |
| TiO ₂ +Pt, Rh | Кисень -O ₂ | — | — |
| Co _{1-x} Mg _x O | Кисень -O ₂ | — | 0,1<x<0,95 |
| ZnO | Кисень -O ₂ | — | Вимірювання атомарного кисню в середовищі H ₂ |
| ZrO ₂ | Кисень -O ₂ | — | Вимірювання термоелектронної емісії або роботи виходу |
| Pb, Sb, Ag+S, Se, Fe | Кисень -O ₂ | 293-473 | Вимірювання O ₂ до 10 ⁻⁴ % |
| SrFeO _{3-x} | Кисень -O ₂ | 623 | Селективність підвищується шляхом введення La, Y |
| SnO ₂ , TiO ₂ , ZnO | Кисень -O ₂ | ≈973 | — |
| Nb ₂ O ₅ , CoO | Кисень -O ₂ | 1273 | Вимірювання до 10 ⁻⁹ Па O ₂ в негорючих газах і в окисно-відновлювальних сумішах |
| TiO ₂ | Кисень -O ₂ | — | Час відклику 20 мкс |
| Сплави з Pb, Ti, V, Zr, Pd, Ta | Водень - H ₂ | 348 | Вимірювання здійснюється на змінному струмі в діапазоні 0 - 10% H ₂ |
| WO ₂ | Водень - H ₂ | 513-553 | Каталізатор Pt, Rh, Ir, An, Pd; нечутливий до вуглеродів; вимірювання у діапазоні 10 - 10 ² Па; час перехідного процесу 5 - 10 хв |
| PdO або PtO | Водень - H ₂ | 293 | — |
| TiO ₂ | Водень - H ₂ | 293 - 573 | Вимірювання в атмосфері чистого H ₂ при концентраціях до 100 ppm; час відгуку 1 с |
| Pd | Водень - H ₂ | — | — |
| 25% Pd + 40% Ag | Водень - H ₂ | — | — |
| In ₂ O ₃ | Водень - H ₂ | >773 | Вимірювання атомарного H до 5·10 ⁻⁵ ppm |
| SnO ₂ | Оксид вуглецю - CO | >523 | Каталізатор 0,5 - 10% Pt; низька чутливість до H ₂ , CH ₄ ; діапазон вимірювання 5·10 ⁻² - 10 ⁵ ppm |
| Ag ₂ O | Оксид вуглецю - CO | 373 | Каталізатор 0,5 - 10% CoO, MnO ₂ ; діапазон вимірювання 100 - 800 ppm |
| Cu ₂ O, CuO | Оксид вуглецю - CO | — | — |

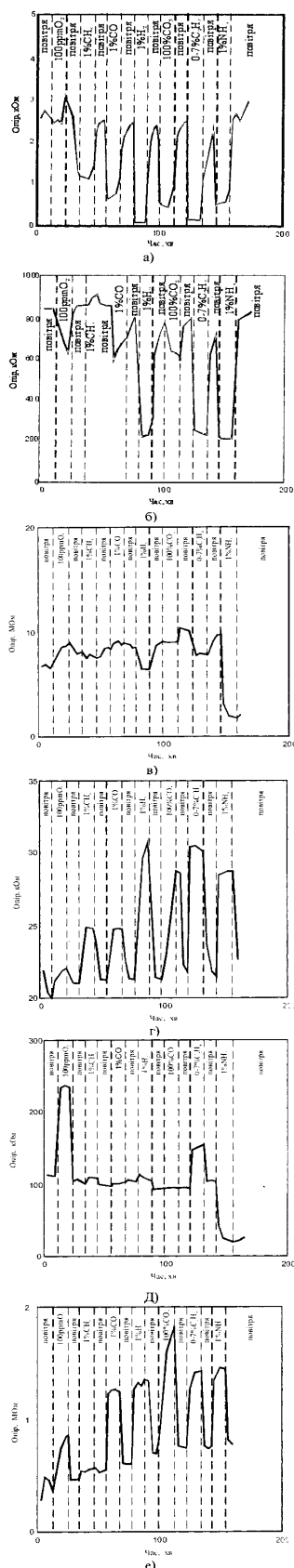


Рис. 16.6. Часові відгуки складних окислів на різні агресивні гази (при робочій температурі 500° С): а) BaSnO₃; б) TiNb₂O₇; в) BaSn_{0,8}Ti_{0,2}O₃; г) Y₂TiFeO₇; д) BaTiO₃; е) Bi₂Sn₂O₃.

| | | | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------|-----------|-----------------------------------------------------------------------------------------------------------------------------|
| SnO ₂ або ZrO ₂ , NiO або Mn | Оксид вуглецю – CO | 623- 823 | Чутливий також і до H ₂ |
| SiO ₂ + SnO ₂ з Sb або Bi + SnO ₂ з Pt або B + Bt з Au | Оксид вуглецю – CO | | Вимірювання CO в присутності NO _x в діапазоні 0 –100 ppm |
| SnO ₂ +Sb ₂ O ₃ або Bi ₂ O ₃ | Оксид вуглецю – CO | — | — |
| V ₂ O ₅ , CoO + MgO, SnO ₂ , ZnO, AFe ₂ O ₄ | Оксид вуглецю – CO | 723 - 823 | A - Mg, Mn, Fe, Co, In, Pb |
| SO ₂ , ZnO, Fe ₂ O ₃ , V ₂ O ₅ , WO ₃ , ThO ₂ , MoO ₂ , CdO, PdCrO ₄ | Оксид вуглецю – CO | 423 - 523 | Проводить вимірювання H ₂ , CO, C _n H _m |
| TiO ₂ | Двооксид вуглецю - CO ₂ | — | — |
| MnO ₂ | Двооксид вуглецю - CO ₂ | — | — |
| BaTiO ₃ | Двооксид вуглецю - CO ₂ | — | Селективний в присутності кисню і парів води |
| SnO ₂ | Пари води - H ₂ O | — | — |
| SiO ₂ | Пари води - H ₂ O | | Вимірювання точки роси при пониженні температури у замкнутому об'ємі при концентрації парів води менше 10 ⁻⁴ ppm |
| Al ₂ O ₃ | Пари води - H ₂ O | | Вимірювання ємності в системі метал-діелектрик-метал Al - Al ₂ O ₃ -Au |

З метою забезпечення селективності і збільшення чутливості напівпровідникових ЧЕ до газу, що вимірюється, застосовують каталізatori (активатори). Так, чутливість сенсора кисню на основі ферооксиду стронцію SrFeO_{3-x} можна підвищити легуванням півки лантаном La або ітрієм Y. Каталізаторами у водневих сенсорах служать паладій, платина або їх солі. Селективність сенсорів оксиду вуглецю підвищують введенням в основу ЧЕ платини або її солей, а також оксиду торію. Легування оксиду срібла оксидами металів більш високої валентності підвищує його окислювальну здатність. Каталізатором може служити також тонка півка, що має адгезією до півки ЧЕ (наприклад, дрібнодисперсна платина або паладій), з півковою острівковою структурою. Для збільшення адгезії півки ЧЕ до підкладки з скла в матеріал ЧЕ іноді включається оксид кремнію SiO₂.

Матеріал ЧЕ напівпровідникового сенсора повинен характеризуватися такими властивостями: високою

чутливістю і селективністю до газу, що вимірюється; низьким ТКО, що виключає або знижує значення температурної погрішності вимірювання, фізико-хімічною сумісністю з матеріалами інших елементів сенсора; високою температурою плавлення, низькою пружністю пари, високою хімічною стійкістю, тобто високою стабільністю оксиду.

16.2.2. Конструкції сенсорів. До основних елементів конструкції сенсорів відносяться чутливий елемент, нагрівач, підкладка, електроди (контактні площадки), засоби контролю і підтримки необхідної робочої температури. Нагрівач, підкладка, електроди і засоби для вимірювання температури дозволяють забезпечити найбільш повну реалізацію властивостей ЧЕ, що визначають метрологічні характеристики сенсора.

Значне поширення отримали конструкції напівпровідникових сенсорів з об'ємними ЧЕ у вигляді спеченого оксиду металу. Всередині такого сенсора вміщується електрод, який одночасно служить нагрівачем прямого і непрямого нагріву.

До недоліків сенсорів з об'ємними ЧЕ відносять недостатню експлуатаційну надійність елементів перетворювачів і металокерамічних сполук в умовах великих температурних градієнтів і різких теплових ударів; складність технології виготовлення, що вимагає забезпечення високих температур (1673-1973 К) і тиску (до 108 Па) в ході технологічних процесів; недостатню відтворюваність властивостей отриманих ЧЕ і необхідність метрологічного контролю всіх сенсорів; високу їх собівартість.

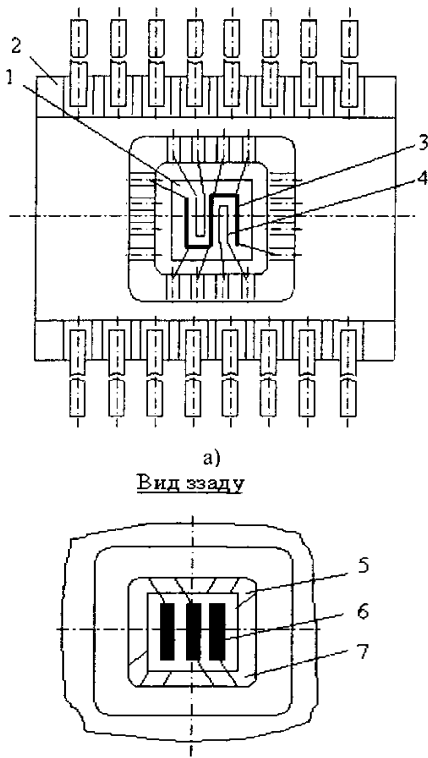
Плівкові сенсори практично позбавлені вказаних недоліків. Вони мають високу відтворюваність характеристик структури, складу і електроопору ЧЕ, а також простіші і дешевші у виготовленні.

Конструкція типового плівкового напівпровідникового сенсора наведена на рис. 16.7. Сенсор включає корпус і модуль, що приварюється до контактних площадок корпусу за допомогою позолочених виводів.

Несучим елементом модулів напівпровідникового сенсора є керамічна підкладка, на якій з обох боків розташовані їх елементи.

На робочому боці напівпровідникового сенсора розташовані чутливий елемент і контактні площадки для виведення сигналу.

При розробці планарних конструкцій сенсорів основною проблемою є забезпечення стабільності їх параметрів з урахуванням роботи в окислювально-відновних середовищах при підвищених температурах. До параметрів, що визначають цю стабільність, відносяться температурні коефіцієнти розширення (ТКР) матеріалів, адгезія матеріалів, робота виходу матеріалів



6)
Рис. 16.7. Типова конструкція напівпровідникового плівкового сенсора: 1 – модуль; 2 – корпус; 3 – нагрівач; 4 – терморезистор; 5 – чутливий елемент; 6 – електроди; 7 – виводи.

для отримання мінімального значення бар'єра Шоттки при їх контакті.

Як показує досвід, відтворюваність і стабільність електрофізичних властивостей плівок багато в чому визначаються правильним вибором матеріал(ів) підкладки. Він повинен мати невеликі питому електропровідність і хімічну інертність, високу діелектричну міцність і теплопровідність, хімічну і фізичну стійкість при термообробці у вакуумі або на повітрі, коефіцієнт теплового розширення, близький до коефіцієнта напилених шарів, шорсткість не нижче за 10 класу чистоти, бути дешевим. З полікристалічних підкладок вказані вимоги задовольняє полікор, що містить 90 % оксиду алюмінію.

Матеріал підкладки і попередня обробка її поверхні багато в чому визначають процес формування плівок і їх адгезію. Перед напилюванням плівок елементів сенсора необхідно провести комплексну обробку поверхні підкладки (очищення хімічним розчинником, термообробку при високій температурі, іонне бомбардування в плазмі аргону або іншого інертного газу).

Матеріали, що використовуються для виготовлення нагрівника, терморезисторів і електродів, повинні відповідати таким основним вимогам: високої електропровідності; співмірності їх КТР з КТР підкладки і ЧЕ; здатності розпилюватися при порівняно невисокій температурі; хорошій адгезії до матеріалу ЧЕ і підкладки; високої корозійної стійкості, термостійкості; низької здатності до рекристалізації під впливом високих температур; можливості приварювання виводів; високої механічної міцності; не взаємодіяти з навколишнім середовищем. Крім того, потрібно, щоб матеріали нагрівача і терморезистора мали високу теплопровідність для забезпечення швидкого прогрівання сенсора і лінійність температурного коефіцієнта опору (ТКО) для забезпечення однакової чутливості терморезистора до температури у всьому діапазоні температур, що досліджується. Процес нанесення матеріалів нагрівача і терморезистора повинен бути сумісний з сучасними методами фотолітографії. Необхідно, щоб матеріали легко взаємодіяли з реактивами, що використовуються для травлення.

У залежності від робочої температури для контактних площадок, терморезистора і нагрівача застосовують різні матеріали. Так, для $T < 573$ К використовують нікель (табл. 16.3), що має хороші адгезійні властивості. Лінійна залежність $R = f(T)$ у цьому випадку спостерігається лише до $T \sim 573$ К. На лінійній ділянці спостерігається досить високе значення ТКО, що важливо при використанні нікелю в якості терморезистора. Застосування нікелю обмежене, оскільки при $P_{O_2} \approx 10^{-2}$ Па

спостерігається його інтенсивне окислення і порушення адгезії до підкладки.

Таблиця 16.3.

| Характеристики матеріалів елементів напівпровідникових сенсорів | | | | | | | | |
|-----------------------------------------------------------------|---------------------------|-------------------------------------|----------------------|-----------------------------|-----------------------------|--------------------------|----------------------------|--------------------------------------|
| Матеріал | Тепло-провідність, Вт/м·К | Питомий опір, Ом·м | ТКО, К ⁻¹ | КТР, К ⁻¹ | Питома теплоємність, Дж/г·К | Температура плавлення, К | Густина, кг/м ³ | Сталя Гратки, а·10 ¹⁰ , м |
| Полікор | 16,66 – 29,30 | 10 ¹¹ – 10 ¹³ | — | 7 – 10·10 ⁻⁶ | 0,838 | 2320 | (3,7-3,9)·10 ³ | 7,9 |
| Молибден | 9,1·10 ⁴ | 5,15·10 ⁻⁸ | 1,1·10 ⁻³ | 5,1·10 ⁻⁶ | 0,251 | 2893 | 10,2·10 ³ | 3,1 |
| Хром | 3,8·10 ⁴ | 14·10 ⁻⁸ | 0,8·10 ⁻³ | 6,5·10 ⁻⁶ | — | 2163 | 7,19·10 ³ | 2,9 |
| Платина | 4·10 ⁴ | 9,4·10 ⁻⁸ | 3,9·10 ⁻³ | (8,9-10,2)·10 ⁻⁶ | 0,133 | 2052 | 21,5·10 ³ | 3,9 |
| Нікель | — | — | 0,1·10 ⁻³ | 16,5·10 ⁻⁶ | 0,439 | 1702 | 8,91·10 ³ | — |

При $T < 1073$ К найдоцільніше застосовувати молибден, залежність опору від температури для якого лінійна до $T = 1073$ К, ТКО становить приблизно $1,1$ К⁻¹.

При $T > 1073$ К молибден окислюється, і його опір змінюється. В цьому випадку необхідно використати багатошарові композиції з декількох матеріалів: нижній шар товщиною 10...20 нм дозволяє забезпечити необхідну адгезію до підкладки, верхній шар товщиною 300...800 нм – високу провідність, можливість припаювання і приварювання виводів. Як матеріал адгезійного шару можуть служити молибден, титан, ванадій, хром, цирконій, тантал і ніобій, як провідний шар платина, золото, срібло, мідь і алюміній.

Металургійна стабільність багатошарової системи визначається характером діаграми фазової рівноваги металів, швидкістю їх взаємної дифузії, адгезією одного до іншого. Існує два підходи до вибору багатошарової системи: матеріали, що створюють при безпосередньому контакті безперервний ряд твердих розчинів з високою швидкістю дифузійного «розмиття» меж (рис. 16.8), і бінарні композиції з мінімальною швидкістю взаємної дифузії, які не створюють проміжних фаз в температурному діапазоні роботи сенсорів (рис. 16.9). З врахуванням вимог до матеріалів підкладки, нагрівача і терморезистора найчастіше використовують такі багатошарові системи на полікорі: оксид алюмінію – молибден, оксид алюмінію – платина, оксид алюмінію – молибден – платина, оксид алюмінію – молибден – хром – платина. Характеристики цих матеріалів наведені в табл. 16.3.

Система оксид алюмінію – молибден $Al_2O_3 - Mo$ може використовуватися при робочих температурах сенсорів до 1273 К, оскільки після термообробки протягом 8 год при $T = 1073$ К механічна цілісність плівки молибдену не порушена, що свідчить про хорошу адгезію, КТР матеріалів близькі, отже, температурні напруження в плівці при нагріванні незначні.

У системі оксид алюмінію – платина $Al_2O_3 - Pt$ ТКО платини лінійний до $T = 1273$ К, при більш високих температурах платина випаровується, і її опір змінюється.

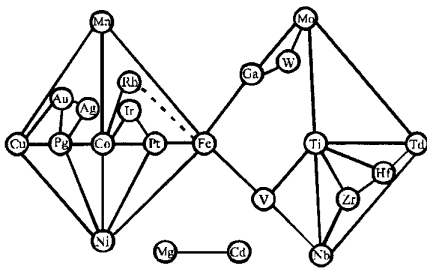


Рис. 16.8. Бінарні металеві композиції з необмеженою розчинністю (суцільною лінією показані системи з переходом без порушення гомогенності; пунктирною – системи з розчинністю в широкому діапазоні температур).

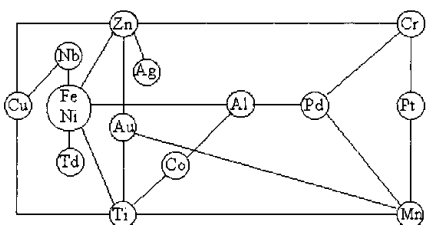


Рис. 16.9. Бінарні металеві композиції з мінімальною швидкістю взаємної дифузії металів.

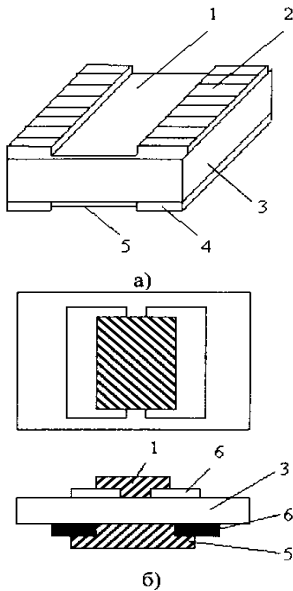


Рис. 16.10. Конструкції напівпровідникових плівкових сенсорів газу:

а) тонкоплівковий;

б) товстоплівковий: 1 – напівпровідник; 2, 4, 6 – молибден утворює безперервні тверді розчини; 3 – підкладка; 5 – нагрівач.

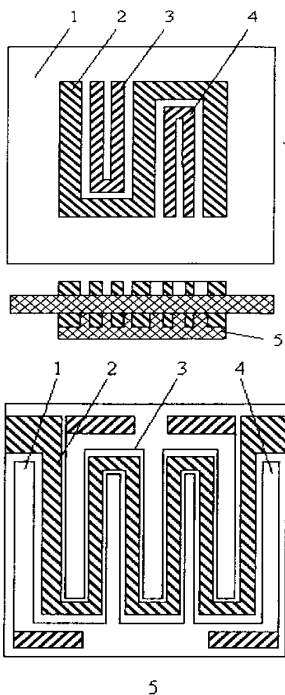


Рис. 16.11. Модулі напівпровідникового плівкового сенсора з різними варіантами топології:

- 1 – підкладка;
- 2 – нагрівач;
- 3, 4 – терморезистори;
- 5 – чутливий елемент.

Після термообробки при $T = 1073$ К протягом 8 год спостерігається погана адгезія платини до полікору, і приварювання виводів приводить до відшарування плівки платини внаслідок низької механічної міцності системи.

При взаємодії молибдену і платини в системі оксид алюмінію – молибден – платина Al_2O_3 -Mo-Pt утворюються обмежені тверді розчини і сполуки. Розчинність платини в молибдені незначна. При $T = 1273$ К вона становить 25%. У разі дифузійної взаємодії молибдену і платини при $T > 1273$ К відбувається утворення крихких інтерметалевих фаз $MoPt_3$, що погіршують зчеплення плівок з підкладкою, внаслідок чого відбувається руйнування плівок. Ця система може використовуватися при температурах до 1273 К.

При температурі вище за 1273 К застосовується система оксид алюмінію – молибден – хром – платина. Хром використовується в якості бар'єрного шару між молибденом і платиною. Система хром-платина утворює обмежені тверді розчини і сполуки. Розчинність платини в хромі при $T = 1173$ К становить 2,5%, розчинність хрому в платині при $T = 1273$ К – 71%. Система хром-платина утворює безперервні тверді розчини, розчинність хрому в молибдені і молибдену в хромі становить 100%.

При виборі матеріалів виводів потрібно враховувати можливість зварювання або паяння, а також забезпечення електропровідності контактів. Для з'єднання виводів і плівкових сенсорів рекомендується застосовувати дріт із золота і платини діаметром 30 ... 70 мкм.

У разі вибору корпусу для кріплення модуля сенсора необхідно враховувати механічну міцність сенсора в зборі з корпусом, термостійкість і мінімальне газовиділення при робочих температурах і тиску, відсутність клейових з'єднань, наявність вікна з розмірами, близькими до розмірів модулів, що дає можливість розварювати модуль сенсора з обох сторін.

Конструкції типових напівпровідникових тонкоплівкового та товстоплівкового сенсорів газу показані на рис. 16.10.

На робочому боці напівпровідникового сенсора розташовані чутливий елемент та контактні площадки для виводу сигналу, на протилежному боці – нагрівач і два терморезистори. На рис. 16.11. показаний модуль сенсора з різними варіантами топології.

Застосування техніки фотолітографії при серійному виробництві плівкових сенсорів дозволяє отримати високу точність при формуванні рисунка елементів, виконувати модуль розміром 1 мм^2 .

Розглянута конструкція уніфікована для виготовлення сенсорів парціального тиску різних газів. При створенні сенсора необхідно провести вибір матеріалу ЧЕ, який має необхідну чутливість і селективність до газу, що

вимірюється, а також визначити робочу температуру сенсора, яка забезпечує задані метрологічні характеристики матеріалів ЧЕ напівпровідникових сенсорів, що використовуються для аналізу кисню, водню, оксидів вуглецю і пари води.

Сенсор розглянутої конструкції дозволяє перетворювати парціальний тиск газу у вихідну величину – зміну опору ЧЕ або ЕРС, встановлювати і стабілізувати температуру ЧЕ, підключати перетворювач до інформаційно-виміральної системи. До переваг уніфікованого сенсора можна віднести мінімальні масогабаритні характеристики (маса в корпусі до 1 г, розміри сенсора в корпусі 18x12x2,5 мм), мінімальне енергоспоживання (до 1 Вт), незначне тепловідведення на елементах монтажу, широкий діапазон робочих температур.

Такі уніфіковані сенсори можуть бути використані для вимірювань полів парціального тиску кисню в суцільних вакуумних системах і в автоматизованих системах керування різними технологічними процесами.

Експлуатаційні і метрологічні характеристики сенсорів залежать від вибору матеріалів елементної бази.

Підкладка перетворювача виконується з газонепроникної, електроізоляційної і теплопровідної кераміки (глинозему, силікату алюмінію, алюмінату магнію, форстериту, цирконату кальцію, кварцу, оксиду кремнію), металів (неіржавіючої сталі або платини, титана, золота, срібла, нікелю, сплавів міді) або з кермету. Вона може складатися з двох частин, між якими вміщують нагрівач.

Нагрівач крім забезпечення необхідної температури виконує знегажування сенсора перед початком його експлуатації. Його виконують з металів з високим питомим опором – платини, вольфраму, молібдену, паладію, танталу або сплавів цих металів. Звичайно він має М- або S-подібну форму.

16.2.3. Технологія виготовлення сенсорів. Вибір технології виготовлення напівпровідникових сенсорів газу зумовлюється специфікою і областю їх застосування, експлуатаційними параметрами і вартістю продукції.

Для побудови і виготовлення напівпровідникових сенсорів газу найчастіше застосовуються чотири технології:

- керамічна технологія (технологія спікання);
- тонкоплівкова технологія;
- товстоплівкова технологія;
- низькотемпературна керамічна технологія (Low Temperature Cofiring Ceramics (LTCC) – технологія).

В табл. 16.4 наведено порівняльний опис методів виготовлення сенсорів газу на основі SnO₂].

Керамічні сенсори. Ця технологія є найбільш ранньою з вищезгаданих. На сьогодні ці сенсори є найбільш

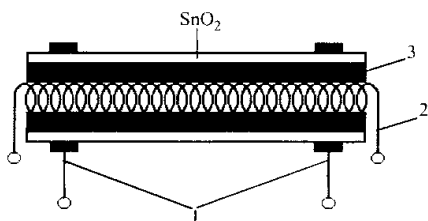


Рис. 16.12. Конструкція керамічного сенсора: 1 – електроди; 2 – нагрівач; 3 – керамічна трубка.

розповсюдженими у світі. Типова конструкція керамічного сенсора газу показана на рис. 16.12.

Таблиця 16.4.
Технології виготовлення, домішки та діапазон робочих температур сенсорів газу на основі SnO₂

| Температура, °C | |
|----------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 100 | 200 300 400 500 600 700 |
| H ₂ : | керамічний SnO ₂ [SO ₂] ↔ товстоплівковий SnO ₂ ↔ ↔ товстоплівковий Sb _{0.95} Sn _{0.02} O ₂ кристалічний SnO ₂ ↔ |
| SnO ₂ [Pd] | |
| SnO ₂ [Sb ₂ O ₃] | ←—————→ |
| Тонкоплівковий | ←—————→ керамічний SnO ₂ |
| NH ₃ : | |
| SnO ₂ [Pd, Bi, AlSiO ₃] | |
| SO ₂ : | ↔ керамічний SnO ₂ [SO ₂] |
| CO : | ↔ товстоплівковий SnO ₂ ↔ товстоплівковий SnO ₂ +ThO ₂ +SiO ₂ |
| SnO ₂ [Pd, ThO ₂] | |
| | ↔ керамічний SnO ₂ [Pt, Sb] |
| | ←—————→ товстоплівковий SnO ₂ [Pd, Cu] |
| SnO ₂ [PdO, MgO, ThO ₂] | |
| | ←—————→ товстоплівковий SnO ₂ [Pt] |
| CH ₃ COOH : | |
| SnO ₂ [Pd, Bi, AlSiO ₃] | |
| C ₂ H ₅ OH : | ↔ товстоплівковий SnO ₂ |
| CH ₄ : | товстоплівковий SnO ₂ ↔ |
| Складові вуглеводів: | |
| керамічний SnO ₂ | ←—————→ [SO ₂ , початкова обробка] |
| AsH ₃ : | шар SnO ₂ ↔ |
| H ₂ S : | керамічний SnO ₂ [SO ₂] |
| NO ₂ : | ←—————→ тонкоплівковий SnO ₂ |

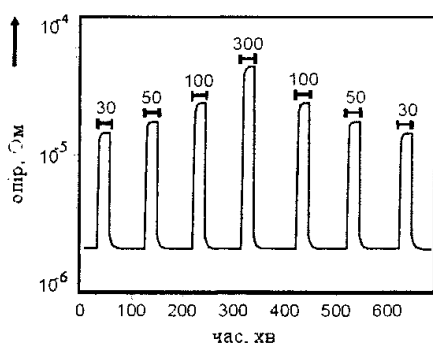


Рис. 16.13. Зміни опору в залежності від часу для керамічного сенсора у синтезованому повітрі у присутності CO різної концентрації при постійній температурі.

Несучим елементом сенсора є керамічна трубка з внутрішнім діаметром порядку 1 мм, яка виготовляється з кераміки Al₂O₃ з високою теплопровідністю. На цю трубку наносяться електроди з золота чи платини. Електроди можна наносити шляхом випалювання відповідної металічної пасти. Далі на приготовану таким чином трубку наноситься шар керамічної пасти (SnC). Після її нанесення на підкладку проводиться процес спікання при температурі 970-1270 К в атмосфері повітря. Всередині трубки знаходиться платинова нагрівна спіраль, завдяки чому сенсор під час роботи нагрівається до відповідної температури 470-620 К. До електродів прикріплюються виводи, які своїми кінцями приєднуються до відповідних кінцівок корпусу.

На рис. 16.13. показана зміна провідності в часі для типового керамічного сенсора у синтезованому повітрі у присутності CO різної концентрації при постійній температурі.

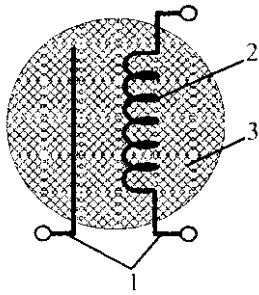


Рис. 16.14. Конструкція керамічного сенсора у вигляді таблетки: 1 – електроди; 2 – нагрівач; 3 – таблетка SnO_2 .

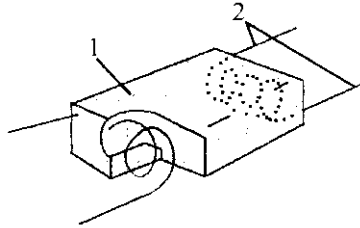


Рис. 16.15. Сенсор на основі SnO_2 : 1- чутливий матеріал; 2- дріт (сплав Ir-Pt), який виконує роль електрода і нагрівача.

Інші методи виготовлення керамічних сенсорів представлено на рис. 16.14 і 16.15.

В сенсорі, показаному на рис. 16.50, дріт із сплаву одночасно виконує роль електрода і нагрівача. Цей тип виконання, однак, застосовується значно рідше, тому що сенсор внаслідок своєї специфічної будови вимагає великої кількості активного матеріалу (невигідне відношення поверхні до об'єму), що призводить до неефективного використання, а в результаті до зменшення його чутливості.

16.2.4. Тонкоплівкові сенсори. Оскільки процеси зміни провідності в тонкоплівкових структурах відбуваються на межі між газовою фазою і твердим тілом чутливого елемента, особлива увага приділяється якості поверхні окислу. Внаслідок вигідного відношення поверхні до об'єму тонкі плівки окислів металів найбільше придатні для виготовлення сенсорів.

При створенні тонкоплівкових сенсорів повною мірою можуть бути використані методи і прийоми технології мікроелектроніки: вакуумне напилювання, фотолітографія, осадження плівок окислів металів різноманітними методами. Цей тип сенсорів також найбільш перспективний із погляду підвищення продуктивності при виготовленні чутливого елемента, зменшення його розмірів і споживаної потужності.

Для створення напівпровідникових сенсорів газів використовуються різноманітні матеріали (SnO_2 , Zn , V_2O_5 , In_2O_3 , Co , Mg тощо), проте основним матеріалом вважається SnO_2 . Плівки SnO_2 мають високі хімічну стабільність, механічну міцність, термостійкість, високу адгезію до скла й інших підкладок. Зміною умов осадження, підбором плівкоутворюючих сполук і легуванням можна одержувати плівки на основі SnO_2 з питомим опором від $1,5 \cdot 10^{-3}$ до $2,5 \cdot 10^9$ Ом·см. Шари SnO_2 , вирощені різними методами, мають великий розкид електричних параметрів, що пов'язано з типом структури (аморфна або полікристалічна) розміром окремих кристалітів плівки. Широкий діапазон розмірів одержуваних кристалів (від одиниць до декількох сотень нанометрів) впливає на електричні властивості SnO_2 , зокрема на рухливість носіїв заряду, що може змінюватися від декількох одиниць до $100 \text{ cm}^3/(\text{V} \cdot \text{c})$.

Як підкладка найчастіше використовується Al_2O_3 , SiO_2 або ферит. Технологічний процес починається з дуже важливої операції, а саме, очистки підкладки, від якої залежить адгезія шарів сенсора.

Основними методами нанесення чутливих покриттів для напівпровідникових тонкоплівкових сенсорів є термічне випаровування, вакуумне напилювання, пульверизація та хімічне осадження з газової фази.

Метод термічного випаровування металу у вакуумі з наступним його окисленням на повітрі є одним із перших

методів одержання плівок окислів металів. Тонкі металоокисні плівки наносять термічним випаровуванням металу в повітрі при тиску приблизно $1,32 \cdot 10^{-4}$ Па на діелектричні підкладки. Потім отриману плівку металу, наприклад, цинку, окисляють при температурі 770 К. Для стабілізації її параметрів проводять подальшу термообробку при температурі 820...970 К. Товщина осадженої плівки – 0,1...1,0 мкм, опір – порядку декількох мегом. Найбільша зміна опору для такої плівки Zn в присутності парів C_2H_5OH спостерігається при температурі 670-770 К. Оптимальним режимом отримання плівок SnO_2 товщиною порядку 1 мкм є: температура окислювання 720...740 К, час окислювання 2,5 години, відносна вологість повітря 60-70% . Для зменшення сталої часу сенсора проводиться активація отриманої тонкої плівки окислу металу. Зазвичай використовують вакуумне напилювання платини з нагрітої платинової нитки. Товщина осаджених плівок платини – від 1/10 до 10 моношарів, один моношар відповідає середній поверхневій густоті атомів $5 \cdot 10^{14} \text{ см}^{-2}$.

Внаслідок ряду суттєвих недоліків цей метод нанесення плівок окислів металів широко не застосовується. Більш відомі його модифікації, зокрема, реактивне термічне випаровування, яке значно покращує електрофізичні параметри плівок на основі SnO_2 .

Відомий також метод, що дозволяє одержувати ультрадрібні частинки осадженої речовини: так званий метод випаровування в газі. У атмосфері реактивного газу при тиску $13,3...13,3 \cdot 10^3$ Па матеріал нагрівається до випаровування, атоми речовини, що випаровується, зустрічаються з залишковими атомами газу, в результаті чого утворюються ультрадрібні частинки матеріалу, кожна з яких являє собою кристал із середнім розміром 0,001...0,100 мкм. Отримана плівка має надзвичайно розвинену поверхню, тому подальша термообробка не проводиться.

Плівки SnO_2 можна одержувати катодним розпиленням металевого олова. Цей спосіб передбачає створення плазми в середовищі інертного газу, наприклад, аргону, при тиску у вакуумній камері від 50 до 10^{-2} Па. До катода кріпиться мішень, до анода – підкладка.

Відомі два варіанти катодного розпилення, що відрізняються природою прикладеної напруги – діодне і височастотне (ВЧ).

При діодному розпиленні між анодом і катодом прикладена постійна напруга; для нанесення шару SnO_2 необхідно проводити процес при залишковому тиску кисню. Проте цей метод не одержав широкого поширення, оскільки швидкість росту плівок SnO_2 невелика (приблизно 1,7 нм/хв) і можна використовувати тільки мішені з провідних матеріалів.

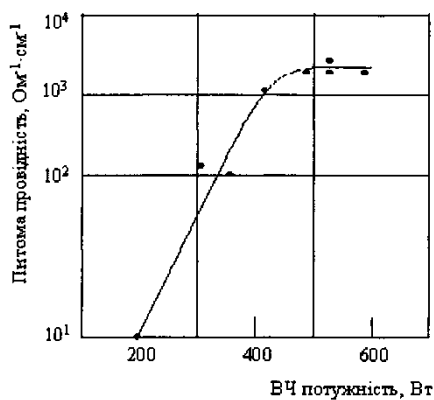


Рис. 16.16. Залежність питомої провідності від потужності ВЧ розряду (товщина плівки – 700 нм, концентрація кисню в плазмі – 1,2%).

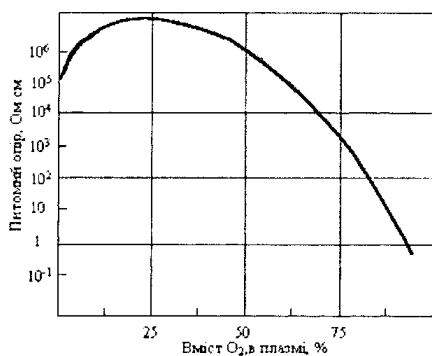


Рис. 16.17. Залежність питомого опору плівок при $T = 298\text{ K}$ від вмісту кисню в плазмі під час реактивного напилювання.

Більш висока швидкість нанесення шарів забезпечується при ВЧ напилюванні. У цьому випадку між анодом і катодом прикладається змінна ВЧ напруга. ВЧ напилювання дозволяє використовувати непровідні мішені, наприклад, спресовані порошки окису олова, окису індію або їхньої суміші. Температуру нанесення плівок можна змінювати в інтервалі 570...870 К. Швидкість напилювання, що залежить, в основному, від прикладеної напруги і парціального тиску кисню й аргону, може досягати 800 нм/хв. Провідність плівок SnO₂ змінюється в діапазоні 10⁻⁹ ... 10² Ом⁻¹·см⁻¹ (рис. 16.16). Цей метод дозволяє осаджувати як чисті, так і леговані індієм, кадмієм і сурмою плівки SnO₂.

Модифікацією ВЧ напилювання є реактивне ВЧ розпилення, що здійснюється в промисловій триелектродній системі. Процес проводиться в плазмі аргону високої чистоти і кисню. Перед напилюванням створюється вакуум 6,65·10⁻⁵ Па; під час напилювання вакуум підтримується на рівні 3,3 Па. Мішень являє собою спресований диск із порошку SnO₂ високої чистоти (99,99%). Відстань між підкладкою і мішенню – 7 см, температура підкладки 370...720 К. Для стабілізації параметрів окислу проводиться його термообробка на повітрі протягом 2 годин при температурі 620...720 К.

Середній розмір зерна для плівок товщиною 0,6 мкм, отриманих реактивним ВЧ розпиленням, дорівнює 0,04 мкм, для плівок товщиною 0,04 мкм – 0,020...0,025 мкм із коливаннями від 0,008 до 0,025 мкм. Дифракційний аналіз показав, що плівки, незалежно від типу підкладки, являють собою неперервні полікристалічні утворення без якої-небудь переважаючої орієнтації.

Питомий опір плівок залежить від вмісту кисню в плазмі під час осадження (рис. 16.17). Низька величина питомого опору плівок, осаджених у чистому аргоні, зумовлена високою концентрацією кисневих вакансій у ґратці. Максимум кривої відповідає повному заповненню цих вакансій. Відомо, що провідність полікристалічних плівок залежить від величини потенціального бар'єра між сусідніми зернами, що, у свою чергу, викликано різним складом зерен, що контактують. Висока провідність плівок, осаджених у чистому кисні пов'язана зі зменшенням густини станів у забороненій зоні. Більшу однорідність за складом мають плівки, отримані в чистому кисні. Крім того, ці плівки характеризувалися найбільшою стабільністю параметрів.

Одним із ефективних методів нанесення плівок окислів металів є магнетронне розпилення. При негативній напрузі на катоді (700...900 В) поблизу його поверхні створюється розряд плазми (рис. 16.18). Під впливом магнітного поля вільні електрони здійснюють коливальні рухи по спіральній траєкторії. В результаті

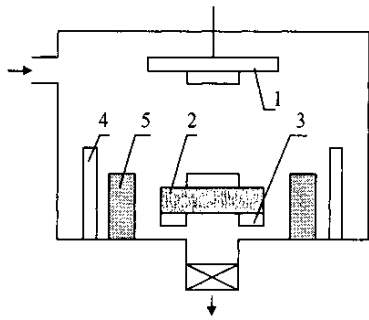


Рис. 16.18. Схема реакційно-розрядної камери магнетронного напилювання: 1 – підкладкотримач; 2 – катод; 3 – постійні магніти; 4 – захисний циліндр; 5 – анод.

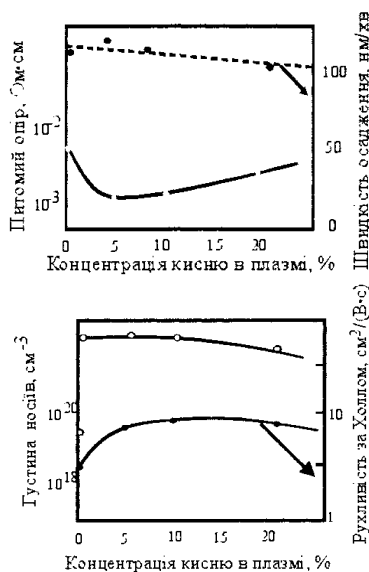


Рис. 16.19. Залежність швидкості осадження й електричних параметрів плівок SnO_2 від концентрації O_2 у плазмі; вміст сурми в мішені 3.5 ат %, температура підкладки = 670K; $P=0.4$ Па; $U_{вч}=1,75$ кВ.

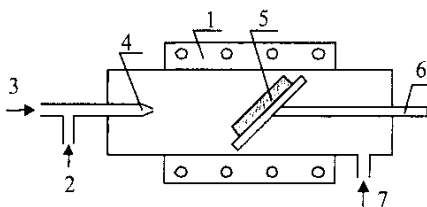


Рис. 16.20. Схема установки для пульверизації аерозолів: 1 – півка; 2 – спиртовий розчин хлоридів; 3 – газ-носіє; 4 – сопло; 5 – підкладка; 6 – термопара; 7 – відведення продуктів реакції.

захоплення електронів плазмою збільшується можливість їхнього зіткнення з атомами залишкових газів, а інтенсивність розряду зростає на два порядки в порівнянні зі звичайними катодними системами; швидкість осадження на підкладці досягає 210 нм/хв.

Існують два різновиди цього методу: магнетронне розпилення на змінному і постійному струмі.

Плівки SnO_2 , нанесені в чистому аргоні при температурі 670 К і тиску $4 \cdot 10^{-1}$ Па мають полікристалічну структуру типу рутилу з площиною $\langle 111 \rangle$, паралельну поверхні підкладки. З ростом температури підкладки і концентрації кисню до 10% збільшується інтенсивність рефлексу $\langle 101 \rangle$, який є характерним для плівок SnO_2 . Залежність їхніх електричних властивостей від парціального тиску кисню в камері (рис. 16.19) пов'язана зі зміною числа вакансій і надлишком металевих іонів внаслідок нестехіометричного складу плівки. Рухливість носіїв заряду різко зростає зі збільшенням концентрації кисню внаслідок зменшення кількості фази Sn і збільшення ступеня кристалічності структури плівки. Максимальна рухливість носіїв заряду спостерігається для плівок із явно вираженою орієнтацією $\langle 101 \rangle$.

Всі фізичні методи нанесення окисних плівок мають недоліки, зумовлені наявністю плазми в робочому обсязі камери та високою вартістю обладнання, що використовується. Це викликало бурхливий розвиток аерозольних методів отримання плівок на основі окису олова, основаних на розкладанні при високій температурі (720-970 К) розчину хлориду або металоорганічної сполуки (МОЗ) олова на поверхні підкладки (рис. 16.20). Як вхідні МОЗ використовуються тетраалкоксипоксидні олова. Пульверизація розчин відбувається під впливом інертного газу, так що розчин переноситься до поверхні підкладки у вигляді дрібних крапель. При цьому сопло пульверизатора переміщається вздовж поверхні підкладки, нагрітої до високої температури.

Змінюючи параметри процесу (швидкість подачі розчину, концентрацію сполук олова і легуючих добавок, температуру підкладки), можна одержувати шари з різноманітними електрофізичними характеристиками. В осадженні беруть участь краплі рідини розчину, складові частини якого можуть перейти в газоподібний стан до того, як досягнуть поверхні підкладки, або ж прореагувати на ній після розбризкування, що суттєво впливає на властивості одержуваних плівок. Швидкість осадження плівок окису олова методом пульверизації досить висока й складає в середньому 100 нм/хв. Цей метод дозволяє відносно просто одержувати плівки з композиційних і легованих окислів, наприклад, SnO_2 : Sb, а також складні шаруваті структури.

Недоліком методу є порівняно мала продуктивність і

велика витрата реагентів у порівнянні з методами вакуумного напилювання та хімічного осадження з газової фази. Висока температура процесу створює умови для виникнення великих внутрішніх напруг у плівках. Питомий опір плівок окису олова для відносно низької температури осадження (менше 570 К) великий внаслідок аморфної структури плівки і пов'язаної з цим малої концентрації носіїв заряду та їх низької рухливості. При підвищенні температури збільшення розмірів кристалітів призводить до росту провідності плівок. Проте при дуже високій температурі підкладки (вище 720 К) питомий опір зазвичай зростає, що пояснюється сильним забрудненням лужними металами з матеріалу підкладки або зміною стехіометрії складу шару. Основним недоліком аерозольних методів нанесення плівок на основі окису олова є великий розкид параметрів одержуваних шарів.

Одним з перспективних методів одержання плівок на основі SnO_2 , який був розроблений спеціально для виготовлення напівпровідникових тонкоплівкових сенсорів газів на керамічній основі, є так званий метод занурення (Dipping). Підкладку занурюють у розчин тетраалкоксипохідного олова (зазвичай тетрабутоксидолова) у розчиннику, що забезпечує необхідні плівкоутворюючі властивості і містить необхідні легуючі добавки; після підсушування сформовану тонку плівку відпалюють при температурі 750 К. При цьому відбувається розкладання МОЗ олова до SnO_2 . Різновидністю даного способу є метод, оснований на використанні дуже в'язких суспензій, тобто паст, які наносять на підкладку тим або іншим способом; термообробку проводять аналогічним способом.

Широке використання для одержання шарів окислів металів знаходять хімічні методи осадження з газової фази. При цьому реагент у вигляді парогазової суміші подається в реакційну камеру (рис. 16.21). У процесі хімічного осадження з газової фази ламінарний потік інертного газу-носія цілком покриває підкладку й утворює над нею граничний шар, у якому газ практично не переміщується.

Швидкість осадження в даному випадку залежить від швидкості дифузії реагентів через граничний шар. На поверхні зразка відбуваються хімічні реакції з утворенням окислу металу. Як реагенти використовуються МОЗ олова. Окислювачами служать кисень, вода або перекис водню. Сполуки олова легко переводяться в газоподібний стан при порівняно низьких температурах у випадку використання МОЗ або при більш високій температурі коли використовуються хлориди олова. Як МОЗ застосовуються також дибутилдіацетатолово. Причому легування сурмою здійснюється спільним піролізом із пентахлоридом сурми, тетраметилолова та легуючих добавок триметилсурми, фосфіну або арсину.

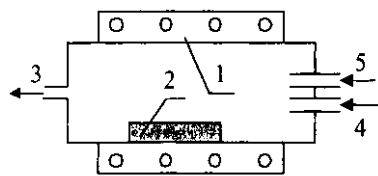


Рис. 16.21. Схема методу хімічного осадження плівок SnO_2 із газової фази: 1 – півка; 2 – підкладка; 3 – відведення продуктів реакції; 4 – $(\text{SnCl}_4, \text{InCl}_3, \text{MO}_3) + \text{N}_2$; 5 – $(\text{O}_2, \text{H}_2\text{O} \dots) + \text{N}_2$.

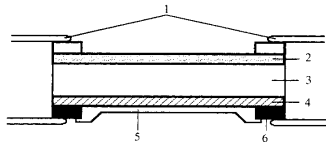


Рис. 16.22. Приклад конструкції тонкоплівкового сенсора газу: 1 – золоті виводи; 2 – SnO_2 ; 3 – підкладка; 4 – нагрівач (RuO_2); 5 – ізоляційний шар; 6 – електрод нагрівача (Pd-Ag).

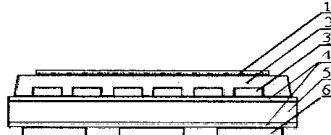


Рис. 16.23. Конструкція тонкоплівкового сенсора газу на підкладці з кремнію:

- 1- каталізатор;
- 2- активний шар SnO_2 ;
- 3- електроди;
- 4- ізолюючі шари SiO_2 ;
- 5 – підкладка Si;
- 6- платиновий нагрівач.

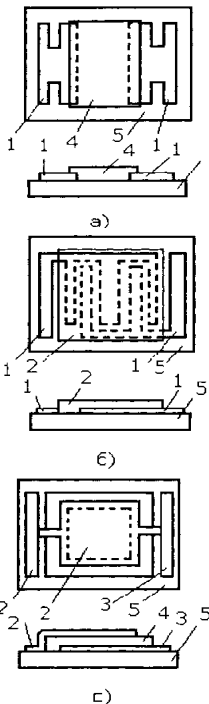
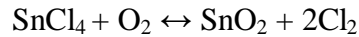
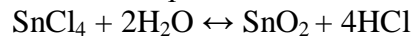


Рис. 16.24. Основні структури товстоплівкових сенсорів: а – планарна; б – «тандем»; в – «сандвіч»: 1 - електрод; 2 - верхній електрод; 3 - нижній електрод; 4 - чутливий матеріал; 5-підкладка (Al_2O_3).

Розглянуті МОЗ олова – технологічно зручні рідини з відносно високим тиском пари при кімнатній температурі, тому перенесення їхніх парів у реакційну камеру достатньо простий. Деякі результуючі рівняння реакції одержання двоокису мають вигляд:

Газ Газ Тверде тіло Газ



Зміною умов осадження, підбором вхідних сполук і легуванням можна отримати опори плівок SnO_2 у діапазоні від $1,5 \cdot 10^{-3}$ до $2,5 \cdot 10^9$ Ом·см. Швидкість нанесення плівки осадження із газової фази зазвичай вища, ніж при вакуумному розпиленні, і складає в середньому 20-100 нм/хв і навіть порядку 1000 нм/хв при температурі 870 К.

На кінцевому етапі на активний шар наносяться вимірювальні електроди і електроди для підведення живлення до нагрівача (рис. 16.22).

Ще одним прикладом конструкції напівпровідникового сенсора газу на основі тонкоплівкової технології є сенсор на кремнієвій підкладці (рис. 16.23).

Обидві сторони Si-підкладки кремнію піддаються типовій технологічній обробці, характерній для моносистем (фотолітографія, травлення тощо). Активний шар SnO_2 напильють на платинові електроди у визначених місцях з використанням масок відповідного типу.

16.2.5. Товстоплівкові сенсори. Товстоплівковий газочутливий сенсор складається з діелектричної підкладки, плівки чутливого матеріалу і електродів. За способом кріплення електродів їх можна розділити на три типи, які вибираються для отримання потрібних характеристик (рис. 16.24).

На рис. 16.25 схематично показані основні процеси виготовлення товстоплівкових сенсорів.

Верхній електрод у сенсорів з «сандвіч» – структурою відпалюється одночасно з матеріалом сенсора, оскільки плівка чутливого шару після відпалу стає пористою і виникає небезпека короткого замикання при нанесенні провідникової пасти. Відпалювання здійснюється при температурі 970-1120 К. Характеристики сенсорів значною мірою залежать від температури відпалювання, яка визначається матеріалом сенсора. Температуру відпалювання для електродів в сенсорах планарної конструкції і типу «тандем» та нижнього електрода сенсора типу «сандвіч» бажано вибрати на 320 К вище температури відпалу матеріалу сенсора.

Як матеріал електродів часто використовується система Ag-Pd; можуть також використовуватись системи з платиною, золотом, сріблом, міддю, нікелем і оксидом

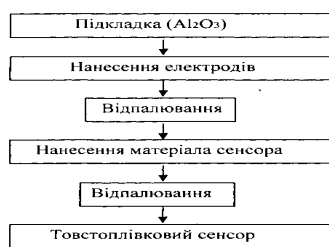


Рис. 16.25. Схеми основних процесів виготовлення товстоплівкових сенсорів планарної конструкції (а) і «сендвіч» – типу (б).

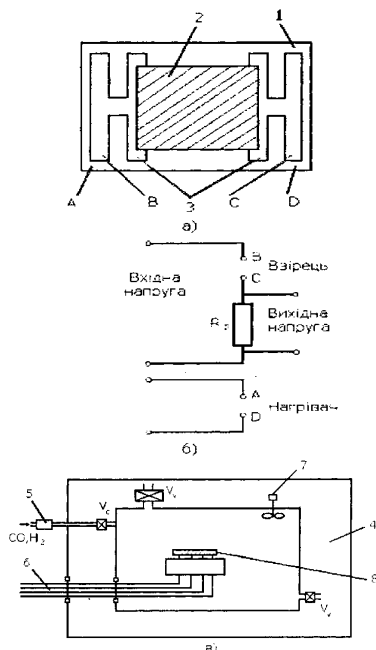


Рис. 16.26. Схематичне зображення товстоплівкового сенсора (а), експериментальної установки для вимірювання концентрації газу (б), і електричної вимірювальної схеми (в): 1- підкладка; 2 – чутливий шар; 3-електроди; 4- термогігростат; 5- електричні з'єднання; 6- вентилятор; сенсор; Vc, Vv – клапани.

рутенію. Всі ці матеріали використовуються у вигляді паст. Багато напівпровідників, що використовуються в сенсорах, мають вищу температуру спікання, ніж у традиційній товстоплівковій технології. В процесі відпалювання напівпровідники формуються не у вигляді суцільної структури з комірками, а розсіюються в отриманій склоподібній фазі. Тому при утворенні плівки сильно зростає її питомий опір та з'являється нелінійна залежність струму від напруги за рахунок тунелювання на границі зерен напівпровідника. Цей недолік частково усувається при легуванні плівок металами.

Товсті плівки використовуються при виготовленні газових сенсорів як зв'язувальний матеріал, який запобігає утворенню скла, щоб не зменшувалась питома провідність плівки і відповідно чутливість сенсора (рис. 16.25). Для збільшення механічної міцності в пасту вводять гідрофільний кремнезем. Сенсори, виготовлені за товстоплівковою технологією, використовуються для визначення таких газів, як CH_4 , CO , H_2 . Матеріалом сенсора є система силікагель-двоокис олова з добавками родію або $\text{PdCl}_2\text{-MgO}$.

Типовий процес виготовлення товстоплівкового сенсора включає такі операції: суміш порошоків (у вагових процентах) 93 – SnO_2 , 1 – PdCl_2 , 1 – $\text{Mg}(\text{NO}_3)_2$, 5 – ThO_2 відпалюють на повітрі при 1270 К протягом 1 год, після чого подрібнюють, змішують з силікагелем, наносять на підкладку і відпалюють на повітрі при 870 К. Товщина утвореної плівки – 20 мкм. Сенсори, що містять гідрофільний кремнезем, чутливі до вмісту вологи в газі, що аналізується. В зв'язку з цим частіше використовується гідрофобний кремнезем, зокрема, в сенсорах CO_2 .

Гідрофобний кремнезем приготують реакцією взаємодії гідрофільного кремнезему і триметилхлоросилану. Після нанесення пасту на підкладку і сушіння протягом дня при температурі 350 К здійснюється відпалювання при 670 К протягом 1 год. Товщина плівки – 20 мкм. Відстань між електродами (Pt-Au) – 0,5 мм. На зворотній стороні підкладки формується товстоплівковий нагрівач. До взірця приєднується вимірювальний опір R_p , величина якого складає 1/10 від опору взірця в відсутності детектованого газу (рис. 16.26). Залежності вихідної напруги від концентрації CO і H_2 показані на рис. 16.27. Характеристики сенсора практично не залежать від вологості навколишнього середовища.

В процесі розробки та дослідження товстоплівкового сенсора пропану на основі SnO_2 встановлено, що плівки SnO_2 , леговані перехідними металами (Nb, V, Ti) і молібденом, не підпорядковуються рівнянню Фрейндліха. Добавки PdCl_2 і $\text{Mg}(\text{NO}_3)_2$ складали по 1 ваг.%, добавки ніобію – до 20 ваг.%, робоча температура сенсора – 550 К.



Рис. 16.27. Залежність вихідної напруги сенсора від концентрації CO (криві 1, 4) і H₂ (криві 2, 3) для нелегованих (криві 2, 4) і легованих ThO₂ (криві 1, 3) взірців при температурі 470 К; вміст кремнезему 15,4 ваг.%, величина R_p=411,5 кОм.

Електричний опір сенсорів зменшується зі збільшенням концентрації газу. Для зразків, які містять більше 20 ваг.% і менше 7 ваг.% ніобію, залежність опору від концентрації газу є пряма лінія, що відповідає рівнянню Фрейндрліха. При концентраціях ніобію 9-17 ваг.% на характеристиці спостерігається злам. Напруга детектування на вимірювальному резисторі збільшується з 13 до 25 В при концентраціях газу 0,6%. Точка згину зсувається в сторону менших концентрацій з збільшенням прикладеної напруги і зовсім зникає при напругах більше 40 В. Введення добавок титану в кількості 5 ваг.% покращує точність визначення концентрації газу з ±0,1 до ±0,03%.

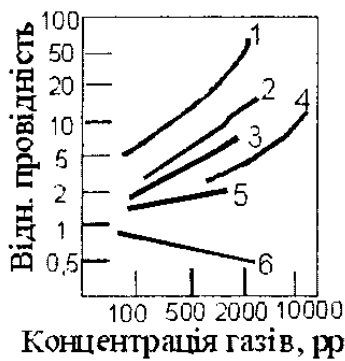
При використанні як основного матеріалу для сенсорів складних окислів (наприклад, типу перовскіту LaNiCb) технологія виготовлення чутливого шару дещо інша. Вихідними матеріалами, наприклад, є ацетати металів чистотою 99,9%. їх водяний розчин випаровується у вакуумі (2,6·10⁻³ Па); отриманий дрібнодисперсний порошок поміщається в кварцовий човник і піддається піролізу в діапазоні температур 420-570 К. Отримана керамічна речовина подрібнюється і використовується як вихідний матеріал для виготовлення сенсора.

Плівки окислів типу перовскіту виявились чутливими до вмісту в повітрі етанолу. Пасту з вихідної речовини, виготовлену за описаною вище методикою, наносили на діелектричну підкладку і відпалювали при 1270 К протягом 10 хв. Товщина плівки – 15-20 мкм. Опір сенсорів вимірювали в інтервалі температур 370-670 К. Концентрацію спирту вимірювали від 30 до 1000 ppm. Калібрування проводили за допомогою методу газової хроматографії. Було виявлено, що в плівках LaMeO₃ основну роль у зміні опору сенсора відіграє метал (Fe, Ni, Co). Плівки LaNiO₃ показали найбільшу зміну опору, плівки LaCoO₃ та LaNiO₃ – найбільшу швидкодію. Максимальний сигнал спостерігали для плівок LaNiO₃ (відносна зміна опору становила 23%) при температурі 540 К; швидкодія при цьому складала 5 с.

Використовуючи товстоплівкову технологію, можна отримувати багатошарові структури з товстих плівок на діелектричній підкладці Al₂O₃. Це призводить до підвищення ступені мініатюризації, збільшення термічного ККД і покращення надійності. Крім цього, товстоплівкова технологія дозволяє сумістити процеси виготовлення сенсора та електричної схеми (підсилювач, перетворювач і тощо).

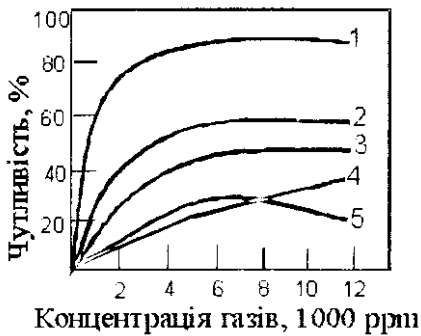
Товстоплівкові сенсори газу є одними з найдешевших. Параметри деяких товстоплівкових сенсорів наведені в табл. 16.5.

Чутливість товстоплівкових сенсорів можна підвищити, збільшуючи пористість структури активного



Концентрація газів, ppm

Рис. 16.28. Залежність провідності окислів (WO_2 -криві 1, 3; SnO_2 – криві 2, 4, 5; $LaNiO_3$ – крива 6) від концентрації етанолу (криві 1, 2, 6), окислу вуглецю (криві 3, 5) і метану (крива 4).



Концентрація газів, 1000 ppm

Рис. 16.29. Чутливість γ - FeO_3 , легованої паладієм, до бутану (крива 1), метану (крива 2), пропану (крива 3), водню (крива 4) і окислу вуглецю (крива 5).

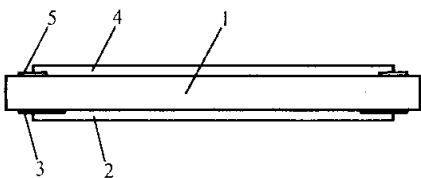


Рис. 16.30. Конструкція товстоплівкового сенсора газу на основі SnO_2 : 1 – підкладка (Al_2O_3); 2 – нагрівач; 3 – електрод нагрівача; 4 – чутливий елемент (SnO_2); 5 – електрод чутливого елемента (Au).

шару. Цьому сприяє використання високої робочої температури сенсора, що приводить до збільшення його швидкодії, оскільки прискорюються процеси поверхневої дифузії і каталітична реакція між включеними в склад шару хімічними елементами. Вибір сенсора для різних газів ілюструється на рис. 16.28.

Товстоплівковий сенсор на підкладці розміром 2,5x5,0 см використовується для виявлення факту втрати газу, його типу і концентрації. До складу сенсора також входять товстоплівковий нагрівач на основі платини і схема в інтегральному виконанні, яка визначає концентрації CO і C_2H_5OH у середовищі, що досліджується.

Таблиця 16. 5.

Параметри товстоплівкових сенсорів концентрації газів

| Основні складові компоненти чутливого елемента | Гази, що вимірюються | Чутливість $\Delta R/R$ (%), при концентрації і 1000 ppm | Робоча температура, К |
|------------------------------------------------|----------------------|----------------------------------------------------------|-----------------------|
| $SrFe^{3+}_{0,7}Fe^{4+}_{0,3}O_{2,65}$ | C_2H_5OH | 0,1 | 743 |
| SnO_2+Pd | $CH_4/CO/C_2H_5OH$ | 2 / 1,1 / 8, 3 | 673 |
| WO_3+Pd | $CH_4/CO/C_2H_5OH$ | 1.5 / 9 / 25 | |
| $ZnO+Pd$ | $CH_4/CO/C_2H_5OH$ | 1 / 2, 2 / 17 | |
| SnO_2ThO_2 +гідрофобний кремнезем | CO | 90 (500 ppm) | 473 |
| SnO_2+Pt, Pd | $C_2H_5OH/CO/H_2$ | -60/-2/+45 | 523 |
| SnO_2+Pt, Pd | $C_2H_5OH/CO/H_2$ | -90/-5/-50 | 673 |

В системах, які регулюють процеси горіння, можна використовувати напівпровідниковий сенсор на основі суміші Al_2O_3 і V_2O_5 . Такий склад чутливий до зміни концентрації окисів азоту (1-1000 ppm) і практично не реагує на зміну концентрації CO і CO_2 . Як сенсор водню і вуглеводню використовується окис цинку і легована паладієм γ - Fe_2O_3 . Залежність чутливості подібного сенсора від концентрації різних газів наведена на рис. 16.29.

Приклад типової конструкції товстоплівкового сенсора газу представлено на рис. 16.30.

Зустрічаються й інші конструкції, які можуть відрізнятися взаємним розміщенням окремих компонентів сенсора, причому сам процес виготовлення залишається незмінним.

Прикладом нестандартного рішення є "одностороння" схема (рис. 16.31). Як активний шар, так і інші компоненти розміщуються на одному боці підкладки. Таке технічне рішення є певного роду компромісом між простотою схеми і хорошими експлуатаційними параметрами сенсора.

Іншою групою сенсорів є схеми з двома активними

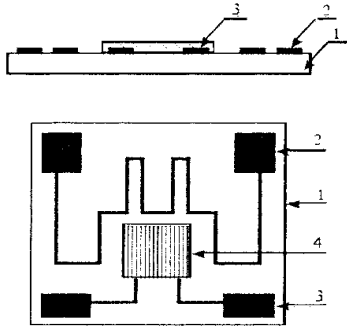


Рис. 16.31. Конструкція товстоплівкового одностороннього сенсора газу: 1- підкладка; 2- нагрівач; 3- платинові електроди; 4- активний шар.

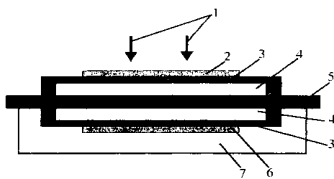


Рис. 16.32. Конструкція товстоплівкового сенсора газу з активним елементом і елементом порівняння: 1 – агресивний газ; 2 – чутливий шар (SnO_2); 3 – ізоляція; 4 – нагрівач; 5 – підкладка; 6 – елемент порівняння; 7 – корпус.

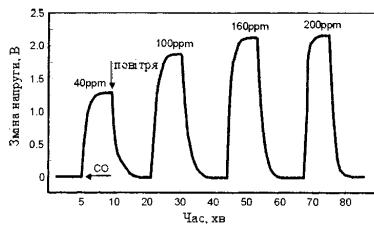


Рис. 16.33. Часовий відгук сенсора, працюючого у схемі моста, на окис вуглецю малої концентрації.

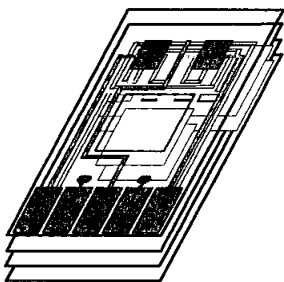


Рис. 16.34. Сенсор газу (SnO_2), отриманий за технологією LTCC.

шарами, причому одним шаром служить елемент порівняння, а другим – вимірювальний елемент. Схема такого сенсора показана на рис. 16.32.

Два шари працюють у схемі зрівноваженого моста. В момент появи агресивного газу міст виходить із стану рівноваги, що викликає реестрацію навіть малих концентрацій газу. Приклад часової відповіді сенсора цього типу на присутність CO малої концентрації показано на рис. 16.33.

Керамічні сенсори, виготовлені за LTCC-технологією.

Головною перевагою цієї технології є можливість формування багатшарових тривимірних структур у вигляді однієї схеми MCM (Multichip Module). В зв'язку з цим зростає щільність пакування як елементів, так і з'єднань між ними. Основними етапами технологічного процесу LTCC є:

- виконання перехідних отворів у послідовних плівках мікросхеми;
- ламінування плівок під високим тиском;
- випалення.

В результаті одержують багатофункціональну мікросхему, яка дуже стійка до зовнішніх впливів і має високу механічну міцність.

Застосування даної технології для газових сенсорів дозволяє досягти як мініатюризації самої схеми сенсора, так і розмістити в одному корпусі додаткові схеми, призначені для обробки даних, які поступають з чутливого елемента. Шляхом застосування "захованих" шарів створюється також можливість задавання відповідного розподілу температури на активній поверхні сенсора, що є дуже важливим. Легкість створення інтегрованих сенсорних матриць є додатковою перевагою цієї технології.

На рис. 16.34 наведено приклад конструкції сенсора газу, отриманого за технологією LTCC.

Він складається з чотирьох шарів. Внутрішні шари містять нагрівальні та з'єднувальні елементи. На верхньому шарі знаходиться активний елемент (шар SnO_2), отриманий за товстоплівковою технологією.

Методи покращення чутливості і селективності напівпровідникових сенсорів газу. Однією з найважливіших вимог до газових сенсорів є чутливість і селективність, тобто здатність реагувати на присутність чи певну концентрацію одного чи кількох газів. Водночас це найскладніша проблема в технологічному плані. Мірою чутливості сенсорів є максимальне відношення їх опору у повітрі до опору у досліджуваному газу за даної температури.

Відомі чотири основних методи покращення селективності і чутливості:

- відповідний підбір робочих температур (температурного поля) сенсора;

- легування поверхні сенсора матеріалами, які мають більшу активність до визначених газів;

- застосування фільтрів, що блокують визначені компоненти газів;

- підбір складу матеріалу, особливо каталізаторів і активаторів. Селективність і чутливість сенсора залежать також від технології його отримання. Наприклад, керамічні, товстоплівкові і тонкоплівкові сенсори, виконані з тих самих матеріалів, мають різну чутливість для визначеного газу, причому вона залежить від товщини шару. Окрім технологічних факторів, використовуються також схемні рішення, які збільшують селективність вимірювання концентрації газу. Для цієї мети використовується багатофункціональна схема, яка складається з кількох сенсорів (матриці), що індивідуально реагують на різні компоненти газу, вихідні сигнали яких служать для взаємної компенсації вимірювальних кіл для кожної компоненти.

Робоча температура сенсора. Необхідною умовою роботи сенсорів адсорбційного типу є підвищена температура чутливого елемента (ЧЕ), яка забезпечує активацію процесів адсорбції-десорбції газу, селективність і чутливість до компонента, що визначається, в зв'язку з відмінністю в енергіях активації процесів адсорбції того або іншого газу. Причому, як показало дослідження взаємодії напівпровідникових сенсорів з газом, що вимірюється, (наприклад, оксиду ніобію з киснем), при дуже низькій температурі ЧЕ продукти реакції не будуть десорбуватися з поверхні перетворювача, а поверхню не можна відновити для нового циклу вимірювань. При високій температурі газ, що вимірюється, не зможе адсорбуватися на поверхні перетворювача, тобто необхідний вибір оптимальної робочої температури ЧЕ.

Напівпровідникові сенсори газу є прекрасним прикладом системного підходу до їх проектування, виготовлення та дослідження. В загальному випадку областями досліджень служать неорганічна хімія, фізика твердого тіла і електроніка у широкому розумінні (технологія, проектування, обробка даних тощо). Однак, температура є головним фактором, який визначає правильну роботу сенсора. Як з хімічної, так і з фізичної точки зору, основні робочі параметри, такі як чутливість і селективність значною мірою залежать власне від температури. З хімічної точки зору кінетика реакції на активній поверхні залежить від роду реагенту (активного газу). Можна стверджувати, що для кожної речовини існує температура максимальної чутливості сенсора (визначена як відношення опору сенсора у синтезованому повітрі до опору у присутності активного газу). Коли температура є занадто низькою, то реакція протікає повільно, а коли температура є занадто високою, реакція

окислення проходить занадто швидко і концентрація реагенту на поверхні зумовлена дифузією. У цьому випадку з точки зору сенсора ця концентрація прямує до нуля і чутливість є невеликою. В табл. 16.6 представлено приклади діапазонів змін температури для сенсора на основі SnO₂ для визначення різних газів .

Таблиця 16.6.

Приклади діапазонів змін температури для сенсора на основі SnO₂ для визначення різних газів.

| Паста | Газ | Діапазон зміни температури °С |
|--------------------------|----------------------------------|-------------------------------|
| SnO ₂ | H ₂ | 340-360 |
| SnO ₂ | CO | 280-310 |
| SnO ₂ Pd, Cu] | CO | 100-220 |
| SnO ₂ Pt] | CO | 130-220 |
| SnO ₂ | C ₂ H ₅ OH | 230-250 |
| SnO ₂ | CH ₄ | 470-520 |

Напівпровідникові сенсори газу зазвичай працюють при підвищеній температурі (навіть до 900 °С), досягнутій за допомогою нагрівача, який інтегрований в конструкцію сенсора. В залежності від технології виготовлення сенсора нагрівачем може служити надрукований чи напильний шар або нагрівальна дротяна спіраль з визначеними властивостями. У даний час найчастіше застосовуються плівкові нагрівачі з резистивної пасти на основі RuO₂ або з провідникових паст на основі Pt, PdAg тощо.

Загальний приклад нагрівача як складового елемента сенсора показаний на рис. 16.35 .

У такому випадку опір сенсора $R_{сeнс}$ є функцією розподілення температури на активній поверхні, концентрації газу (V_g) і параметрів матеріалу, наприклад, густини (ρ):

$$R_{сeнс} = f [T(x, y); V_g; r] \quad (16.1)$$

Із залежності (16.19) випливає сильна залежність опору сенсора, а значить, і його чутливості і селективності від розподілу температури на його поверхні. Тому слід вірно підібрати нагрівач (чи систему нагрівачів) і спосіб його живлення для отримання якнайбільш рівномірного розподілу температури поля, що є основною проблемою вже на етапі проектування.

Приклади конфігурацій нагрівачів, виконаних з резистивних і провідникових паст показано на рис. 16.36 і 16.37.

Як показують експерименти, у більшості випадках окремі сенсори не гарантують відповідного розподілення температурного поля (рис. 16.38 і 16.39).

Тому для знаходження задовільного розподілення температурного поля мають значення як сама конструкція нагрівальних систем, так і спосіб їх живлення. У першому

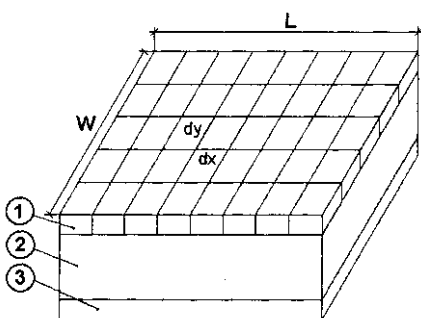


Рис. 16.35. Плівкова конструкція газового сенсора: 1- активний шар; 2- підкладка; 3- нагрівач; L – довжина, w – ширина.

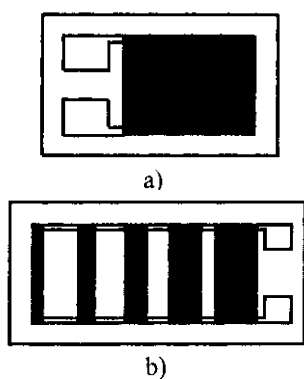


Рис. 16.36. Резистивні нагрівачі: а) один сенсор; б) матриця з п'яти сенсорів.



Рис. 16.37. Нагрівачі, виконані з провідникових паст у вигляді меандру.

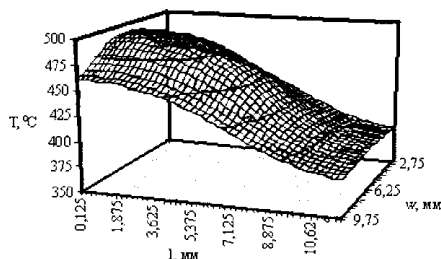


Рис. 16.38. Приклад розподілу температурного поля на поверхні окремого сенсора.

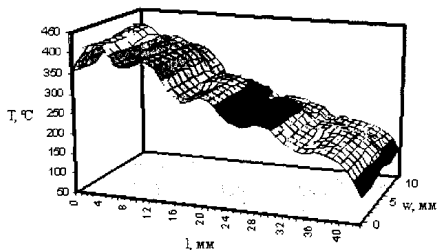


Рис. 16.39. Приклад розподілу температурного поля на поверхні матриці сенсорів (у схемі з 5 нагрівачів).

випадку одним з основних методів є застосування додаткових нагрівачів, які коректують температурне поле (рис. 16.40). Приклад розподілення температурного поля для групи нагрівачів показаний на рис. 16.41.

Як видно з наведеного прикладу, додаткові нагрівачі, розташовані на ізоляційному шарі, який покриває основний нагрівач, достатньо добре коректують розподіл температурного поля. Градієнт температурного поля на активній поверхні є значно меншим, ніж для випадку одного сенсора. Цей метод, на жаль, є досить складним для реалізації у випадку матриць сенсорів, де часто необхідно мати зміни значень температури порядку кількох сот градусів на малих поверхнях.

Іншим методом формування температурного поля є вид живлення нагрівальних елементів. Система (мозаїка) нагрівачів може отримати живлення двома методом: постійним чи імпульсним струмом. Від систем газових сенсорів вимагається максимальне плоске (а на границях різко спадаюче) температурне поле. Згідно з експериментами живлення постійною напругою не дозволяє отримати такого розподілу температурного поля. Публікації останніх років показують, що великий градієнт температури на відносно малих поверхнях можна отримати з використанням імпульсного живлення чи модуляції температури. В окремих випадках найкращим є квазістаціонарний розподіл температурного поля, коли навіть при дуже різких змінах напруги живлення поле буде рівномірним і максимально крутим на його границях.

Легування поверхні сенсора. Іншим методом хімічного впливу на властивості напівпровідникового газового сенсора є його легування, суть якого полягає у введенні чужорідного атому елемента (домішки) у кристалічну решітку металу (наприклад, олова у випадку сенсорів на основі SnO_2). Легування впливає насамперед на такі параметри сенсора як селективність, опір, стабільність і чутливість. Види домішок, що застосовуються для легування окислів металів, наведено у табл. 16.7.

Як вже згадувалося, провідність активного шару сенсора залежить від його поверхневої провідності і від приєднаної паралельно провідності внутрішньої частини зерен. Як провідність внутрішньої частини зерна, так і поверхнева провідність можуть змінюватися в залежності від потреби, причому чутливою на присутність домішок, що утворюють рівні пасток є провідність внутрішньої частини зерен. Для прикладу, введення невеликої кількості домішки в структуру викликає десятикратне пониження чутливості на етанол при незмінній чутливості на CO_2 . Це означає, що даний сенсор буде характеризуватися збільшеною селективністю на CO_2 .

Таблиця 16.7.

Порівняння деяких напівпровідникових хімічно чутливих окислів

| Основний матеріал | Домішки | Газ, що аналізується |
|---------------------|--------------------------|----------------------------------------------------------------|
| $Ba_{1-x}Sr_xTiI_3$ | $MgCr_2O_4$ | H_2O |
| CoO | -, MgO | O_2 |
| Co_3O_4 | - | CO |
| Fe_2O_3 | Pd, Pt | CO, HC, H_2 |
| фталоціамін олова | - | NO_2 |
| In_2O_3 | CeO_2 | O_3 (озон) |
| $LaCrO_3$ | SrO | O_2 |
| $MgCr_2O_4$ | TiO_2 | H_2O, H_2S |
| NiO | сульфанілінова кислота | NO_2 |
| SnO_2 | - | $O_2, CO, H_2, HC, H_2S, R-OH, NO_x,$ анестезіологічний газ |
| SnO_2 | Cr_2O_3, In | O_2, NO, NO_2 |
| SnO_2 | окисли Mg, Pd, Nb, V, Ti | HC |
| SnO_2 | Pt, Pd/Cu, Au/Pt | H_2, HC, CO, H_2S |
| SnO_2 | Pd, Pt, V | $H_2HC, CO, H_2S, R-OH$ |
| SnO_2 | Sb | H_2, CO, H_2S |
| SnO_2 | ZnO | CO |
| $SnO_2-Bi_2O_3$ | Sb_2O_3 | CN, CO |
| TiO_2 | -, Pt | O_2, CO |
| TiO_2 | Cr | NH_3 |
| TiO_2 | Ru | $(CH_3)_3N$ |
| $TiO_2 - SnO_2$ | Ta_2O_5, Sb_2O_5 | H_2O |
| ZnO | - | O_2, CO, HC |
| ZnO | Ag, Pd, Pt, Au | HC, CO, R-OH |
| ZnO | CuO | H_2, O_2 |
| ZrO_2 | MgO | $H_2O, R-OH, H_2CO,$ |
| ZrO_2 | Y_2O_3 | CN |
| WO_3 | - | O_2 |
| WO_3 | Rh, Pd, Pt, Ir | HC, R-OH HC, H_2, NH_3, H_2S |

Примітка: у колонці " Газ, що аналізується " символом HC позначено вуглеводні, R – органічні радикали, H_2O - водяна пара.

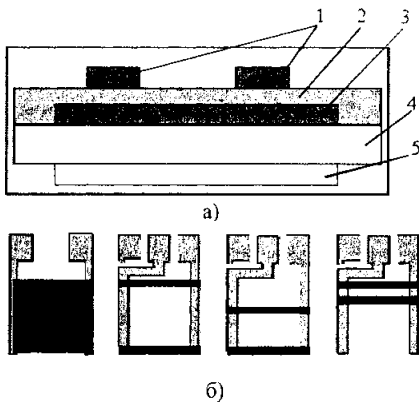


Рис. 16.40. Багатошарова схема нагрівачів: а) структура; б) конфігурація нагрівачів: 1 – додатковий нагрівач; 2 – ізоляційний шар; 3 – основний нагрівач; 4 – підкладка; 5 – чутливий шар.

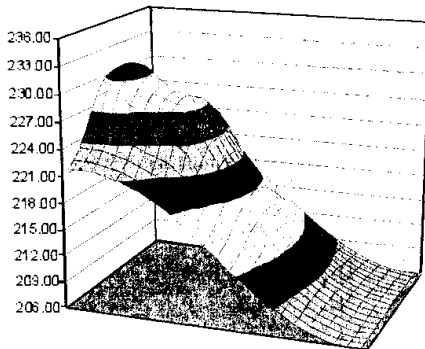


Рис. 16.41. Приклад температурного поля для групи сенсорів.

В загальному випадку, при малій товщині чутливого шару електропровідність може бути визначена методом Ван-дер-Ваальса залежністю:

$$S = \frac{p \cdot l}{\ln 2 \cdot R \cdot b}$$

де: l – довжина плівки; R – опір плівки; b – ширина активного шару.

На рис. 16.42 представлено характеристику залежності електропровідності від температури товстоплівкової структури, легованої лантаном (де σ_0 початкова електропровідність).

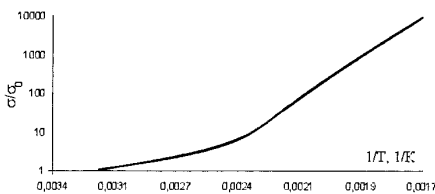


Рис. 16.42. Залежність електропровідності від температури для шару $SnO_2 + La$.

Електричні властивості напівпровідників надзвичайно чутливі до найменших змін хімічного складу. Сторонні включення або надлишкові відносно стехіометричного складу атоми речовини стають джерелами вільних зарядів. Цим фактом пояснюється використання легованих напівпровідників як ЧЕ сенсорів. Сторонні домішки, а також надлишок тих або інших іонів в порівнянні зі стехіометричним складом не тільки різко

мінюють електропровідність, але й визначають знак носіїв струму в напівпровіднику. Зміни адсорбції частинок, що спостерігаються на поверхні оксидів металів залежать як від колективних, так і локальних властивостей поверхні, причому домінуюча роль механізму визначається концентрацією та хімічною природою введеної в об'єм або на поверхню домішки. При адсорбції молекул, хімічна активність яких значно менше, ніж атомів, переважає колективний механізм.

Домішковими можуть бути атоми того ж металу, який утворює оксид, або ж вакансії кисню в ґратці оксиду. Як домішки можуть виступати іони іншого металу, які знаходяться у валентному стані, відмінному від валентного стану іонів в оксиді. Такий стан викликає перехід деяких іонів оксиду в інший валентний стан (наприклад, іон літію Li^+ в оксиді нікелю NiO викликає перехід деяких іонів Ni^{2+} в стан Ni^{3+} , іони титану Ti^{4+} в оксиді заліза Fe_2O_3 викликають перехід деяких іонів Fe^{3+} в стан Fe^{2+}). Метод поверхневого легування плівок оксидів металу атомами металів показав, що домішкові атоми металів незалежно від їх валентності збільшують як електропровідність оксидів n-типу, так і їх чутливість, тобто вони є центрами адсорбції акцепторних частинок. Значення чутливості двоокису олова SnO_2 , що має провідність n-типу, в залежності від типу легуючої домішки з вказівкою оптимальної робочої температури матеріалу наведені в табл. 16.8.

Таблиця 16.8

Чутливість (відносні одиниці) напівпровідникових сенсорів в залежності від типу легуючої домішки

| Напів- провідн ик | Газ, що вимірюється, та температура | | | | | | | |
|-------------------------|-------------------------------------|------|-------------------------|------|----------------------------------------|------|--------------------------|------|
| | 0,02 % CO | T, K | 0,8 % H ₂ | T, K | 0,2 % C ₃ H ₈ | T, K | 0,5 % CH ₄ | T, K |
| SnO ₂ | 4 | 623 | 37 | 73 | 49 | 623 | 20 | 723 |
| +Pt | 136 | 293 | 3600 | 293 | 38 | 548 | 19 | 573 |
| +Pd | 12 | 293 | 419 | 423 | 75 | 523 | 20 | 598 |
| +Ag | 8 | 373 | 666 | 373 | 89 | 623 | 24 | 673 |
| +Cu | 7 | 423 | 98 | 573 | 48 | 598 | 20 | 623 |
| +N | 7 | 473 | 169 | 523 | 67 | 573 | 9 | 623 |

Збільшення домішкової провідності оксидів n-типу при хемосорбції донорних частинок пояснюється тим, що адсорбційними центрами таких частинок можуть бути електронегативні атоми матеріалу, наприклад регулярні іони кисню ґратки.

Легування полікристалічних плівок селеніду кадмію $CdSe$ атомами індію In (донор) і селену Se (акцептор) показало, що в разі легування донорною домішкою спостерігається збільшення адсорбційної чутливості поверхні селеніду кадмію $CdSe$ завдяки зменшенню енергії в комплексі «іонізована молекула кисню – центр адсорбції» через диполь-дипольну взаємодію між

комплексами $\text{In}^+\delta\text{O}_2^{-\delta}$ (δ – міра перенесення заряду в комплексі) і утворення кластерів індію. При зменшенні концентрації комплексів «адсорбована молекула кисню – центр адсорбції» завдяки легуванню поверхні акцепторами енергія зв'язку хемосорбційного кисню з центром адсорбції зростає внаслідок зменшення взаємодії між адсорбційними комплексами.

Існуючі теорії ще не дозволяють оптимально підібрати матеріал ЧЕ і легуючу домішку. Але можна виділити деякі загальні правила підбору найбільш активних напівпровідникових матеріалів для прискорення тих або інших реакцій:

- каталітична активність основного матеріалу ЧЕ залежить від значення і знаку електропровідності, яка визначається положенням рівня Фермі;

- для збільшення швидкості акцепторної реакції необхідно ввести в напівпровідник домішки, що підвищують його електронну провідність, а для прискорення донорної реакції – домішки, що підвищують його діркову провідність;

- каталітична активність тим вища, чим вища робота виходу електрона, тобто чим нижче рівень хімічного потенціалу електронів;

- каталітична активність напівпровідника обернено пропорційна до ширини забороненої зони, яка визначається ступенем іонного зв'язку, атомною вагою елементів, що утворюють напівпровідник, їхньою поляризованістю.

Відомо також, що чутливість плівкових ЧЕ набагато вища, ніж об'ємних монокристалів, а у полікристалічних плівок вона значно вища, ніж у монокристалічних. Це й визначило широке використання полікристалічних напівпровідникових в якості ЧЕ сенсорів.

Гетерогенна структура полікристалічних плівок з атомарно розупорядкованими поверхнями чутливіша до хемосорбції, ніж монокристалічна, оскільки навіть при парціальному тиску 10^{-6} Па густина поверхневих станів у полікристалічній плівці досить велика для помітної зміни провідності.

Лекція 17

АКТИВНІ НАПІВПРОВІДНИКОВІ ФОТОЕЛЕКТРИЧНІ ЕЛЕМЕНТИ.

План

- 17.1. Світло діоди
 - 17.1.1. Основні параметри та характеристики світлодіодів
 - 17.1.2. Принцип роботи інжекційних світлодіодів з р-n-переходами
 - 17.1.3. Світлодіоди з антистоксовими люмінофорами
- 17.2. Напівпровідникові лазери
 - 17.2.1. Лазери на гомоструктурах
- 17.3. Фоторезистори
 - 17.3.1. Принцип дії фоторезисторів
 - 17.3.2. Основні параметри фоторезисторів
 - 17.3.3. Конструкція фоторезисторів
 - 17.3.4. Області використання фоторезисторів
- 17.4. Фотодіоди
 - 17.4.1. Класифікація фотодіодів
 - 17.4.2. Р-n-фотодіоди
 - 17.4.3. Р-i-n-фотодіоди
 - 17.4.4. Матеріали і конструкція р-i-n-фотодіодів
 - 17.4.5. Лавинні фотодіоди
 - 17.4.6. Високочастотні фотодіоди
 - 17.4.7. Інжекційні фотодіоди
 - 17.4.8. S-фотодіоди
 - 17.4.9. Використання фотодіодів

Література

- 17.1. Ишанин Г.Г., Панков С.Д., Радайкин В.С. Источники и приемники излучения. – М.: Машиностроение, 1982. – 222 с.
- 17.2. Источники и приемники излучения / Г.Г. Ишанин, Е.Д. Панков, А.А. Андреев и др. – СПб: Политехника, 1991. – 240 с.
- 17.3. Фізичні основи електронної техніки / За ред. З.Ю. Готри. – Львів: Видавництво «Бескид-Біт», 2004. – 880 с.
- 17.4. Ишанин Г.Г., Панков С.Д., Радайкин В.С. Источники и приемники излучения. – М.: Машиностроение, 1982. – 222 с.
- 17.5. Корнійчук В.І., Мосорін П.Д. Волоконно-оптичні компоненти, системи передачі та мережі. – Одеса: Друк, 2001. – 364 с.

17.1. Світлодіоди

Світлодіодом називається випромінюючий напівпровідниковий прилад з р-n-переходом, призначений для безпосереднього перетворення електричної енергії в енергію некогерентного оптичного випромінювання. Принцип дії світлодіода оснований на явищі електролюмінесценції під час протікання струму в структурах з р-n-переходом.

Властивість випромінювати в оптичному діапазоні спектра мають деякі складні напівпровідникові з'єднання, основою яких є галій, миш'як, карбід кремнію, що

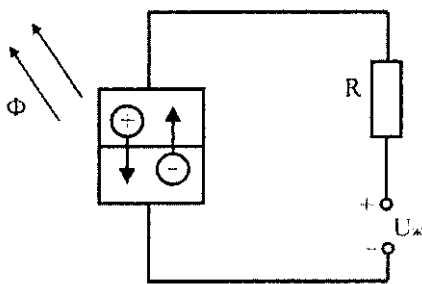


Рис. 17.1. Схема для пояснення роботи світлодіода ($U_{ж}$ – напруга живлення).

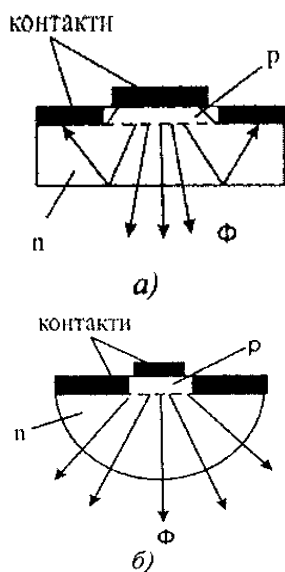


Рис. 17.2. Будова світлодіодів з плоскою (а) та напівсферичною (б) поверхнею випромінювання.

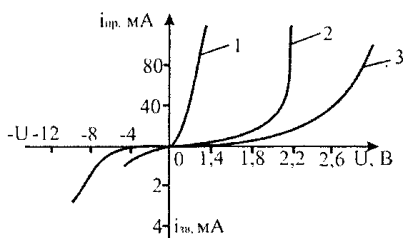


Рис. 17.3. Вольт-амперна характеристика: 1 – світлодіод на основі GaAs; 2 – GaP; 3 – SiC.

служать вихідним матеріалом для виготовлення світлодіодів.

В результаті подачі на р-п-перехід прямої напруги (рис. 17.1) виникає інжекція неосновних носіїв заряду через перехід і наступна рекомбінація їх з основними носіями з виділенням енергії у вигляді оптичного випромінювання.

Рекомбінаційне випромінювання у світлодіоді генерується у шарі, що прилягає до межі розділу р-п-напівпровідників, і виводиться на поверхню через п-область. Схематичну будову світлодіода з плоскою або напівсферичною випромінюючою поверхнями наведено на рис. 17.2. Напівпровідникова структура світлодіода знаходиться у пластмасовому або металічному корпусі з вікном (лінзою) для виходу випромінювання.

17.1.1. Основні параметри та характеристики світлодіодів. *Ефективність* перетворення електричної енергії у світлову ($\eta_{ке}$) що називається інколи **внутрішнім квантовим виходом**, – це відношення кількості фотонів, які виділились в результаті рекомбінацій, до кількості носіїв, що з'явилися у р-п-переході. Практично $\eta_{ке}$ становить величину 0,7...0,9.

Відношення кількості квантів світла, що вийшли зі світлодіода, до кількості електронів, що рекомбінували при цьому, називають **зовнішнім квантовим виходом** η_k . Інакше η_k можна виразити як відношення вихідної потужності випромінювання до струму через р-п-перехід. η_k менше від $\eta_{ке}$ внаслідок поглинання частини випромінювання у товщі напівпровідника і втрат випромінювання на межі напівпровідник-повітря за рахунок повного внутрішнього відбивання для променів, кут падіння яких на границю перевищує критичний. Величина критичного кута $\alpha_{крит}$ визначається як

$$\alpha_{крит} = \arcsin(n_2 / n_1),$$

де n_2 – показник заломлення зовнішнього середовища (для повітря $n_2 = 1$); n_1 – показник заломлення напівпровідника ($n_1 > n_2$).

Зовнішній квантовий вихід, а отже, потужність випромінювання світлодіода значною мірою визначається його геометричною формою. Напівсферична форма на відміну від плоскої практично повністю виключає повне внутрішнє відбивання. Збільшенню вихідної потужності сприяє нанесення просвітлюючого покриття на вивідну поверхню.

На рис. 17.3 наведена вольт-амперна характеристика світлодіода.

Для світлодіодів характерна висока густина **спектральної енергетичної яскравості**. Основна частина випромінювання приходить на спектральний інтервал $\Delta\lambda$ шириною 400...1000 Å (рис. 17.4). Довжина хвилі випромінювання λ_{max} визначається енергетичною шириною зон напівпровідника, частіше шириною

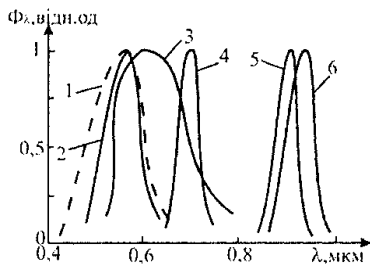


Рис. 17.4. Спектральні характеристики світлодіодів: 1 – відносна спектральна світлова ефективність монохроматичного випромінювання; 2 – GaP; 3 – SiC, легований B та N; 4 – GaP легований Zn, Fe та O₂; 5 – GaAs, легований Zn та Fe; 6 – GaAs, легований Si та Fe (Φ – функція розподілу спектральної густини потоку випромінювання; λ – довжина хвилі).

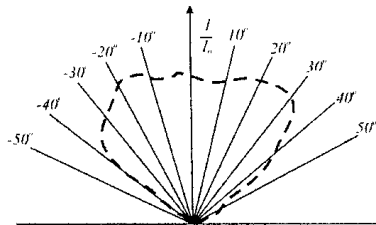


Рис. 17.5. Індикатриса випромінювання арсенід-галієвого світлодіода.

забороненої зони E_g .

З ростом температури значення λ_{max} росте за лінійним законом, причому $d\lambda_{max}/Dt \approx 3 \text{ \AA/градус}$.

Суттєвою перевагою світловодів є їх швидкодія, що дозволяє реалізувати імпульсний режим роботи і отримати короткі світлові імпульси не з крутими фронтами наростання і спаду з частотою повторення до 10^8 Гц .

Імпульсний режим роботи найбільш ефективний: в цьому режимі прямий імпульсний струм може досягати одиниць ампер, а яскравість – $10^4 \text{ кд}\cdot\text{м}^{-2}$, що значно перевищує яскравість свічення при неперервному режимі.

Випромінювання світлодіодів є ненаправленим (рис. 17.5). Тому для концентрації випромінювання в певному напрямку використовуються зовнішні фокусуючі лінзи або відбиваюче покриття, яке наноситься безпосередньо на кристал напівпровідника.

Найбільш суттєвими недоліками світлодіодів є розкид характеристик від зразка до зразка та їх залежність від температури.

Світлодіоди мають малі габарити, підвищену економічність, тривалий строк служби, стійкість до механічної взаємодії, високу швидкодію, що дозволяє здійснювати модуляцію випромінювання струмом живлення. Всі ці позитивні властивості сприяють успішному застосуванню світлодіодів у якості твердотільних цифрових індикаторів, у пристроях оптичного зв'язку, в оптоелектронних приладах тощо.

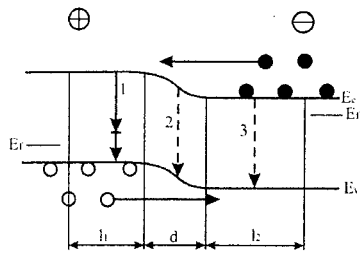
17.1.2. Принцип роботи інжекційних світлодіодів з р-п-переходами.

Світловипромінювальний напівпровідниковий діод (або світлодіод) є одним з основних джерел випромінювання в оптоелектроніці. Він являє собою увімкнений у прямому напрямку р-п-перехід (рис. 17.6), в якому відбувається рекомбінація електронів і дірок як в області об'ємного заряду завширшки d , так і по обидва боки від цієї області на віддалі дифузійних довжин електронів в р-області (l_1) і дірок в п-області (l_2). Зовнішня напруга знижує потенціальний бар'єр на межі р- і п-областей і створює умови для інжекції (введення) електронів в р-область і дірок в п-область.

У міру підвищення прямої напруги U струм через перехід експоненціально зростає і при $qU \gg kT$ визначається виразом $I = I_n \cdot \exp\left(\frac{qU}{kT}\right)$, який отримуємо

із загального виразу для вольт-амперної характеристики р-п-переходу (q – заряд електрона; I_n – струм насичення; k – стала Больцмана).

Загальна кількість рекомбінацій щосекунди визначається силою струму. Частина цих рекомбінацій відбувається з випромінюванням, і потік випромінювання Φ , виражений у кількості фотонів, випромінених щосекунди, становить:



$$\Phi = h_{\kappa} \frac{I}{q} = h_{\kappa} \frac{I_n}{q} \left(\frac{qU}{kT} \right).$$

Випромінювальним є один з боків р-п-переходу (наприклад, р-область), тому бажано, щоб частка електронного струму, що потрапляє в цю область (коефіцієнт інжекції γ), була максимальною. Під час інжекційної електролюмінесценції величина γ , а отже, і $\eta_{\kappa} = \gamma P$, можуть залежати від струму (рис. 17.6, б). У випадку низьких U та I переважає рекомбінація в області просторового заряду, де ймовірність випромінювальних переходів у потрібній спектральній області невелика, тому з підвищенням напруги і струму γ і η_{κ} спочатку зростають, а потім стають приблизно постійними. При дуже великих U та I можна зменшити η_{κ} збільшенням кількості безвипромінювальних Оже-рекомбінацій, а інколи і підвищенням температури зразка.

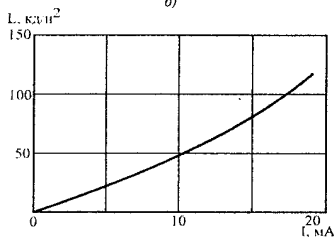
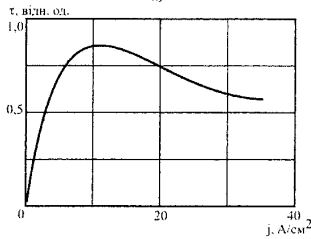


Рис. 17.6. Рекомбінація електронів і дірок в р-п-переході, який ввімкнено в прямому напрямку (а), відносний квантовий вихід GaP-світлодіода з червоним випромінюванням (б) і яскравість GaAs-світлодіода як функція струму (в).

Існування залежності $\eta_{\kappa}(I)$ призводить до того, що залежність $\Phi(I)$ є лінійною тільки в області середніх струмів, коли $\eta_{\kappa} \approx const$. В області низьких струмів Φ та яскравість L зростають з підвищенням I надлінійно (рис. 17.6, в), а в області великих I - сублінійно.

Зовнішній квантовий вихід світлодіода $\eta_{\kappa} = \eta_{\kappa e} \cdot K_0$ завжди нижчий від внутрішнього, що зумовлено, перш за все, втратами світла на виході з діода (K_0 -коефіцієнт втрат світла). Через повне внутрішнє відбиття при кутах падіння, більших від θ_0 , через поверхню напівпровідника вийде тільки світло в межах конуса з напівкутом θ_0 поблизу вершини (рис. 17.7, а). Цей кут визначається з умови $\theta_0 = \arcsin(1/n)$ (n - показник заломлення матеріалу) і, наприклад, для GaP становить 17,70. За такого θ_0 через межу напівпровідник-повітря може пройти тільки приблизно 2% світла, що випромінюється р-п-переходом рівномірно у всі боки. Покращити умови виходу світла з напівпровідника можна, надаючи йому сферичну форму або додаючи прозору пластмасову лінзу сферичної форми (рис. 17.7, б). Якщо втрати на переході в лінзу усунуті, а світло, що випромінюється в нижню півсферу, поглинається в невипромінювальному матеріалі (контактний шар), то максимальне значення $K_0 = 0,5$.

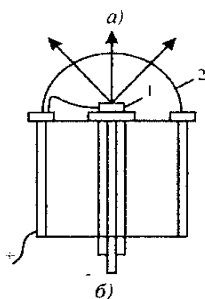
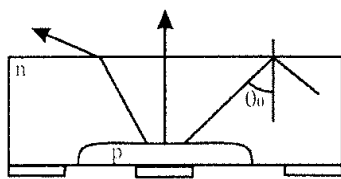


Рис. 17.7. Відбивання і заломлення світла в пластинці напівпровідника (а) і одна з конструкцій світлодіода (б): 1 - кристал з р-п-переходом; 2 - лінза.

Розглянемо як приклад чинники, які впливають на величину $\eta_{\kappa e}$ у світлодіодів GaP:Zn та O з червоним світінням. Основне випромінювання цих діодів ($h = 1,8$ eВ, де h - стала Планка; ν - частота коливань) з'являється внаслідок рекомбінації екситонів, зв'язаних на комплексах Zn-O (після захоплення електрона цей комплекс отримує заряд і притягає дірку, утворюючи екситон). Випромінювальною є р-область, і найбільше досягнуте значення коефіцієнта інжекції в цю область $\gamma = 0,8$. Рекомбінацією в області просторового заряду можна знехтувати, оскільки її товщина становить приблизно $1 \cdot 10^{-5}$ см, а дифузійна довжина електронів $l_1 \approx 10^{-4}$ см. Вимірювання люмінесценції р-області дають

значення ймовірності випромінювальних рекомбінацій $P_g = 0,3$ при 300 К (тобто $2/3$ рекомбінацій відбувається без випромінювання). Отже, при $K_0 = 0,5$ зовнішній квантовий вихід $\eta_{ке} = 0,8 \cdot 0,3 \cdot 0,5 = 12 \%$.

Діоди з GaP з домішкою азоту дають зелене свічення і мають менше $\eta_{ке}$. Промислові світлодіоди з GaP з червоним або зеленим свіченням мають яскравість 50-100 кд/м² і працюють при прямій напрузі 3-5 В.

Найвищий вихід спостерігається у діодів з арсеніду галію, які випромінюють інфрачервоне світло. Залежно від конструкції діодів зовнішній вихід за кімнатної температури становить 8-20%, а за низької температури вихід досягає 40%. Це свідчить про те, що внутрішній квантовий вихід інжекційної електролюмінесценції у цій речовині може наближатись до 100%.

Зовнішній квантовий вихід промислових діодів GaAs:Si ($h\nu = 0,94$ eV) за кімнатної температури дорівнює приблизно 8%. Якщо активна область сферична, то вихід підвищується до 20-30%. Відповідний зовнішній енергетичний вихід становить близько 15%. Інерційність світлодіодів визначається в основному часом дифузії неосновних носіїв у товщині напівпровідника і дорівнює 10^{-7} - 10^{-6} с.

Діоди з GaAs є мініатюрними приладами розміром в кілька міліметрів і масою $\sim 0,5$ г. Потужність випромінювання при струмі 100 мА становить приблизно 0,5 мВт, робоча напруга $\sim 1,7$ В.

До цього часу йшлося про гомопереходи, тобто р-п-переходи, створені в одній і тій самій речовині. Особливі властивості мають гетеропереходи, отримані між р- і п-областями матеріалів з різною шириною забороненої зони. Вибір таких матеріалів обмежується необхідністю точного узгодження сталих ґраток обох компонент (на межі утворюється велика кількість дефектів, і відповідні до них локальні рівні погіршують характеристики переходів). Прикладом матеріалів з добре узгодженими ґратками є GaAs і AlAs, які утворюють неперервний ряд твердих розчинів $Ga_{1-x}Al_xAs$. Зміною x можна змінити ширину забороненої зони цього потрібного з'єднання, наприклад, при $x = 0,3$ $E_g = 1,8$ eV (у GaAs $E_g = 1,4$ eV). Вводячи необхідні домішки в обидві речовини, можна створити р-п-переходи з різними властивостями.

На рис. 17.8ображено одну з можливих енергетичних схем гетеропереходу. У випадку увімкнення його в прямому напрямку відбувається практично тільки інжекція електронів в р-область, оскільки для дірок потенціальний бар'єр виявляється надто високим. У результаті з широкозонного матеріалу можна отримати однібічну інжекцію електронів ($\gamma = 1$) у вузькозонний люмінесцентний матеріал і підвищити таким чином квантовий вихід світлодіода.

Іншою перевагою гетеропереходу є можливість

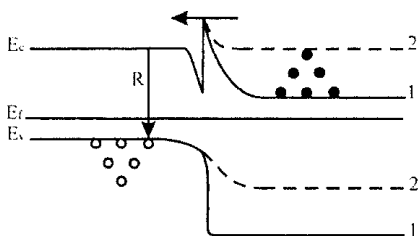


Рис. 17.8. Гетероперехід за відсутності зовнішньої напруги (1) і після увімкнення в прямому напрямку (2).

виведення випромінювання без втрат на поглинання через широкозонний матеріал. Випромінювання GaAs при $T=197\text{ K}$ має максимум при $1,38\text{ eV}$, тобто воно сильно поглинається самим GaAs, але слабше – більш широкозонним матеріалом $\text{Ga}_{1-x}\text{Al}_x\text{As}$.

Гетеропереходи використовують також для отримання когерентного випромінювання.

17.1.3. Світлодіоди з антистоксовими люмінофорами. Оскільки квантовий вихід діодів GaAs:Si, що випромінюють інфрачервоне світло, значно вищий від виходу діодів, що випромінюють у видимій області, існує можливість використання інфрачервоних діодів для отримання видимого випромінювання за допомогою так званих *антистоксових люмінофорів*.

За правилом Стокса частота фотолюмінесценції звичайно менше від частоти, що збуджує випромінювання. У деяких випадках спостерігається обернене співвідношення між цими частотами, що і відбувається у антистоксових люмінофорів. Були отримані антистоксові люмінофори з максимумом збудження близько $h\nu = 1,3\text{ eV}$, тобто в області випромінювання діодів GaAs:Si. Якщо на поверхню GaAs-діодів нанести шар порошкоподібного антистоксового люмінофора, то під дією ІЧ-випромінювання діода він буде випромінювати у видимій області завдяки послідовному поглинанню і підсумовуванню енергії двох квантів ІЧ-випромінювання. Ефективність перетворення ІЧ-випромінювання у видиме зростає зі збільшенням рівня збудження люмінофора (струму через діод).

Як стоксові люмінофори використовують фториди, оксисульфід лантану та інші сполуки, активовані ітербієм і ербієм, наприклад, люмінофор $\text{La}_2\text{O}_3\text{S:Yb, Er}$, за допомогою якого може бути отримане зелене свічення яскравістю до 400 кд/м^2 та густиною струму через діод 5 A/см^2 . Загальний квантовий вихід такого подвійного перетворення електричної енергії у світлову не перевищує 1%, але приблизно такий самий квантовий вихід мають і світлодіоди з GaP із зеленим випромінюванням. Отже, використання діодів з антистоксовим перетворенням енергії може бути виправданим, особливо у випадку необхідності підвищеної яскравості зеленого випромінювання. Швидкодія діодів з антистоксовими люмінофорами визначається процесами в люмінофорах і може бути нижче ніж у звичайних світлодіодів.

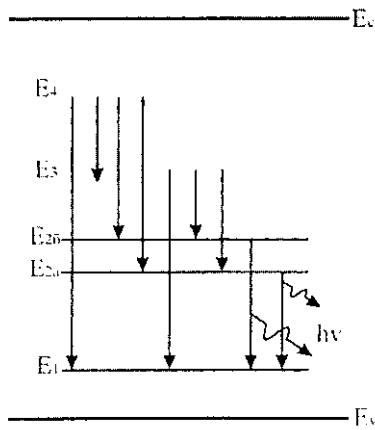


Рис. 17.9. Енергетичні рівні хрому в забороненій зоні рубіну: E_1 – основний стан; E_2 , E_3 , E_4 – збуджений стан.

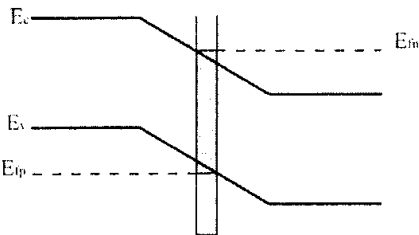


Рис. 17.10. Реалізація інверсної заселеності в тунельному переході.

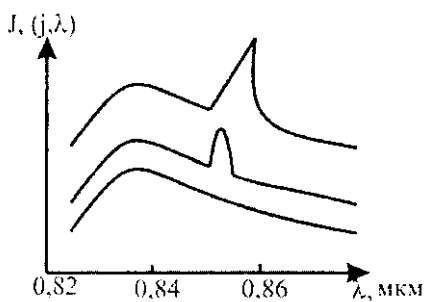


Рис. 17.11. Спектр випромінювання GaAs-лазера.

17.2. Напівпровідникові лазери

Як робоче середовище в перших твердотільних лазерах використовувались штучно вирощені кристали рубіну, тобто кристали окису алюмінію Al_2O_3 з вмістом домішки хрому Cr приблизно 0,05%. В результаті збудження атома хрому електрони з рівня E_1 переходять на рівні E_3 , E_4 (рис. 17.9), час перебування на яких малий (близько 10^{-8} с). Тому дуже швидко здійснюється не випромінювальний перехід на рівні E_2 , де електрон затримується на триваліший час (близько 10^{-3} с), що й дає змогу створювати на рівнях E_2 вищу концентрацію електронів, ніж на рівні E_1 , тобто реалізувати інверсну заселеність. Переходи $E_2 \rightarrow E_1$ випромінювальні (оскільки рівень E_2 є вузьким дублетом, вони супроводжуються двома лініями випромінювання $\lambda_1 = 0,6943$ мкм і $\lambda_2 = 0,6929$ мкм). Спонтанне випромінювання, що виникає при цьому, є стимулюючим і підсилюється індукованими переходами.

На оптично відполіровані та строго паралельні торці рубінового стрижня наносяться дзеркальні покриття частково прозорі для виходу випромінювання. Під час проходження випромінювання вздовж стрижня відбувається багаторазове відбивання від торців – у результаті інтенсивність випромінювання підсилюється, здійснюється додатний зворотний зв'язок і реалізується режим генерації. Слід підкреслити, що оптичне підсилення – процес, обернений до поглинання. В цьому випадку коефіцієнт поглинання $\alpha < 0$ і інтенсивність зростає експоненціально. Крім того, в результаті багаторазового відбивання від торців зі стрижня виходять лише промені, нормальні до торців, тобто створюється вузько напрямлений пучок випромінювання великої інтенсивності. Збудження рубіну (накачування лазера) здійснюється за допомогою потужної лампи-спалаху.

Необхідна умова для одержання інверсної заселеності рівнів у напівпровіднику полягає у забезпеченні в заданому об'ємі напівпровідника одночасного виродження електронного та діркового газів. Такий стан можна реалізувати в області виродженого переходу, увімкненого в прямому напрямку (рис. 17.10).

Однак індуковане випромінювання в цьому випадку виникає лише тоді, коли струм через перехід буде перевищувати деяке порогове значення (рис. 17.11). При значеннях струму, нижчих від порогових, спектр випромінювання дуже широкий (процес спонтанного випромінювання). У випадку зростання струму спектр випромінювання значно звужується. Оскільки перехід зміщується постійним струмом, лазер працює в неперервному режимі. Однак внаслідок значного розігрівання переходу такий режим роботи можливий лише за низьких температур.

На рис. 17.12, а наведено енергетичну схему лазера за

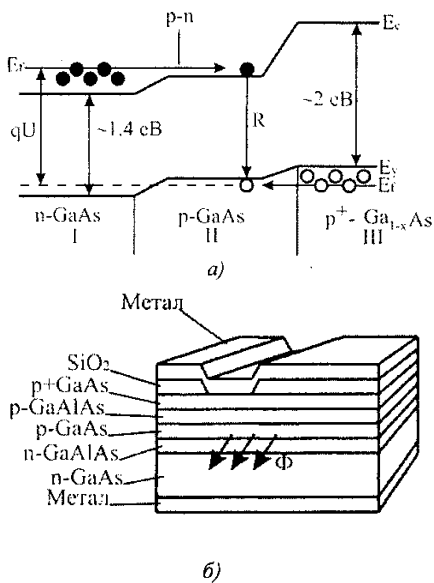


Рис. 17.12. Енергетична схема інжекційного лазера (а) і структура лазера з двома переходами (б).

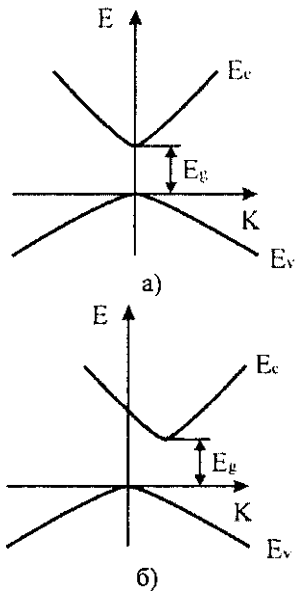


Рис. 17.13. Енергетична зонна структура напівпровідників з прямою (а) і непрямою (б) структурою енергетичних рівнів.

наявності зовнішньої напруги U . Через р-п-ієрохїд їнжєктуютьє електрони з п-область I в активну область II (завтовшки ~ 1 мкм), де вїдбуваєтьє випромїнювання фотонїв з енергїєю $h\nu \approx 1,4$ еВ. Перехїд типу р-р $^+$ мїж р-областью II і р $^+$ -областью III створює бар'єр для електронїв, що потрапили в активну область II , і сприяє накопиченню електронїв в цїй області. Лазерний ефект досягаєтьє за певних порогових значень струму через перехїд (приблизно 1 кА/см 2 при 300 К). Починаючи з цих значень струму, спектральна смуга випромїнювання значно звужуєтьєся. Для зменшення робочих струмїв і ослаблення нагрївання активний шар часто скорочують до смуги завширшки $5-20$ мкм, що проходить вїд однїєї вїдбивальної поверхнї до їншої. Цього досягають застосуванням вузького металевого електрода (верхнього на рис. 17.12, б). У подїбних пристроях знижуєтьєся як пороговий струм (приблизно до 100 мА за кїмнатної температури), так і їнерцїйнїсть внаслїдок зменшення ємностї переходїв.

Матерїалом, який їнжєктує електрони, може бути бїльш широкозонний $Al_xGa_{1-x}As$ (рис. 17.12, б). У цьому випадку активний шар GaAs р-типу розташовуєтьєся мїж двома широкозонними напївпровїдниками, якї мають нижчї коефїцїєнти заломлення. Це призводить до пїдсилення вїдбивання свїтла вїд бокових шарїв і, отже, до зменшення втрат свїтла.

Потужнїсть випромїнювання лазерїв, що працюють в неперервному режимї, становить близько $0,1$ Вт. У випадку їмпульсного збудження потужнїсть може бути значно пїдвищена, оскїльки нагрївання приладу буде слабке. ККД їнжєкцїйних лазерїв сягає 50% , їнерцїйнїсть становить $1 \cdot 10^{-9}$ с, напруга живлення не перевищує 3 В, а розмїри – декїлька мїліметрїв. Модуляцїя свїтлового потоку може здїйснюватїсь змїною напруги.

У напївпровїдникових лазерях випромїнювальнї перехїди вїдбуваютьєся не мїж вузькими дискретними енергетичними станами, а мїж широкими енергетичними зонами, тому перехїд електрона їз зони провїдностї у валентну зону з наступною рекомбїнацїєю зумовлює випромїнювання, яке лежить у широкому спектральному їнтервалї (кїлька десяткїв нанометрїв), набагато бїльшому, нїж ширина смуги випромїнювання газових лазерїв.

Властивостї напївпровїдникїв описуютьєся поведїнкою електронїв поблизу «дно» зони провїдностї E_c і дїрок поблизу «стелї» валентної зони E_v . Кїнетична енергїя E_v електронїв у зонї провїдностї (або дїрок у валентнїй зонї) залежить вїд хвильового вектора $|\mathbf{k}| = \frac{2p}{l}$.

Ця залежнїсть для прямої (а) і непрямої (б) структури енергетичних зон наведена на рис. 17.13.

Кїлькїсть енергетичних рївнїв, якї знаходятьєся в

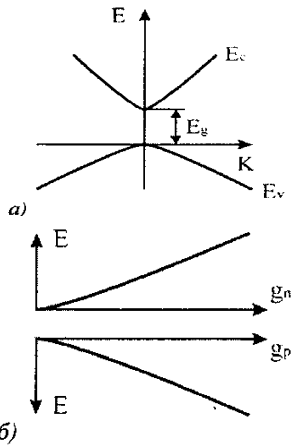


Рис. 17.14. Енергетична зонна структура напівпровідника (а) та функція густини станів (б) у зоні провідності та у валентній зоні.

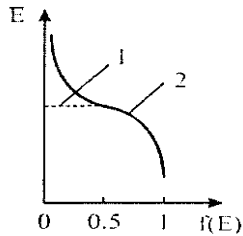


Рис. 17.15. Функція ймовірності заповнення квантового стану з енергією E для $T=0$ (1) і $T>0$ (2).

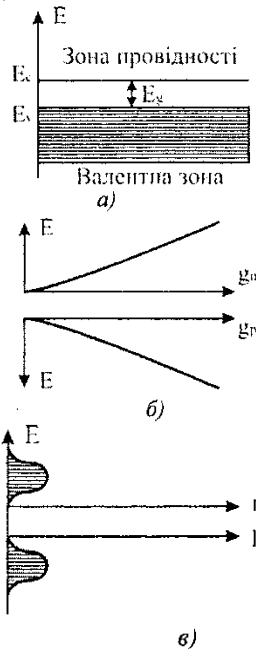


Рис. 17.16. Схема заповнення енергетичних рівнів у напівпровіднику: а – структура енергетичних зон; б – функції густини станів; в – концентрація носіїв струму у власному напівпровіднику.

інтервалі енергій від E до $E+dE$, визначається через функцію густини станів

$$dN(E) = q(E)dE.$$

Функції густини станів наведені на рис. 17.14.

У випадку теплової рівноваги ймовірність заповнення квантового стану з енергією E для ферміонів визначається функцією розподілу Фермі – Дірака:

$$f(E, T) = \frac{1}{e^{\frac{E-E_F}{kT}} + 1}$$

На рис. 17.15. наведено графіки ймовірності заповнення квантового стану з енергією E для $T=0, K(1)$ і $T>0, K(2)$.

Ймовірність заповнення квантового стану діркою рівна ймовірності відсутності в цьому стані електрона:

$$f(E, T) = 1 - \frac{1}{e^{\frac{E-E_F}{kT}} + 1} = \frac{1}{e^{\frac{E_F-E}{kT}} + 1}$$

Для власного напівпровідника концентрація електронів у зоні провідності або дірок у валентній зоні визначається з формул:

$$n_0 = \int_{E_c}^{\infty} f_0(E) q_n(E) dE;$$

$$p_0 = \int_{E_c}^{\infty} f_{0p}(E) q_p(E) dE;$$

Графіки функції густини станів концентрації носіїв струму для власного напівпровідника наведено на рис. 17.16, а, б, в. У випадку домішкових напівпровідників ті самі залежності мають вигляд (рис. 17.17, 17.18) для напівпровідника р- і n-типів.

Розглянутий ефект оптичного підсилення можна використовувати для реалізації лазерного ефекту у разі виконання двох умов: підсилення повинне перевищувати втрати; повинен існувати додатний зворотний зв'язок.

Розглянемо функції розподілу у квазірівноважному стані для електронів.

1. Енергія накачування невелика (концентрація електронно – діркових пар невелика). Квазірівні Фермі для електронів попадають у заборонену зону, тобто $E_{Fn} - E_{Fp} < E_g$ (рис. 17.19, а)).

Ймовірність заповнення електронних станів, які відповідають дну зони провідності, буде меншою, ніж 0,5, а ймовірність заповнення стану, який дорівнює стелі валентної зони – більшою від 0,5, тобто стани в зоні провідності заселені з меншою ймовірністю, ніж стани валентної зони (на рисунку заповнені стани заштриховані), тобто інверсної населеності не виникає.

2. Енергія накачування велика (концентрація електронно-діркових пар велика). Квазірівні Фермі потрапляють у зону провідності і у валентну зону (рис. 17.19, б).

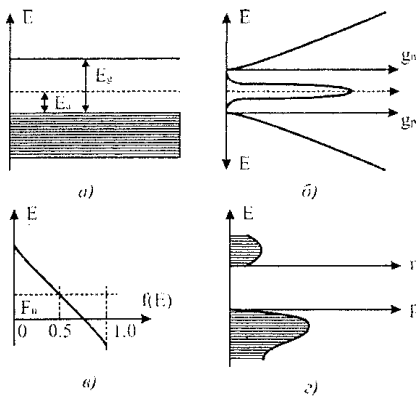


Рис. 17.17. Схема заповнення енергетичних рівнів в напівпровіднику р-типу:

- а) зонна діаграма;
- б) функції густини станів;
- в) розподіл Фермі;
- г) концентрація електронів.

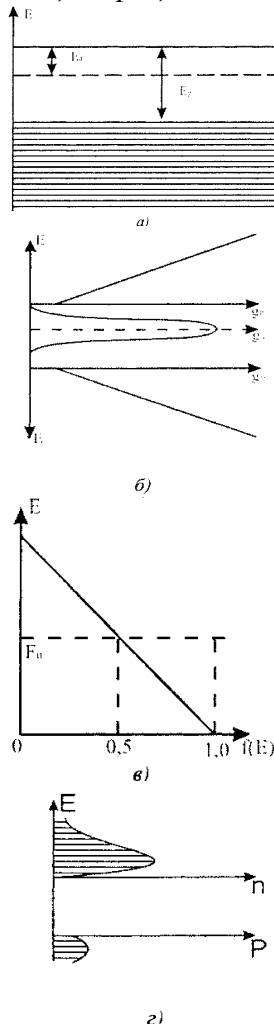


Рис. 17.18. Схема заповнення енергетичних рівнів в напівпровіднику n-типу: а) зонна діаграма; б) функції густини станів; в) розподіл Фермі; г) концентрація електронів.

В цьому випадку ймовірність заповнення електронних станів, що відповідають дну зони провідності, більша, ніж 0,5, а станів, що відповідають стелі валентної зони – менша, ніж 0,5, тобто виникає ситуація з інверсною населеністю.

Стан з інверсною населеністю реалізується:

- оптичним накачуванням;
- інжекцією неосновних носіїв через р-п-перехід;
- збудженням електронним променем (електронно-променево накачування);
- збудженням в сильному електричному полі.

Критерій реалізації стану з інверсною населеністю на міжзонних переходах

$$E_{Fn} - E_{Fp} < E_g.$$

17.2.1. Лазери на гомоструктурах. Інверсія населеності досягається за рахунок інжекції електронів у р-п-перехід. У вироджених напівпровідниках рівні Фермі знаходяться у дозволених зонах і при рівновазі для електронів і дірок збігаються. В області р-п-переходу утворюється бар'єр, який не дозволяє переходити в зону основним носіям. Створити в заданому об'ємі напівпровідника одночасне виродження електронного і діркового газу – означає одержати інверсну населеність рівнів. Такий стан можна одержати в області виродженого р-п-переходу у випадку вмикання його в прямому напрямку. Область р-п-переходу є підсилювальним середовищем для світла, яке поширюється в площині р-п-переходу. Для одержання генерації необхідно ввести зворотний зв'язок, який можна одержати, помістивши р-п-перехід між двома дзеркалами. Роль останніх здебільшого виконують гладенькі грані самого напівпровідникового кристала. Гомоструктурні лазери виготовляються з GaAs або $Ga_{1-x}Al_xAs$. Формуються вони за рахунок епітаксійного росту шару р-типу на підкладці n-типу. Паралельні торці виготовляються сколюванням вздовж кристалографічної осі для роботи дзеркалами резонатора. Внаслідок великого показника заломлення напівпровідникового матеріалу коефіцієнт відбивання від граней становить 30...50%. Бічні грані залишаються шорсткими.

Розглянемо коефіцієнт підсилення напівпровідникового лазера, виготовленого на напівпровідниковому кристалі з параметрами, наведеними на рис. 17.20. Тут: l – довжина кристала; w – ширина кристала; d – ширина моди; t – час життя електрона в р-області.

Коефіцієнт підсилення дорівнює:

$$G(n) = \frac{c^2 (n_2 - n_1) q(\nu)}{8\pi n^2 v^2 t_p V},$$

де t_p – час життя електрона в р-області; $V = lwd$ – об'єм моди.

У випадку низьких температур $n_1 = 0$.

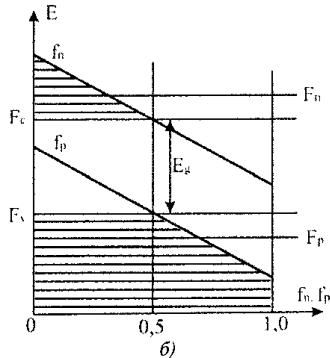
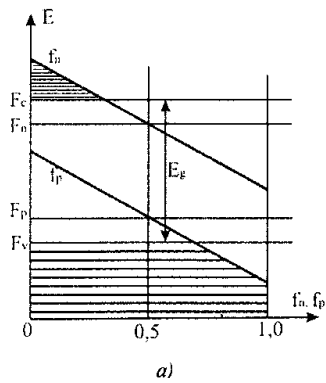


Рис. 17. 19. Ймовірність заповнення нерівноважними носіями заряду енергетичних зон у напівпровіднику при малому (а) і великому (б) рівнях збудження.

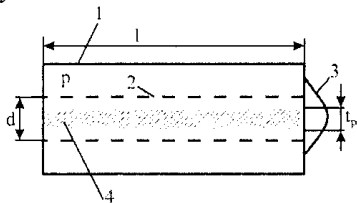


Рис. 17. 20. Схема напівпровідникового лазера:

- 1 – активний елемент;
- 2 – об'єм моди;
- 3 – розподіл інтенсивності;
- 4 – активний шар.

Умова самозбудження лазера:

$$e^{(G_r - a)l} R_D = 1,$$

де a – коефіцієнт втрат; G_r – пороговий коефіцієнт підсилення; R_D – коефіцієнт відбивання дзеркал.

Потужність, затрачена на збудження

$$P_{збуд} = V_{прик} I,$$

де $V_{прик}$ – прикладений потенціал; I – струм.

Коефіцієнт корисної дії

$$h = \frac{P}{P_{збуд}},$$

де P – вихідна потужність.

17.3. Фоторезистори

Напівпровідникові приймачі оптичного випромінювання, у яких під дією потоку випромінювання внаслідок внутрішнього фотоефекту змінюється опір, називаються фоторезисторами.

17.3.1. Принцип дії фоторезисторів. Принцип дії фоторезисторів ґрунтується на внутрішньому фотоефекті. Випромінювання, що падає на напівпровідник, частково поглинається в його об'ємі, взаємодіючи з атомами кристалічної ґратки або домішок. Поглинання фоторезистором фотонів супроводжується збільшенням провідності, приріст якої називають фотопровідністю. Для підвищення фотопровідності необхідно вибрати матеріал з якомога більшими рухливістю та часом життя носіїв фотона. Якщо до фоторезистора під'єднати джерело напруги, то відповідно до зміни провідності фотоприймача під час освітлення його через контакти із зовнішнього кола проходитиме струм, що перевищує кількість початково утворених фотонів. Підсилюється фотострум.

Фоторезистори – це плівки або пластинки фоточутливого напівпровідникового матеріалу, що мають два невідпрямляючі контакти для включення їх в електричне коло. Фоторезистори неполярні, вони

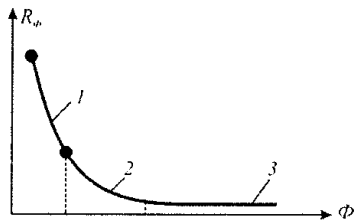


Рис. 17.21. Зміна опору чутливого шару фоторезистора зі зміною падаючого випромінювання:

- 1) $R_\phi = R_T - k\Phi$;
- 2) $R_\phi = R_T + (B_c/\Phi)$;
- 3) $R_\phi = \text{const}$.

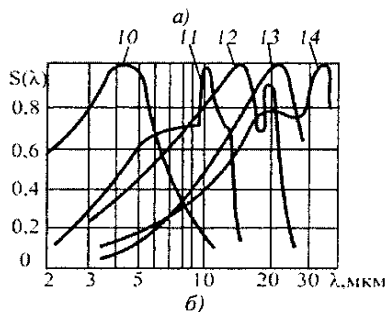
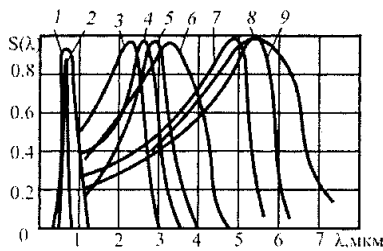


Рис. 17.22. Відносна спектральна чутливість фоторезисторів на основі власної (а) та домішкової (б) фотопровідностей: 1 – CdS; 2 – CdSe; 3 – PbS; 4 – PbS (195 K); 5 – PbS (77 K); 6 – PbSe (295 K); 7 – InSb (77 K); 8 – PbSe (77 K); 9 – InSb (295 K); 10 – Ge:Au (60 K); 11 – Ge:Hg (27K); 12 – Ge:Cd (23 K); 13 – Ge:Cu (15 K); 14 – Ge:Xn (4K).

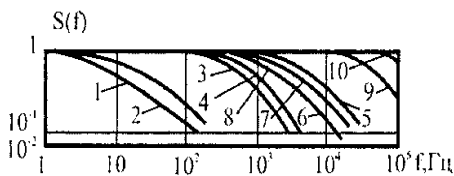


Рис. 17.23. Частотні характеристики фоторезисторів: 1 – CdS; 2 – CdSe; 3 – PbS (77 K); 4 – PbS (195 K); 5 – PbSe (295 K); 6 – PbS (295 K); 7 – Ge.Au (K); 8 – PbSe (77 K); 9 – PbS (295 K); 10 – InSb (77 K).

однаково проводять струм у будь-якому напрямку, тому їх можна живити постійним та змінним струмом. Фоторезистори створюють на базі власного та домішкового (для ІЧ-області спектра) поглинання. Найбільше поширення отримали неохолоджувані фоторезистори на базі напівпровідників без домішок з власним поглинанням.

Коефіцієнт підсилення, як і фотопровідність, пропорційний до рухливості та часу життя носіїв. Значення коефіцієнта підсилення також залежить від напруги та досягає $10^3 - 10^5$.

17.3.2. Основні параметри фоторезисторів. Властивості фоторезисторів оцінюються загальними для всіх приймачів параметрами і характеристиками.

Окрім загальноприйнятих параметрів, **чутливість** фоторезистора (виражену в 1/Вт або 1/лм) інколи характеризують **відносною зміною його опору під дією випромінювання**, яке не залежить від схеми вмикання:

$$S_R = \frac{\Delta R_\phi}{R_\phi \Delta \Phi}, \quad (17.1)$$

де R_ϕ – опір фоторезистора при потоці Φ ; ΔR_ϕ – зміна його опору при прирості потоку на $\Delta \Phi$.

На рис. 17.21 наведена залежність $R_\phi = f(\Phi)$ яку можна апроксимувати на різних ділянках 1,2,3 залежностями, наведеними на рисунку (k_c , B_c – коефіцієнти, що характеризують крутизну характеристики; R_T – темповий опір фоторезистора).

Спектральна чутливість фоторезистора залежить від матеріалу чутливого шару і наявності охолодження. Вона охоплює широкий спектральний діапазон 0,4...40 мкм (рис. 17.22).

Абсолютна спектральна чутливість неохолоджуваних фоторезисторів зі зниженням температури підвищується, оскільки зменшується теплова генерація носіїв струму.

Часова роздільна здатність фоторезисторів визначається їхньою сталою часу τ , зв'язаною з часом життя носіїв у напівпровідниковому чутливому шарі. В найпростішому випадку її можна вважати рівною часу життя носіїв.

Частотна характеристика чутливості фоторезистора залежить від його сталої часу і проявляється як зменшення його інтегральної чутливості з ростом частоти модуляції потоку випромінювання f (рис. 17.23).

Чутливість фоторезисторів на основі CdS, CdSe не перевищує 30% від максимальної на частоті модуляції в 1000 Гц. Сірчисто-свинцеві фоторезистори характеризуються такою чутливістю на частоті приблизно 10 кГц. Ще ширшою є частотна характеристика у селенисто-свинцевих фоторезисторів.

Граничну частоту модуляції потоку випромінювання часто вибирають зі співвідношення

$$f_{cp} \cong 1/3t.$$

Інерційність фоторезисторів залежить від ряду факторів. Розглянемо вплив цих факторів на власні сталі часу для наростання (τ_n) та спадання (τ_{cn}) фотоструму. При $t = 20$ °С, освітленості в 200 лк та короткозамкненому вмиканні фоторезистора $\tau_n > \tau_{cn}$. Зі зростанням освітленості сталі часу зменшуються, причому швидше змінюється τ_n . В середньому величини сталих змінюються на порядок в інтервалі двох декад освітленості.

Залежність сталої часу від провідності фоторезистора G_ϕ апроксимується виразом

$$t = BG_\phi^{-m},$$

де B і m – сталі, величини яких визначаються типом фоторезистора.

Температура довкілля також впливає на інерційність фоторезисторів. Сталі часу максимальні на нижній межі робочої області і зменшуються у випадку зменшення температури. Якщо $t = 20$ °С, значення τ_n , τ_{cn} для граничних температур можуть відрізнитися вдвічі-втричі.

Зі збільшенням опору навантаження фоторезисторів з деякого значення помітне значне зростання сталої часу спадання фотоструму та зменшення сталої часу наростання. Для CdS, CdSe фоторезисторів це навантаження становить 5-10 кОм.

Витримка фоторезисторів у темряві протягом 1 години і більше призводить до зростання сталої часу в 3-4 рази. Значення сталої τ_c практично не змінюється.

Додаткове підсвічування помітно впливає на інерційність фоторезисторів. Найбільше вона проявляється під час роботи зі слабкими освітленостями. Наприклад, додаткове підсвічування в 20 лк зменшує сталу t_n в середньому в 4 рази з робочою освітленістю 20 лк, вдвічі при $E = 200$ лк і в 1,1 рази при $E = 2000$ лк. Стала часу τ_{cn} менше залежить від фонового підсвічування.

Порогова чутливість фоторезисторів, тобто мінімальний променистий потік, визначається рівнем власних шумів. Джерелами шумів можуть бути фоточутливий елемент та площа контакту з електродами. Рівень шуму прямо пропорційний до прикладеної напруги за винятком області, близької до граничної напруги, де на характеристиці з'являється ділянка насичення. Спектр шуму має розподіл, який у першому наближенні визначається як $1/f$. На частотах, вищих, ніж 400 Гц, рівень шуму CdS, CdSe-фоторезисторів обернено пропорційний до квадрата частоти. Порогова чутливість плівкових сірчисто-кадмієвих фоторезисторів досягає 10^{-10} лм.

Енергетична (люкс-амперна) характеристика – залежність фотоструму від потоку випромінювання, що

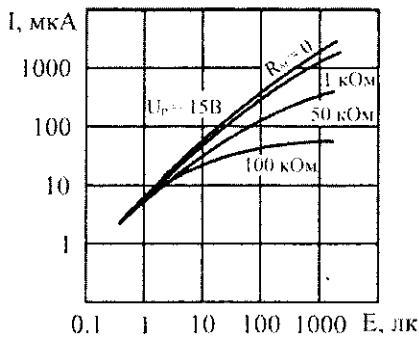


Рис. 17.24. Енергетичні характеристики фоторезисторів.

падає на фотоприймач. Залежність прийнято апроксимувати виразом:

$$I_{\phi} = c_m U E^{\zeta},$$

де E – освітленість; U – напруга; c_m – стала, що визначається властивостями матеріалу; ζ – коефіцієнт нелінійності, значення якого для більшості фоторезисторів знаходиться в інтервалі $[0,5...1,0]$.

Загалом характеристики нелінійні, хоча нерідко містять лінійну ділянку в межах однієї декади освітленості. Нахил характеристик і, отже, інтегральна чутливість максимальні в області слабких освітленостей. Форма кривих мало змінюється зі зростанням прикладеної до фоторезистора напруги.

Енергетичні характеристики знімаються для короткозамкнутого вмикання фоторезистора. Якщо опір навантаження збільшується до 50 кОм і більше, відбувається перерозподіл напруги між фотоприймачем та навантаженням. Внаслідок цього виникає ділянка насичення люкс-амперної характеристики в області сильних освітленостей (рис. 17.24).

$$I_{\phi} = W U^{\gamma},$$

де $W = c_m E^{\zeta} \cdot \gamma$ – коефіцієнт.

У більшості фоторезисторів вольт-амперні характеристики лінійні ($\gamma = 1$) до граничнодопустимих значень. Порушення лінійної залежності може відбуватися при малих напругах (значно менших від робочих) або при великих напругах, особливо у разі сильного засвічення. Робочу напругу фоторезистора встановлюють нижчою від пробивного значення з урахуванням рівня освітленості та температури довкілля, щоб не перевищити допустимої потужності розсіювання.

Температурна характеристика темного струму фоторезистора являє собою залежність темного струму від температури $I_T(T)$. Робоча область температур для більшості фоторезисторів обмежена значеннями -60 і $+70^{\circ}\text{C}$. В межах цієї області струм фоторезисторів зі зростанням температури збільшується, що взагалі характерно для напівпровідників. Абсолютна зміна темного струму становить одиниці мікроампер і не має помітного впливу на загальний струм фоторезистора.

Вольт-амперна характеристика $I(U)$ – залежність електричного струму від напруги, прикладеної до фотоприймача для фіксованого потоку випромінювання:

Температурна характеристика опору фотоприймача – залежність опору від температури $R(T)$. Зі зростанням температури опір освітленого фоторезистора в робочій області зростає. Відповідно загальний струм зменшується.

Температурна залежність фотоструму (чутливості) фоторезистора практично не відрізняється від залежності загального струму, оскільки температурна нестабільність темного струму, який є частиною загального струму,

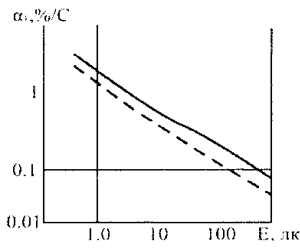


Рис. 17.25. Залежність температурного коефіцієнта фотоструму від освітленості: — область додатних температур; — область температур від $+20^{\circ}\text{C}$ до -60°C .

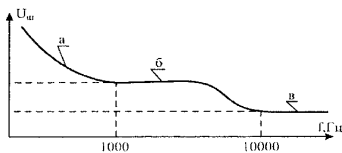


Рис. 17.26. Спектр потужності шуму фоторезисторів: а) — струмовий шум; б) — генераційно-рекомбінаційний шум; в) — тепловий та радіаційний шум.

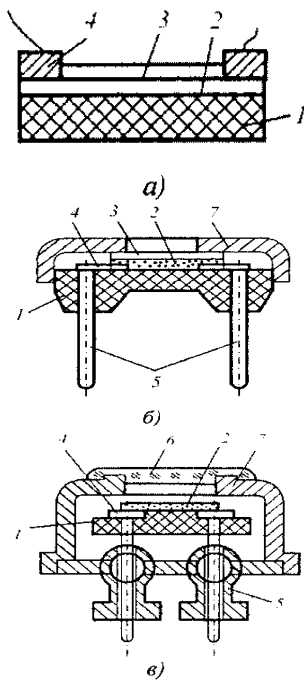


Рис. 12.27. Конструкція фоторезисторів: а) безкорпусні; б) в пластмасовому корпусі; в) в герметизованому корпусі: 1 — ізолююча підкладка; 2 — фоточутливий елемент; 3 — захисний шар лаку; 4 — контакти; 5 — виводи; 6 — скло; 7 — корпус.

має незначний вплив.

Кількісно вплив температури оцінюють за температурним коефіцієнтом фотоструму α_T . Для всіх типів фоторезисторів він від'ємний. У більшості фоторезисторів температурний коефіцієнт фотоструму зростає зі зменшенням освітленості. У межах двох декад величина α_T змінюється на порядок і більше (рис. 17.25).

Шуми в фоторезисторах. У фоторезисторах спостерігається тепловий, генераційно-рекомбінаційний, струмовий ($1/f$ -шум), радіаційний і фотонний шуми. Вони мають спектр шуму, наведений на рис. 17.26.

Електроди у фоторезисторах з найменшим рівнем шуму виготовляються із золота та індію.

17.3.3. Конструкція фоторезисторів. Конструктивно фоторезистори розміщуються в пластмасових або металоскляних герметичних корпусах. Деякі типи фоторезисторів випускаються без корпусів. У них чутливий елемент наклеюється на скляний підшар. Деякі модифікації фоторезисторів мають три електроди, що дає змогу використовувати диференційне вмикання.

У фоторезисторах, призначених для роботи у відбитому світлі, в центрі підшару робиться отвір. Чутливий елемент покритий прозорим шаром лаку. Фоторезистори з пластмасовими корпусами не можна використовувати в умовах підвищеної вологості.

Фоторезистори, в герметичних корпусах можна використовувати в умовах підвищеної вологості, рідких непровідних середовищ.

Фоторезистор складається з тонкого шару фоточутливого напівпровідникового матеріалу з електродами у вигляді плівок, стійких до корозії, нанесених випаровуванням золота, платини або срібла у вакуумі (рис. 17.27). Конструкція охолоджуваних фоторезисторів наведена на рис. 17.28. Як холодоагенти використовуються: фреон (263...145 К), тверда вуглекислота (195 К), рідкий азот (77 К), рідкий неон (27 К), рідкий гелій (4,2 К).

Вологість довкілля впливає на негерметизовані фоторезистори. Відносна вологість 85% призводить до втрати чутливості на 10-15%. Втрати збільшуються до 90% і стають незворотними при тривалій експлуатації в умовах вологості, близької до 100 %.

При розгляді стабільності параметрів фоторезисторів необхідно виділити два варіанти використання: тривала неперервна експлуатація та чергування експлуатації з відпочинком. У першому випадку протягом перших 50-100 годин експлуатації загальний струм зменшується, відповідно і чутливість фотоприймача, а потім параметри стабілізуються.

В окремих типах фоторезисторів загальний струм, навпаки, збільшується. Зменшення чутливості прямо пропорційне до виділеної потужності розсіювання і може

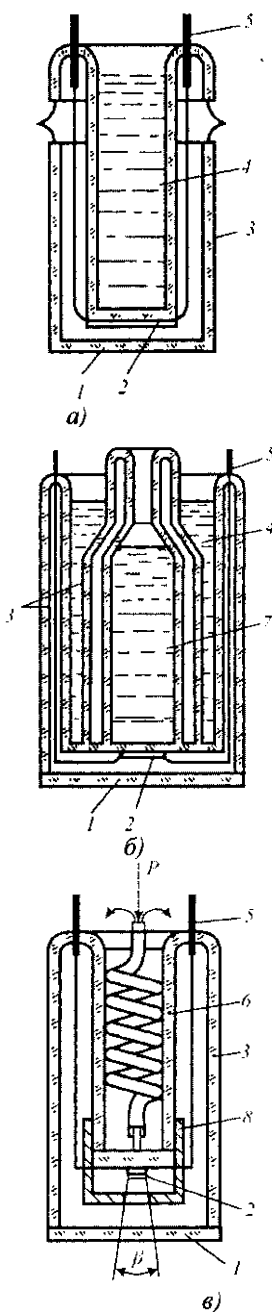


Рис. 17.28. Конструкція пристроїв для охолодження фоторезисторів: а) рідким азотом; б) рідким гелієм; в) стисненим азотом: 1 – вхідне вікно (фільтр); 2 – фоточутливий елемент; 3 – посудина Дьюара; 4 – рідкий азот; 5 – електричний вивід; 6 – змійовик; 7 – рідкий гелій; 8 – діафрагма, що охолоджується.

сягати 20-30%, що зв'язано, насамперед, з процесом втоми. Поряд з втомою відбуваються повільні незворотні зміни параметрів фоторезисторів, викликані старінням.

Втома викликається розігріванням фоточутливого елемента струмом, що проходить через нього. Це зворотне явище. В темряві повне відновлення властивостей елемента відбувається до трьох діб.

Якщо робота фоторезистора чергується з його вимкненням, затемненням, то нестабільність параметрів визначається втомою, яка, як вказувалось, залежить від виділеної потужності розсіювання. Чутливість змінюється незначною мірою, якщо потужність розсіювання мала. Навпаки, функціонування з максимальною потужністю розсіювання освітленого протягом кількох годин фоторезистора призводить до зміни загального струму до 15%. нестабільність значення загального струму збільшується, якщо в неробочому стані фотоприймач зберігається в темряві.

Для стабілізації параметрів фоторезистори нерідко піддають штучному старінню при максимальній або вищій від максимально допустимої потужності розсіювання. Необхідні для режиму старіння значення прикладеної до цього типу фоторезистора напруги та освітленості підбираються експериментально, процес триває приблизно 100 годин. У результаті штучного старіння чутливість фоторезистора практично стабілізується, однак величина її значно зменшується.

У конструкціях з фоторезисторами слід прагнути до повного засвічення чутливої площі. У випадку часткового засвічення параметри фоторезистора відрізняються від паспортних і залежать від положення та розміру освітленої поверхні. Під час освітлення вузькою світловою плямою, що переміщується поверхнею чутливого елемента вздовж електродів, захоплюючи їх, загальний струм пропорційний до площі освітленої поверхні і нерідко може змінювати свою величину внаслідок нерівномірної чутливості поверхні елемента.

Якщо вузька світлова пляма пересувається перпендикулярно до електродів, опір фотоприймача являє собою послідовне з'єднання опорів освітленої та неосвітленої частин робочої поверхні, останній з яких змінюється з пересуванням плями. Змінюється загальний струм і швидкодія фоторезистора, причому максимум загального струму спостерігається, коли світлова пляма проходить центральну частину площі. Переміщення широкої світлової плями в напрямку, перпендикулярному до електродів, дає зростання загального струму, залежність якого близька до квадратичної.

17.3.4. Області використання фоторезисторів. Сферу застосування фоторезисторів визначають такі позитивні якості: найбільша серед напівпровідникових фотоприймачів чутливість, наближена до чутливості

фотоелектронних помножувачів, низький пороговий струм, значно вищі, ніж у фотодіодів, потужності розсіювання, лінійна вольт-амперна характеристика, можливість передавання двополярних сигналів. Основний недолік – інерційність фоторезисторів. За способом одержання вимірюваної інформації перетворювачі поділяються на такі, що контролюють параметри джерела випромінювання та параметри оптичного сигналу. За формою вихідного сигналу – на ті, що дають неперервну інформацію, та порогові перетворювачі. Фотометри, люксметри, пірметри, експонетри фото- та кіноапаратури є пристроями, що вимірюють потік випромінювання самого джерела. Фотореле можуть приймати інформацію як від джерела випромінювання, так і від оптичного каналу, але, на відміну від вищезазначених пристроїв, виробляють вихідний сигнал у дискретній формі. У системах автоматики існують багато різновидів фотореле: пристрої вмикання освітлення вулиць, тунелів, приміщень, світлових вогнів аеропортів, сигналізатори вогню та диму, схеми стеження за положенням об'єкта, фотореле різних промислових автоматів тощо.

Велику групу становлять прилади для контролю оптичних параметрів середовища, яким поширюється променистий потік. В результаті проходження через прозору речовину (газ, рідина, тверде тіло) променистий потік залежно від свого спектрального складу, властивостей речовини, довжини шляху зазнає змін. За величиною поглинутого, заломленого або відбитого потоку відповідної області спектра, за зміною його поляризації визначають концентрацію контрольованої речовини або її компонент. Цей принцип покладено в основу приладів, що вимірюють, наприклад, концентрацію суспензій у спорудах для очищення стічних вод, концентрацію луку під час варіння сульфітної целюлози, наявність нітробензолу в аніліні, вміст сірководню в газових сумішах, що контролюють якість авіаційного пального, заповищеність повітря, наявність диму чи оптично активних речовин. У пристроях використані фоторезистори видимої та ближньої інфрачервоної областей спектра. Численними та різноманітними є перетворювачі з фоторезисторами, призначені для контролю місцезнаходження або геометричних параметрів об'єкта. Залежно від конструкції пристрою світло, яке падає на фоторезистори, перекривається самим контрольованим об'єктом або зв'язаним з ним елементом. Типовими представниками таких перетворювачів є пристрої для визначення кількості рухомих предметів з продуктивністю до 50-80 тис. відрахувань за годину, вимірювання кількості обертів двигуна за допомогою диска з прорізом або прапорця, зв'язаних механічно з валом двигуна та оптично з

фотоприймачем, пристрої охорони приміщень та територій, перетворювачі, що контролюють температуру в заданому діапазоні, тиск газу, рівень рідини, сипких речовин, дозувальні автомати, прилад для вимірювання амплітуди та частоти низькочастотної вібрації тощо.

Фоторезистори на основі сульфїду свинцю та селенїду свинцю використовуються для дистанційного вимірювання температури нагрітих до 100-700 °С об'єктів, які безперервно виділяють променистий потік інфрачервоного діапазону. Такі фотоприймачі функціонують у пірометрах, пристроях виявлення та наведення літальних апаратів, приладах контролю та керування технологічними процесами (наприклад, управління прокатними станами, вугільними, газовими котлами), захисних сигналізаторах. Характерними рисами фотоприймачів є висока виявлювальна здатність та мала інерційність (залежно від глибини охолодження стала часу 0,1-5,0 мкс). Фоторезистори служать для детектування слабких швидкозмінних променистих потоків і успішно використовуються в спектрометрії. Застосування InSb, HgCdTe-фоторезисторів обмежується в діапазоні низьких температур.

Фоторезистори на основі германію, легованого домішками міді або ртуті, працюють у спектральному діапазоні від 1 до 25 мкм. Робоча температура становить 4,2 або 27 К.

17.4. Фотодіоди

Фотодіодами називають напівпровідникові прилади, що ґрунтовані на внутрішньому фотоефекті і використовують односторонню провідність р-п-переходу, під час освітлення яких з'являється ЕРС (фотогальванічний режим) або (за наявності живлення) змінюється величина зворотного струму (фотодіодний режим).

17.4.1. Класифікація фотодіодів. Один з можливих варіантів класифікації фотодіодів наведений на рис. 17.29.

За **принципом дії** розрізняють дифузійні та дрейфові фото діоди. **Дифузійні фотодіоди** являють собою структуру з однорідним розподілом домішки в р-п-областях. В результаті опромінення фотоприймача потоком спектрального складу, що відповідає області власного поглинання, в його об'ємі генеруються електронно-діркові пари. У **дрейфового фотодіодів** в базі присутнє внутрішнє електричне поле, що прискорює рух носіїв до р-п-переходу.

Фотодіоди можуть працювати у **двох режимах** роботи: фотогальванічному та фотодіодному. Режим роботи фото діода без зовнішнього зміщення називається **фотогальванічним**, з від'ємним зміщенням – **фотодіодним**. У фотогальванічному режимі фотодіод безпосередньо перетворює енергію світла в електричну енергію, у фотодіодному режимі до нього прикладають

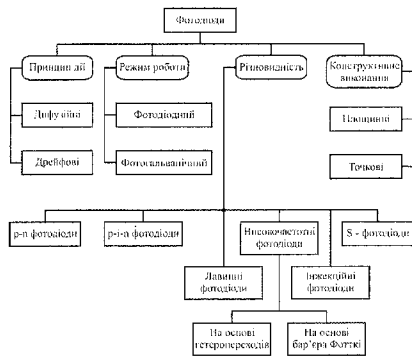


Рис. 17.29. Класифікація фотодіодів.

зворотну напругу.

Відомі такі основні різновидності фотодіодів: р-п-фотодіоди, р-і-п-фотодіоди, лавинні фотодіоди; високочастотні фотодіоди (на основі гетеропереходів та бар'єрів Шотткі), інжекційні фотодіоди; s-фотодіоди.

Р-п-фотодіоди функціонують на основі односторонньої провідності (вентильного фотоефекту), що виникає внаслідок освітлення однієї або обох областей р-п-переходу напівпровідникової структури.

Р-і-п-фотодіод являє собою тонкі низькоомні р-п-області, між якими вміщено протяжний і-шар, у якому поглинається до 90% падаючої потужності. Функціонує р-і-п-фотодіод при зворотному змищенні.

Принцип дії лавинних фотодіодів ґрунтується на явищі лавинного множення фотоносіїв як у поверхневих бар'єрах, так і в р-п-переходах.

Високочастотні фотодіоди використовуються для зменшення сталої часу реалізуються на основі гетеропереходів та бар'єрів Шотткі (метал – напівпровідник).

У високочастотних **фотодіодах з гетеропереходом** матеріал освітленого напівпровідника підбирається так, щоб реєстроване випромінювання проходило через нього і поглиналось у самому р-п-переході і поблизу за ним.

У **фотодіодах на основі бар'єра Шотткі** реєстрований потік випромінювання поглинається в р-п-переході, безпосередньо прилягаючому до зовнішнього напівпрозорого матеріалу електроду, через який він освітлюється.

Інжекційні фотодіоди належать до фотоприймачів з внутрішнім підсиленням. Структуру діода характеризує довга базова область. Довжина бази з високоомного напівпровідника у кілька разів перевищує довжину дифузії неосновних носіїв струму. Фотодіод працює при прямому змищенні.

S-фотодіоди за структурою та принципом дії аналогічні до інжекційних фотодіодів. Фотони світла створюють носії струму в довгій базовій області, що стимулює ріст інжекційного струму та призводить до помноження фотоструму.

Здебільшого фотодіоди мають планарну конструкцію (світло падає перпендикулярно до р-п-переходу). За **конструктивним виконанням** розрізняють **площинні** та **точкові** фотодіоди. Найпоширенішими є площинні кремнієві фотоприймачі. Нарис. 17.30 показано структури кремнієвих фотодіодів.

Антивідбивальна поверхня дає змогу зменшувати втрати випромінювання внаслідок відбивання на величину до 30%. Поверхня виконується у вигляді тонких шарів двоокису кремнію або халькогенідного скла.

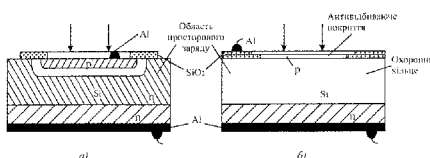


Рис. 17.30. Будова кремнієвих фотодіодів: а) дрейфовий; б) р-типу.

17.4.2. Р-п-фотодіоди. Принцип дії. Як було сказано вище, за принципом дії розрізняють дифузійні та дрейфові фотодіоди. *Дифузійні фотодіоди* являють собою структуру з однорідним розподілом домішки в р-п-областях. Під час опромінення фотоприймача потоком спектрального складу, що відповідає області власного поглинання, в його об'ємі генеруються електронно-діркові пари. Так пари на віддалі, яка дорівнює або менша від дифузійної довжини L від р-п-переходу, розділяються полем переходу. Неосновні носії переходять область переходу, а основні – залишаються. В результаті утворюється додаткове електричне поле, скероване протилежно до поля р-п-переходу, що призводить до зниження висоти потенціального бар'єра. Різниця потенціалів, що виникає внаслідок опромінення фотодіода, залежить від інтенсивності потоку опромінення та опору зовнішнього кола.

У випадку розімкненого кола всі генеровані носії нагромаджуються біля переходу, знижуючи потенціальний бар'єр на максимальну величину U_{np} . Якщо діод замкнено накоротко, то всі розділені носії протікають у зовнішнє коло, утворюючи найбільше можливе значення струму короткого замикання $I_{кз}$.

Не всі генеровані під дією фотонів носії розділяються потенціальним бар'єром р-п-переходу. Для оцінки ефективності розділення носіїв струму використовують *коефіцієнт збирання* носіїв Q :

$$Q = \frac{I_{кз}}{qN_{\phi}},$$

де N_{ϕ} – кількість поглинутих напівпровідником фотонів; q – заряд електрона.

У дифузійному фотодіоді важко досягти високого коефіцієнта збирання. Так, у короткохвильовій області спектрального діапазону діода частина електронно-діркових пар утворюється у приповерхневому шарі бази на значній віддалі l від р-п-переходу ($l < L$). Ці носії рекомбінують. У довгохвильовій області діапазону випромінювання проникає всередину напівпровідника через перехід, слабо поглинаючись.

Коефіцієнт збирання вищий у діодів з базою р-типу, оскільки коефіцієнт дифузії електронів перевищує дифузійну сталу дірок.

У *дрейфового фотодіода* в базі присутнє внутрішнє електричне поле, що прискорює рух носіїв до р-п-переходу. Поле дорівнює нулю біля поверхні бази і досягає максимального значення біля р-п-переходу. Напруженість поля E пропорційна до градієнта концентрації домішки в точці, що розглядається. Необхідний профіль розподілу домішки в базі створюється під час виготовлення діода. Гауссовий профіль розподілу одержують іонно-променевим та дифузійним методами легування. Ефективне розділення

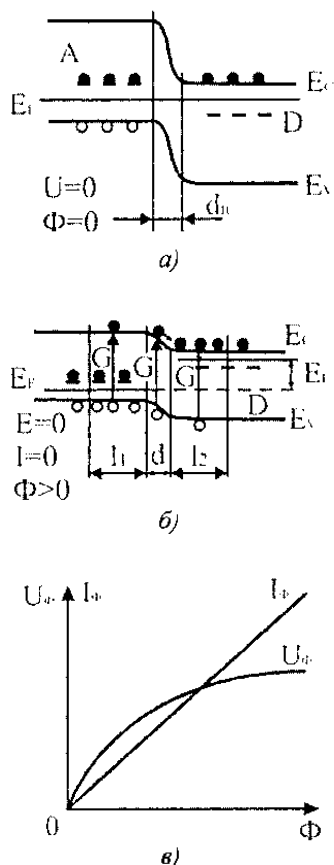


Рис. 17.31. р-п-перехід у стані рівноваги (а) і при освітленні за відсутності зовнішньої напруги (б), залежність фотоструму I_ϕ і фото-ЕРС U_ϕ від світлового потоку Φ (в).

генерованих носіїв досягається, якщо ширина області просторового заряду більша від середньої глибини проникнення в напівпровідник. Для р-п-фотодіода в спектральному діапазоні 0,45-0,6 мкм товщина області просторового заряду повинна становити 1-3 мкм. Коефіцієнт збирання носіїв дрейфового фотодіода вищий, ніж дифузійного.

У фотодіодах світлочутливим елементом є перехідна область між матеріалами з електронною і дірковою провідністю. На рис. 17.31, а зображено енергетичну схему р-п-переходу у рівновазі. Рівень Фермі E_F у всіх частинах системи має одне і те саме положення. Область з дірковою провідністю містить акцепторні домішки A , на які електрони теплом переводяться з валентної зони, область з електронною провідністю – донорні домішки D які дали свої електрони в зону провідності. В області завширшки d_0 присутнє контактне поле, що утворилось в результаті переходу деякої кількості електронів з електронного напівпровідника в дірковий. Це поле утруднює потрапляння електронів і дірок у шар d_0 , який збіднений носіями, і визначає струм через діод.

В результаті освітлення р-п-переходу світлом з енергією фотонів, більшою від ширини забороненої зони $E_g = E_c - E_v$, по обидва боки від переходу і в самому переході виникають пари, електрон-дірка, в області d_0 вони розділяються електричним полем, причому електрони рухаються в область з електронною провідністю, а дірки – в область з дірковою провідністю (рис. 17.31, б). Окрім того, в перехід потрапляють електрони, створені світлом у шарі l_1 ліворуч від переходу, оскільки після звільнення електрони протягом часу життя τ_e встигають проходити середню віддаль l_1 (дифузійна довжина). Те саме стосується дірок у шарі l_2 праворуч від переходу. При $\tau = 1 \cdot 10^{-6}$ с і коефіцієнті дифузії носіїв $D_n = 25 \text{ см}^2 \cdot \text{с}^{-1}$ довжина $l = 5 \cdot 10^{-3}$ см ($l = \sqrt{D_n \cdot \tau_e}$).

Якщо на перехід не подана зовнішня напруга і коло розімкнуте (рис. 17.31, б), то освітлювання призводить до накопичення фотоелектронів у п-області і дірок у р-області. У результаті утворюється різниця потенціалів U_ϕ , тобто з'являється фото-ЕРС. Якщо зовнішнє коло замкнуте, то виникає фотострум. У таких умовах діод працює як фотоелемент.

Напруга U_ϕ зміщує перехід у прямому напрямку, понижує висоту бар'єра для електронів і дірок, що полегшує перетікання швидких електронів у р-область. У міру збільшення Φ зростання U_ϕ сповільнюється (рис. 17.31, в). Вольт-амперну характеристику р-п-переходу при освітлюванні можна записати так

$$I = I_n \cdot \left[\exp\left(\frac{qU}{kT}\right) - 1 \right] - I_\phi, \quad (17.2)$$

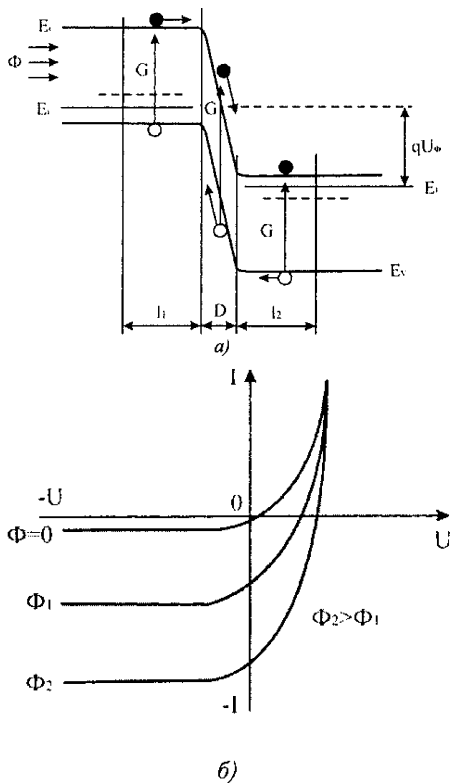


Рис. 17.32. р-п-перехід, увімкнений у зазірному напрямку (а), і вольт-амперні характеристики фотодіода при різних світлових потоках Φ (б).

де I_n – струм насичення в темноті; I_ϕ – фотострум, що проходить через перехід; U – зовнішня напруга на переході.

Якщо у зовнішньому колі $I = 0$, а $U = U_\phi$, то з (17.2) отримаємо

$$U_\phi = \frac{kT}{q} \ln \left(\frac{I_\phi}{I_n} + 1 \right). \quad (17.3)$$

Оскільки $I_\phi \sim \Phi$, то з (17.3) маємо нелінійну залежність U_ϕ ("напруга неробочого ходу") від Φ . Якщо фотоелемент входить у зовнішнє коло з малим опором, то фотоелектрони не накопичуються в п-області і $U_\phi = 0$. Оскільки зовнішня напруга також відсутня $I = I_\phi$, тобто в цьому випадку струм (так званий **струм короткого замикання**) $I_\phi \sim \Phi$ (рис. 17.31, в).

Якщо р-п-перехід увімкнений в зазірному напрямку (на р-область подати мінус джерела напруги), то практично вся напруга буде падати на збіднену область завширшки d і схема енергетичних зон набере вигляд, наведений на рис. 17.32, а.

Тепер розподіл електронів і дірок, що виникає під час освітлення, виконується сильнішим полем, а ширина області високого опору зростає ($d : \sqrt{U}$). Це збільшує ту частину кристала, з якої електрони і дірки швидко відносяться полем. Якщо дрейфова швидкість електронів $v_d = 1 \cdot 10^6$ см/с, а $d = 10^{-4}$ см, то час прольоту електронами області поля $\tau_i = 7 \cdot 10^{-10}$ с. З такою самою швидкістю ця складова фотоструму буде прямувати за зміною інтенсивності світла. Але через перехід проходять і електрони (дірки), народжені світлом у шарах l_1 і l_2 , які досягли переходу дифузиею. Відповідні часи значно більші, і саме вони можуть визначати інерційність фотодіода. Для послаблення цієї складової фотоструму р-п-перехід необхідно формувати біля найбільш освітленої поверхні (зменшуючи, наприклад, товщину x_1 шару р-типу).

Оскільки більша частина світла поглинається на глибині $x^* = 1/\alpha$ (де α – коефіцієнт поглинання), то при $x_1 \approx 0$ і $d = x^*$ роль світла, що поглинається в шарі l_2 , і дірок, що приходять, також послаблена (рис. 17.32, а). Щоб повністю виключити поглинання світла праворуч від переходу, можна збільшити d до $2x^*$. Цього досягають у спеціальних р- і п-переходах, в яких між шарами з високою концентрацією акцепторів і донорів розміщують шар з власною провідністю (i).

Графіки вольт-амперних залежностей р-п-переходу при освітлюванні зображено на рис. 17.32, б. Під час увімкнення переходу в зазірному напрямку ($U < 0$) і при $qU \gg kT$ з (17.2) випливає, що $I = -(I_n + I_\phi)$. У міру підвищення Φ фотострум I_ϕ зростає і може значно перевищити темновий струм I_n .

Якщо весь світловий потік Φ , що падає на фотодіод,

поглинається в шарі товщиною $x_1 + d + l_2$ (тобто ця товщина більша, ніж $2x$), то всі електрони і дірки, що виникли, візьмуть участь у створенні фотоструму і

$$I_{\phi} = q \cdot h_1 \cdot \Phi / (lm),$$

де η_1 – квантовий вихід генерації електронно-діркових пар.

У цьому випадку квантовий вихід фотодіода $\eta_d = \eta_1$. Якщо частина електронів і дірок народжується за межами шару $l_1 + d + l_2$, то $\eta_d = \eta_1 \cdot K_c$, де K_c – коефіцієнт збирання фотоелектронів і дірок ($K_c < 1$). Відбивання світла від поверхні напівпровідника також зменшує квантовий вихід фотодіода. На відміну від фоторезисторів розглянуті фотодіоди не мають внутрішнього підсилення і їх загальна чутливість становить приблизно 20 мА/лм.

17.4.3. Р-і-n-фотодіоди. Структура та принцип дії збільшення ширини потенціального бар'єра можна отримати в р₊-і-n₊-структурах, що представляють собою власний напівпровідник і з великим питомим опором (в 106... 107 разів більш високим, ніж n- і р-області), обмежений з двох сторін сильно легованими шарами n- і р-типу. Такий діод має широку р-область із власного напівпровідника між n- і р-областями (рис. 17.33).

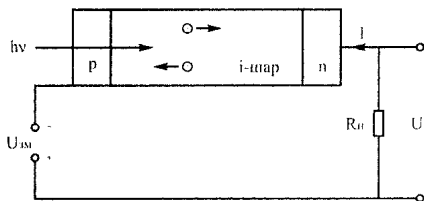


Рис. 17.33. Структура та схема ввімкнення р-і-n-фотодіода

Сильне рівномірне внутрішнє електричне поле електронно-діркового переходу зосереджено в і-області. Якщо товщина освітленого шару р- або n-типу невелика (0,5 мкм), падаюче випромінювання поглинається в і-шарі. Стала часу такого р-і-n-фотодіода визначається часом прольоту носіїв через перехід (при і-області 0,1 мм $\tau = 1,7...1,3 \cdot 10^{-9}$ с). При невеликій площі переходу (приблизно $2 \cdot 10^{-4}$ см²) і значній товщині і-шару ємність переходу невелика і час схемної релаксації рівний $10^{-11}... 10^{-12}$ с.

Область власної провідності не має вільних носіїв заряду, тому її опір досить великий і електричне поле в ній має велику напруженість. Оскільки область власної провідності дуже широка, то існує висока ймовірність поглинання фотонів в і-області, а не в тонких р- або n-областях, що збільшує квантову ефективність і швидкість відгуку в порівнянні з р-n-фотодіодом.

На р і n-області наносяться контакти. Ці фотодіоди отримали назву **р-і-n-фотодіодів**. У таких діодах після прикладання до контактів напруги зміщення

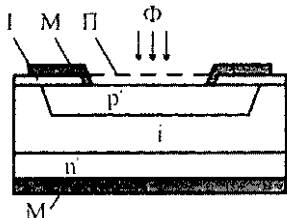


Рис. 17.34. Конструкція р-і-n-фотодіода: М – металеві контакти; I – ізолятор (SiO₂); П – просвітлювальне покриття.

У кремнієвому р-і-n-діоді, який є одним із найпоширеніших фотодетекторів (рис. 17.34), товщина і-області становить приблизно 50 мкм, а шару р⁺ – всього 3 мкм. У випадку освітлення такого діода світлом з довжиною хвилі $\lambda = 0,9$ мкм (від GaSte-випромінювача) $x^* = 30$ мкм і близько 80% світла поглинається в і-шарі. Отже, практично всі фотоелектрони і дірки виникають в і-шарі, і швидкодія діода визначається часом τ_i їх прольоту через і-шар. Дрейфова швидкість електронів у кремнії спочатку зростає з підвищенням напруженості поля, а

далі досягає насичення.

В загальному випадку враховують ще одну складову, τ_{RC} , сталої часу, що пов'язана з опором R і ємністю C кола. При малому опорі нейтральних областей діода, а також зовнішнього кола, та при широкому переході маємо $\tau_{RC} < \tau_i$.

Вольт-амперна характеристика р-і-п-фотодіодів. Сім'я вольт-амперних характеристик кремнієвого р-і-п-фотодіода з струмовим відгуком 0,5 А/Вт показана на рис. 17.35.

17.4.4. Матеріали і конструкція р-і-п-фотодіодів. У таблиці 17.1 наведені параметри р-і-п-фотодіодів із найбільш поширених напівпровідникових матеріалів.

Таблиця 17.1

Типові параметри р-і-п-фотодіодів

| Матеріал | Область спектральної чутливості, мкм | Довжина хвилі максимального відгуку, мкм | Максимальний струмовий відгук, А/Вт |
|----------|--------------------------------------|------------------------------------------|-------------------------------------|
| Si | 0,3...1,1 | 0,8 | 0,5 |
| Ge | 0,5...1,8 | 1,55 | 0,7 |
| JnGaAs | 1,0.-1,7 | 1,7 | 1,1 |

Фотодіоди монтуються в корпус такими основними способами:

- кристал фотодіода встановлюють у стандартний корпус транзистора, аналогічний корпусу світлодіода. У металевий ковпачок вбудовується прозоре скляне вікно або лінза. Остання фокусує світло на активну площадку приймача. Лінза може збирати світло з торця волокна, діаметр серцевини якого більший, ніж діаметр фотодіода, що збільшує ефективність збору променів. У деяких випадках ковпачок усувають, щоб забезпечити доступ безпосередньо до діода;

- у корпус фотодіода вмонтовують гнучкий одноволоконний кабель із (або без) з'єднувачем на протилежному кінці;

- фотодіод розміщується всередині корпусу приймального модуля. При цьому можлива його інтеграція в єдину інтегральну схему або сполучення двох чіпів (діода і попереднього підсилювача) при гібридному виконанні приймального модуля.

17.4.5. Лавинні фотодіоди. Принцип дії та структура. Лавинні фотодіоди (ЛФД) дозволяють реалізувати внутрішнє підсилення фотоструму за рахунок

електричного пробію р-п-переходу і утворення лавинного процесу розмноження неосновних носіїв, що надходять у р-п-перехід. При достатньому зворотному електричному зміщенні р-п-переходу утворені в результаті освітлення неосновні носії прискорюються електричним полем з ґраткою кристала в області р-п-переходу, утворюючи додаткові електронно-діркові пари, які в свою чергу, створюють такий же процес. В

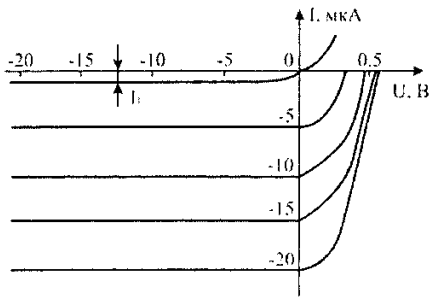


Рис. 17.35. Сім'я вольт-амперних характеристик кремнієвого р-і-п-фотодіода.

результаті фотострум I_{ϕ} підсилюється в M разів, де $M = I/I_{\phi}$ – коефіцієнт помноження лавинного фотодіода.

Лавинне множення відбувається так. Поглинений в збідненій області фотон створює вільну електронно-діркову пару (ЕДП). Високе електричне поле в збідненій області, змушує носії заряду рухатись із прискоренням, що збільшує їхню кінетичну енергію. Коли ці носії заряду зіштовхуються з нейтральними атомами, вони створюють додаткові ЕДП. Частина їх кінетичної енергії у цьому випадку використовується, щоб передати зв'язаним (з атомами) електронам енергію, достатню для подолання забороненої зони. Один прискорений заряд може створити кілька вторинних. Вторинні заряди також можуть прискорюватися і створювати ще більшу кількість ЕДП. Це явище називається лавинним множенням носіїв заряду.

Лавинні фотодіоди є напівпровідниковими аналогами вакуумних фотоелектронних перетворювачів з кращим відношенням сигнал/шум. Коефіцієнт помноження темного I_T і світлового I_{ϕ} струмів залежить від прикладеної напруги і змінюється по-різному. Величина коефіцієнта помноження M лавинного фотодіода залежить від опору, фонового засвічення і величини темного струму, і для кремнієвих лавинних фотодіодів рівна $10^4 \dots 10^6$, а для германієвих – $10^2 \dots 10^3$.

Коефіцієнт підсилення повного струму I зростає зі зростанням зворотної напруги, а коефіцієнт помноження M світлового потоку I_{ϕ} має максимум при напрузі пробою. У лавинних фотодіодах використовується або широкий р-n-перехід, або р-i-n-перехід, або поверхневий бар'єр Шотткі. Оскільки дірки мають великий коефіцієнт ударної іонізації, освітлювана область стає областю n-типу.

Лавинні фотодіоди можна розглядати як модифікацію р-i-n-фотодіодів. Використовуються однакові матеріали та області спектральної чутливості.

На рис. 16.36 показані структура та принцип роботи одного з варіантів ЛФД – із проникненням поля.

Шари p^+ та n^+ є високолегованими (низькоомними) областями, на яких відбувається дуже мале падіння напруги. Напівпровідникова π -область тільки злегка легована, тобто має майже власну провідність. Більшість фотонів поглинаються в цій області. Внаслідок великої напруги зворотного зміщення фотоелектрони рухаються до р-області, збідненої вільними носіями заряду. Збіднена область р-n⁺-переходу "проникає" у π -область. Найбільше падіння напруги відбувається в області р-n⁺-переходу, де велике електричне поле спричиняє лавинне множення.

У наведеному приладі множення починається електронами. Дірки, створені в n^+ -області, дрейфують до р⁺-електрода і не беруть участі в процесі множення. Структура, в якій процес множення починається носіями

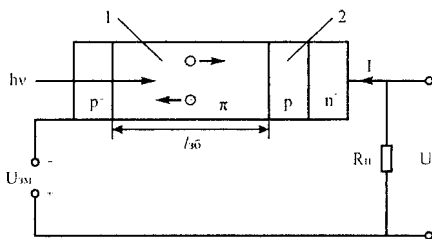


Рис. 16.36. Структура та принцип дії лавинного фотодіода з проникненням поля: 1 – генерація електронно-діркової пари; 2 – область множення; $l_{зб}$ – ширина збідненої області.

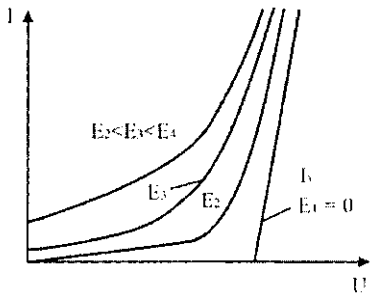


Рис. 17.37. Вольт-амперні характеристики лавинного фотодіода.

заряду одного типу, має кращі шумові показники.

Характеристики лавинних фотодіодів. Вольт-амперні характеристики лавинних фотодіодів наведені на рис. 17.37.

Робоча область характеристик розміщена зліва від вертикальної прямої, що перетинає вісь у точці $U_{проб}$. У цій області коефіцієнт підсилення збільшується від 1 до $M_{макс}$. Ще швидше зростають шуми лавинного фотодіода, внаслідок цього значно знижується порогова чутливість.

Квантовий вихід іонізації залежить від коефіцієнтів ударної іонізації (кількості пар, створених електроном на 1 см шляху) для електронів (α_i) і дірок (β_i), а також від ширини d області поля. При $\alpha_i = \beta_i$ та однорідному полі завширшки d маємо $N = \alpha_i \cdot d$, у випадку неоднорідного поля $N = \int_0^d a_i dx$, де N – кількість іонізацій, що припадає на

один електрон, який пройшов область поля. N можна подати у вигляді $N = a_1 \exp(-b_1/U)$, де a_1 , b_1 і c_1 – сталі. Значення N змінюється від 0 при малих U до 1 при напрузі пробою U_n , коли струм через діод різко зростає ($M \rightarrow \infty$). На практиці часто використовується степеневе подання залежності $N(U)$, а саме: $N = (U/U_0)^m$, де $m = 2 - 4$ для різних матеріалів.

Коефіцієнт підсилення ЛФД можна визначити за допомогою апроксимуючого виразу Міллера

$$M = \frac{n}{n_0} = \frac{1}{1 - (U_{зм} / U_{пр})^m},$$

де n та n_0 – кількість вторинних та первинних носіїв заряду відповідно; $U_{зм}$ – напруга зміщення на фотодіоді; $U_{пр}$ – зворотна пробійна напруга фотодіода для $M \rightarrow \infty$; m – експериментальний коефіцієнт, що набуває значення в діапазоні 1,5...6 залежно від матеріалу фотодіода та типу р-п-переходу.

Коефіцієнт підсилення загального струму зростає пропорційно до прикладеної напруги, а залежність підсилення фотоструму має максимум при $U_{зм} = U_{пр}$. Граничні значення коефіцієнта підсилення $M_{макс}$ для лавинних фотодіодів на основі Si та GaAs дорівнюють $10^3 - 10^4$, Ge – 10^2 . Ці значення тим вищі, чим нижчим є рівень темного струму та фонового засвічення, а оскільки темновий струм германієвого ЛФД вищий, ніж у кремнієвого, то відповідно проявляється й різниця в підсиленні. Коефіцієнт підсилення ЛФД залежить від температури, зменшуючись при її підвищенні. Це пояснюється тим, що середня довжина вільного пробігу носіїв між зіткненнями знижується зі зростанням температури. При цьому багато носіїв заряду втрачають можливість досягнення високих швидкостей, потрібних для створення вторинних носіїв. У прийнятному пристрої з ЛФД, що працює в широкому інтервалі температур, може знадобитися стабілізація температури

або компенсація зміни коефіцієнта лавинного підсилення шляхом автоматичного регулювання напруги зворотного зміщення.

Оскільки коефіцієнт підсилення (множення) M може досягати значень $1 \cdot 10^3$, лавинний діод придатний для реєстрації дуже слабких світлових сигналів. Водночас, через сильну залежність $M(U)$ використання лавинних діодів утруднюється необхідністю застосування високостабільної напруги.

Інтенсивність лавинних фотодіодів становить близько $1 \cdot 10^9$ с.

Внутрішнє підсилення в лавинному фотодіоді не пов'язане зі значним підвищенням інерційності. Лавинний пробій, як і всі польові процеси, відбувається швидко. Гранична частота досягає сотень мегагерц, а добуток смуги частот на коефіцієнт підсилення – 10^{11} Гц. Так, наприклад, для кремнієвих фотодіодів $M\Delta f$ становить 80 ГГц, для германієвих – 50 ГГц.

Струм, що генерується ЛФД із коефіцієнтом лавинного підсилення M

$$I = \frac{MhqP}{hv} = \frac{MhqI_p}{hc},$$

де η – квантова ефективність при $M=1$; q – заряд електрона; h – стала Планка; ν – частота; c – швидкість світла у вакуумі.

Струмова чутливість (відгук) ЛФД

$$S_1 = \frac{Mhe}{hv} = \frac{MhqI}{hc}$$

і становить 20...80 А/Вт в лавинному режимі роботи.

Як і у фотодіоді без множення, **швидкість відгуку** ЛФД обмежена часом прольоту носіїв заряду і RC-сталою часу. Час прольоту в ЛФД, що відповідає часу наростання, становить кілька десятих часток не. Час наростання, менший ніж 100 пс, досягнутий і для кремнієвих, і для германієвих ЛФД.

Лавинні фотодіоди мають добру лінійність відносно оптичної потужності в інтервалі від часток нВт до кількох мкВт.

Кремнієвий ЛФД з робочою поверхнею $2 \cdot 10^{-5}$ см² і зворотною напругою 8...15 В має діапазон **спектральної чутливості** 0,35...1,13 мкм і $S_\lambda = 0,5$ А/Вт для $\lambda = 0,9$ мкм; ємність переходу приблизно 1,5 пФ, що дозволяє реалізувати частотний діапазон до 10 ГГц, темповий струм $I_T \approx 10^{-9}$ А. Для ЛФД характерні такі основні види шумів:

■ **дробовий**, який залежить від коефіцієнта помноження:

$$I_{op}^2 = 2qI_0\Delta fM^x,$$

де q – заряд електрона; I_0 – середнє значення струму; Δf – смуга частот; M – коефіцієнт підсилення; $x = 2...3$, що визначається експериментально;

■ **струмовий**, який збільшується більше ніж в M разів порівняно зі звичайним фото діодом, що погіршує поріг чутливості.

Шуми збільшуються ще й через наявність мікроплазм – виникнення окремих ділянок пробою на р-п-переході раніше ніж на решті поверхні, які виникають внаслідок неоднорідностей у напівпровіднику. В неробочій області ($U_p > U_{np}$) множення фотоструму падає, і різко зростає темповий струм пробою. Напруга пробою залежно від матеріалу та конструкції фотоприймача становить від кількох десятків до сотень вольт.

Шумові характеристики фотоприймача можна покращати охолодженням.

Конструкції лавинних фотодіодів. Використовуються дві конструкції ЛФД: планарна та меза. В **планарній конструкції** світло падає перпендикулярно до р-п-переходу. Оскільки в результаті кривини р-п-переходу величина $U_{проб}$ на його краях менша, ніж у центрі fotocутливої площі, то для запобігання виникненню мікроплазм по периметру р-п-переходу створюється охоронне кільце з меншим градієнтом розподілу концентрації домішки. Для одержання високої однорідності площа переходу обмежується розмірами 10^{-3} - 10^{-4} см².

В **мезадіоді** світло падає паралельно до р-п-переходу. Конструкція ефективна у разі роботи в інфрачервоному спектральному діапазоні. Глибина поглинання фотодіодів дорівнює довжині переходу. Фотодіод має конічну форму, що перешкоджає виникненню зон локального лавинного пробою.

Недоліки та переваги лавинних фотодіодів. До недоліків лавинних фотодіодів слід віднести: необхідність високих напруг живлення при жорстких вимогах до їх стабільності (0,01-0,1%), оскільки навіть незначні коливання призводять до значної зміни фотоструму; залежність напруги пробою від температури, а також розкид її величини в різних приладах; залежність підсилення фотоструму від фонового засвічення; низький ККД. Вказані недоліки обмежують використання ЛФД в оптронах та деяких інших пристроях некогерентної оптоелектроніки.

Лавинний фотодіод має найбільший, порівняно з напівпровідниковими фотоприймачами, добуток смуги частот на коефіцієнт підсилення, що дає змогу ефективно використовувати ЛФД для прийому слабких оптичних сигналів. Коефіцієнт підсилення лавинного фотодіода, шбрашк з умови одержання найбільшого відношення сигнал/шум ($\frac{\partial \Phi_n}{\partial M} = 0$), називають оптимальним.

17.4.6. Високочастотні фотодіоди. Для зменшення постійної часу фотодіода зменшують товщину бази для того, щоб утворені на поверхні в результаті освітлення

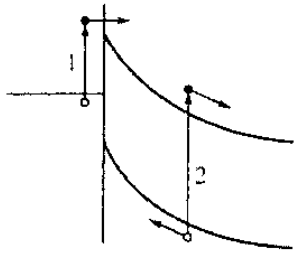


Рис. 17.38. Фотоефект в поверхневому бар'єрі, який ввімкнено в зазірному напрямку.

носії швидше дійшли до р-n-переходу і рекомбінували.

Зменшення сталої часу фотодіода отримують, використовуючи бар'єри Шотткі (метал-напівпровідник) та гетеропереходи.

Фотодіоди на основі бар'єрів Шотткі.

Фотоприймачі з поверхневим бар'єром Шотткі також мають високу швидкодію і ефективність. Подібні бар'єри, що утворюються на контакті металу з напівпровідником (рис. 17.38), можуть бути отримані і на матеріалах, в яких неможливо створити р-n-переходи.

Якщо електронний напівпровідник контактує з металом, у якого робота виходу електронів менша від роботи виходу для напівпровідника, то певна кількість електронів переходить із напівпровідника в метал. Іонізована донорна домішка в напівпровіднику утворює шар позитивного просторового заряду, з високим опором. У разі увімкнення діода в зазірному напрямку (мінус на металі) ширина бар'єра збільшується відповідно до формули:

$$d = \left[\frac{1 \epsilon \epsilon_0 (U_k + U)}{q N_d} \right]^{1/2},$$

де ϵ – діелектрична проникність; ϵ_0 – електрична стала; U_k – контактна різниця потенціалів; U – зовнішня напруга; N_d – концентрація донорів.

Якщо $N_d = 1 \cdot 10^{17} \text{ см}^{-3}$, то при напрузі в кілька вольт $d \approx 1 \cdot 10^{-5} \text{ см}$. Тонкий шар металу може бути нанесений на напівпровідник методом вакуумного розпилювання. Світло спрямовують на кристал крізь цю майже прозору плівку (у випадку діода Au-Si шар Au пропускає 95% випромінювання з $\lambda = 0,63 \text{ мкм}$). Якщо $d > x^*$, то основна частина світла з енергією фотонів $h\nu > E_g$ поглинається в області, в якій присутнє сильне електричне поле (перехід 2 на рис. 17.38). Пари, які виникають, швидко розподіляються полем, і час прольоту t_i може бути дуже малим ($\tau_i = 10^{-10} - 10^{-10} \text{ с}$), особливо в тому випадку, коли x і товщина d менші ніж 1 мкм.

З підвищенням частоти зменшується x^* , і область поглинання світла зсувається до металевому шару. Але внаслідок близькості бар'єра до поверхні це поглинання залишається в межах поля бар'єра, і дифузійна складова струму відсутня. Чутливість діодів Шотткі, як і р-і-n-переходів, висока і досягає $S_I = 0,5 \text{ А/Вт}$. Особливістю бар'єрів Шотткі є можливість реєстрації фотонів з $h\nu > E_g$. Якщо енергія фотонів більша від висоти бар'єра з боку металу, електрони можуть вводиться у напівпровідник з металу (перехід 1 на рис. 17.38). Відповідно розширюється область спектральної чутливості фотодіода.

Щоб попередити сильне відбивання світла від поверхні діода, на неї наносять просвітлювальне покриття, наприклад, плівку сульфїду цинку з

коефіцієнтом заломлення 2,30 для $\lambda = 0,63$ мкм. Товщину плівки вибирають такою, щоб внаслідок інтерференції відбитих від меж плівки променів, отримати мінімум, тобто відбивання було відсутнє для променів завдовжки l .

Бар'єр Шотткі (контакт метал-напівпровідник) утворюється нанесенням на напівпровідник n- або р-типу тонкої прозорої металевої плівки. Матеріалами можуть слугувати різні напівпровідники, що використовуються в оптоелектроніці, та багато металів, однак промислове використання мають кремнієві, германієві, арсенід-галієві фотодіоди із золотою плівкою завтовшки 0,01 мкм. Внаслідок різниці робіт виходу матеріалів контакту в базі діода утворюється область об'ємного заряду, збіднена носіями. Аналогічно р-і-n-фотодіоду на області об'ємного заряду падає майже все прикладене до фотоприймача зворотне зміщення.

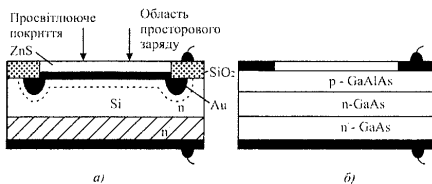


Рис. 17.39. Будова фотодіодів на основі бар'єрів Шотткі (а) і гетеропереходів (б).

Типова структура AuSi фотодіода Шотткі зображена на рис. 17.39, а. Охоронне кільце використовується для зниження струму втрат та підвищення пробивної напруги. Існують конструкції, в яких кільце не використовується. Просвітлювальне покриття зменшує втрати на відбивання.

Залежно від довжини хвилі випромінювання фотони поглинаються в області об'ємного заряду бази (короткохвильова частина спектрального діапазону), вони можуть також надавати металам енергію, достатню для подолання бар'єра переходу (довгохвильова частина).

Відмінністю фотодіодів Шотткі від звичайних р-п-приладів є висока чутливість до ультрафіолетового випромінювання ($\lambda \approx 0,2$ мкм), яке поглинається в приповерхневому шарі бази.

Фотодіоди Шотткі мають близьку до граничної інтегральну чутливість $S_{\text{інтегр}} = 0,3-0,5$ А/Вт і високу швидкодію $10^{-9}-10^{-10}$ с, яка визначається лише часом прольоту носіїв області об'ємного заряду (під дією сильного однорідного електричного поля $E = 10^5$ В/см).

Частотна характеристика таких фотодіодів поширюється до сотень мегагерц. Для того щоб випромінювання поглиналось у більшій мірі в самому переході, його необхідно розширити, що досягається використанням високоомних матеріалів і прикладанням високої зворотної напруги. В цьому випадку випромінювання поглинається в р-п-переході, а утворені при цьому пари носіїв розділяються існуючим сильним полем і там же рекомбінують. Стала часу таких фотодіодів рівна приблизно 10 нс.

Високочастотні фотодіоди з гетеропереходом. У високочастотних фотодіодах з гетеропереходом матеріал, освітленого напівпровідника, підбирається так, щоб реєстроване випромінювання проходило через нього і поглиналось у самому р-п-переході і відразу за ним. Глибина залягання р-п-переходу в цьому випадку

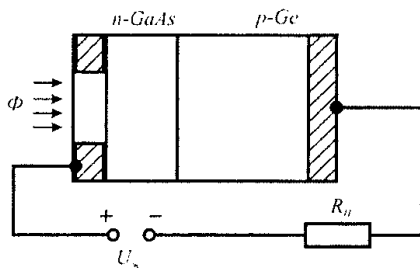


Рис. 17.40. Структура фотодіода з гетеропереходом.

невелика (гетерофотодіод з переходом n-GaAs-p-Ge не має глибини залягання р-n-переходу 50 мкм). Структури фотодіодів з гетеро-переходами показані на рис. 17.39, б та рис. 17.40.

У гетерофотодіодах р-n-перехід утворюють два напівпровідники з різною шириною забороненої зони. Для створення гетеропереходу потрібен точний збіг кристалічних ґраток матеріалів, які використовуються. Характерним прикладом є структура GaAs-GaAlAs, зображена на рис. 17.39, б. На підшарі з арсеніду галію n^+ -типу нарощуються шари n-GaAs і p^+ -GaAlAs. Шар p^+ -GaAlAs з широкою забороненою зоною пропускає падаючий світловий потік. Поглинання випромінювання відбувається у вузькозонному шарі n-GaAs, товщина якого повинна бути досить великою. Якщо до фотодіода прикласти зворотне зміщення, він працюватиме як р-і-n-діод.

Наявність широкозонного "вікна" у гетерофотодіодів дає ряд переваг: р-n-перехід можна виготовляти далеко від поверхні, оскільки немає обмеження залежністю глибини поглинання від довжини хвилі; можна створювати широку спектральну характеристику, що охоплює максимуми ($\lambda_{\text{макс1}}$, $\lambda_{\text{макс2}}$) обох напівпровідників.

Гетерофотодіоди мають високу чутливість та швидкодію, які залежать від параметрів середнього шару пристрою. Ємність гетеропереходу дещо більша, ніж у р-і-n-діода.

У фотогоальванічному режимі фотодіод з гетеропереходом генерує ЕРС, яка дорівнює 0,8-1,1 В, що у 2...3 рази перевищує напругу ненавантаженого режиму кремнієвих зразків.

17.4.7. Інжекційні фотодіоди. Інжекційні фотодіоди належать до фотоприймачів із внутрішнім підсиленням. Структуру діода характеризує довга базова область. Довжина бази з високоомного напівпровідника у кілька разів перевищує довжину дифузії неосновних носіїв струму. Фотодіод працює при прямому зміщенні.

У темновому режимі більша частина прикладеної до діода напруги падає на опорі бази. За наявності зміщення приймача світла фотони, поглинаючись у базі, утворюють нові носії струму, опір бази зменшується, що призводить до перерозподілу напруги між базою та р-n-переходом. З підвищенням напруги р-n-перехід інjektує більший струм, який, у свою чергу, ще більше зменшує опір бази і т. д. Внаслідок цього процесу відбувається підсилення первинного фотоструму.

В інжекційних фотодіодів з високоомного матеріалу (Ge, Si, GaAs) чутливість на 5-6 порядків вища, ніж у звичайних фотодіодів, а пороговий струм та інерційність приблизно такі, як і у фоторезисторів. Водночас, порівняно з лавинними фотодіодами знижуються вимоги до стабільності напруги живлення та струму.

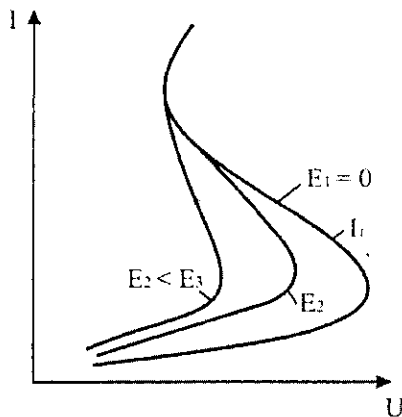


Рис. 17.41. Вольт-амперна характеристика S-фотодіода.

17.4.8. S-фотодіоди. Структура та принцип дії S-фотодіодів аналогічні до інжекційного фотодіода. Фотони світла створюють носії струму в довгій базовій області, що стимулює ріст інжекційного струму та призводить до помноження фотоструму.

При освітленні S-фотодіода значно знижується напруга зриву вольт-амперної характеристики (рис. 17.41). Залишкова напруга практично не змінюється. Чутливість фотоприймача максимальна у випадку слабкої освітленості.

S-фотодіод може працювати як аналоговий фотоприймач на початковій ділянці вольт-амперної характеристики (ВАХ). Перетворення є нелінійним. На S-фотодіоді зручно створювати фотореле. Для цього на фотоприймач подають напругу зміщення, яка відповідає вибраному порогу освітленості, при перевищенні якого різко зростає струм через S-діод. Фотоприймач переходить у другий стійкий стан (верхня гілка ВАХ), який зберігається після затемнення приладу.

17.4.9. Використання фотодіодів. Фотодіоди є наймасовішим типом фотоприймачів. У фотодіодному режимі найповніше реалізуються такі позитивні властивості, як широкий спектральний діапазон, лінійна енергетична характеристика, висока швидкодія, часова стабільність характеристик. У фотогальванічному режимі роботи фотодіоди мають нижчий рівень шумів та темпових струмів. Основний недолік розглянутих фотоперетворювачів – температурна залежність темного струму, який у фотодіодному режимі може наближатися до значень фотоструму. На відміну від фоторезисторів, у фотодіодів не спостерігається внутрішнього підсилення фотоструму.

Фотодіоди використовуються у перетворювачах лінійних та кутових переміщень. На контрольовану поверхню прозорого диска або стрічки наносять відповідно до необхідного кутового або лінійного розділення рисунок. Швидкість обертання (переміщення) можна визначити підрахунком кількості світлових імпульсів, які потрапляють на фотоприймач, проходячи через контрольовану поверхню та вузьку діафрагму (щілину) від світлодіода. Інформацію про напрямок одержують за допомогою муарових картин. У цьому випадку встановлюються 2-4 фотодіоди, кожний з яких приймає світлові імпульси, зсунуті на 90° за просторовою фазою відносно сусіднього фотоприймача. Амплітуда імпульсів обернено пропорційна до густини муарової картини.

Кремнієвий фотодіод – найкращий фотоприймач у приладах та інформаційно-вимірювальних системах контролю параметрів випромінювачів та індикаторів візуального відображення інформації. Оскільки пристрої вимірюють фотометричні параметри, то фотоелектричний

перетворювач повинен мати спектральну характеристику близьку за чутливістю до чутливості людського ока. Для коригування спектральної чутливості фотодіодів до складу перетворювача включають світлофільтри.

Фотодіоди використовуються в різних системах контролю та керування технологічними процесами, наприклад, для контролю параметрів процесів осадження та травлення плівок у технології виробництва інтегральних схем.

Кремнієві фотодіоди з р-п- та р-і-п-структурами є основними фотоприймачами оптронів.

Лекція 18

НАПІВПРОВІДНИКОВІ НАНОМАТЕРІАЛИ, НАНОТЕХНОЛОГІЇ ТА НАНОЕЛЕКТРОНІКА

План

Вступ

18.1. Фізичні основи квантово-розмірних структур

18.2. Методи отримання наноструктурованих матеріалів

18.2.1. Літографія

18.2.2. Молекулярно-променева епітаксія

18.2.3. Осадження наноструктур із металоорганічних сполук

18.2.4. Іонне бомбардування та імплантація

18.2.5. Молекулярне нашарування – метод хімічної зборки

18.2.6. Методи колоїдної хімії

18.2.7. Осадження газодинамічного потоку пари

18.3. Штучні атоми і кристали

18.4. Елементи наноелектроніки

18.4.1. Нанолазери

18.4.2. Одноелектронні транзистори

18.4.3. Одноелектронне тунелювання

Висновки

Література

- 18.1. Д.М. Фреїк. Напівпровідникові наноматеріали, нанотехнології та наноелектроніка (огляд) // Прикарпатський вісник НТШ. – 2009. – Число 1(1). – С. 74-111.
- 18.2. Р.А. Андриевский, А.В. Рогоул Наноструктурные материалы. М.Ж Академия. – 2005. – 1952 с.

Вступ

Розробка і використання міліметрової технології почалося в середині ХХ ст. і надалі розвивалось у зв'язку із виникненням промислової електроніки. Послідовне зменшення розмірів елементів більш, ніж у 10^3 разів привело людство у світ твердотільної мікротехнології. Її виникнення і розвиток пройшов червоною лінією через другу половину ХХ ст., що дозволило отримати величезну кількість різноманітних елементів обчислювальної техніки. Кристали кремнію, які є основою інтегральних мікросхем, мініатюризація яких і обумовлює швидке зростання ефективності і швидкодії на одиницю вартості обчислювальних машин з періодом подвоєння два роки. Рекордний результат по мініатюризації мікросхем на кремнії були досягнуті в 90-х роках ХХ ст. за допомогою електронної літографії, що дозволило отримати будь-яку структуру на мікронному і субмікронному рівнях. Але навіть при таких розмірах їх робота описується класичними законами, наприклад законом Ома, який однаково вірний як для побутового

електронагрівача, так і для інтегральної мікросхеми.

Це правило починає порушуватись при розмірах структур, що складають нанометри ($1 \text{ нм} = 10^{-9} \text{ м}$) (табл. 18.1).

Таблиця 18.1

Область функціонування наноматеріалів

| Розміри | Назва |
|-----------------|---------------------------------------|
| 0,1 нм – 0,5 нм | атоми |
| 0,5 нм – 10 нм | молекули |
| 10 нм – 100 нм | кластери металів, напівпровідників |

Вперше виразно концепція наноматеріалів була сформульована Р. Глейтером, який ввів у науковий обіг і сам термін (спочатку як нанокристалічні матеріали, потім наноструктурні, а також нанофазні, нанокомпозитні і т.д.). З 1992 року регулярно, кожні два роки, проводяться міжнародні конференції (NANO), а щорічна кількість симпозіумів, семінарів і т.д. по даній тематиці вже складає декілька десятків. Відрадно відзначити, що до цього напрямку долучилися і наші міжнародні конференції з фізики і технології тонких плівок, які з XI (2007 р.) стали включати в назві і змісті доповіді із наноструктур – МКФТТПН.

18.1. Фізичні основи квантово-розмірних структур

Наноструктуровані матеріали можна розділити на наступні три основні класи:

– квантові точки (КТ) (quantum dots (QD)) – структури, розміри яких у всіх трьох напрямках складають декілька міжатомних відстаней (залежно від масштабу розгляду структура вважається нульмірною або тримірною);

– квантові дроти (КД) (quantum wires (QWr)) — структури, розміри яких в двох напрямках складають декілька міжатомних відстаней, а розмір в третьому напрямі є макроскопічною величиною;

– квантові стінки (КС), інакше квантові ями (quantum wells (QW)) — структури, розмір яких в одному напрямі складає декілька міжатомних відстаней, а розміри в двох інших напрямках є макроскопічною величиною.

У чому особливість таких нанооб'єктів? По-перше, внаслідок таких малих розмірів вони мають дуже високу величину поверхні (наприклад, частинки діоксиду кремнію – аеросила – практично непористі і розміром декілька десятків нанометрів, можуть мати питому поверхню до $400 \text{ м}^2/\text{г}$). Як наслідок, такі структури проявляють високу фізико-хімічну активність. По-друге, розміри наночасток порівнянні з довжиною вільного пробігу електрона, що впливає не просто на зміну властивостей, але на їх стрибкоподібну зміну (ефект розмірного квантування).

Зупинимося на деякій деталізації цих особливостей.

Відомо, що енергія E вільного електрона визначається співвідношеннями

$$E = \frac{m_0 v^2}{2} = \frac{p^2}{2m_0}, \quad (18.1)$$

де m_0 – ефективна маса електрона, v – швидкість та p – квазіімпульс. Тобто він може мати будь-яке значення енергії. У кристалі реалізується енергетична структура із дозволеними і забороненими зонами. Якщо по одному із напрямків рух електрона обмежений (наприклад тонка фольга товщиною d і $d < \lambda$, де λ – довжина хвилі де-Бройля електрона), тоді енергія електрона E_n як і імпульс p_n може приймати тільки визначені значення:

$$E_n = \frac{p_n^2}{2m} = \frac{p^2 \hbar^2 n^2}{2md^2}, \quad p_n = \frac{p \hbar n}{d}, \quad (18.2)$$

$$n = 1, 2, 3.$$

Таким чином мінімальна енергія E_{\min} ($n = 1$) обернено пропорційна до квадрату лінійних розмірів d ($E_{\min} \sim \frac{1}{d^2}$) і зростає із зменшенням цього характерного розміру.

У тому випадку, коли відстань між рівнями помітно більша теплової енергії, такий електронний спектр носіїв у квантових об'єктах дає можливість усунути основну проблему при виробництві приладів сучасної мікро- і оптоелектроніки – «розмивання» носіїв заряду в енергетичному вікні порядку $k_B T$ (де k_B – стала Больцмана; T – абсолютна температура; підвищення робочої температури приводить до деградації провідних властивостей приладів).

Для характеристики електронів у кристалах вводиться поняття густини станів $g(E)$, яка визначається як число станів які приходяться на одиничний інтервал енергії E . У двовимірному випадку для кожного з квантово-розмірних рівнів з енергією E_N повне число станів дорівнює $G_N(E) = m(E - E_N)S / (2p\hbar^2)$, де S – площа зразка. Густина станів, яка в цьому випадку обчислюється на одиницю площі, визначається сумою по всіх рівнях, енергії яких E_N лежать нижче E :

$$g_2(E) = \frac{2}{S} \sum_N \frac{dG_N}{dE} = \sum_N \frac{m}{p\hbar}. \quad (18.3)$$

Нарешті, для кожного рівня E_i квантової нитки довжиною L $G_1 = \sqrt{2m(E - E_i)}L / (\pi\hbar)$ і густина станів на одиницю довжини

$$g_1 G = \frac{2}{L} \sum_i \frac{dG_i}{dE} = \frac{\sqrt{2m}}{\hbar} \sum_i \frac{1}{\sqrt{E - E_i}}. \quad (18.4)$$

На рис. 18.1 схематично показано функції густини станів для всіх випадків, що обговорювалися. Видно, що вони носять якісно різний характер. У тривимірному випадку густина станів монотонно зростає з енергією (рис. 18.1,а), у двовимірному випадку має вигляд

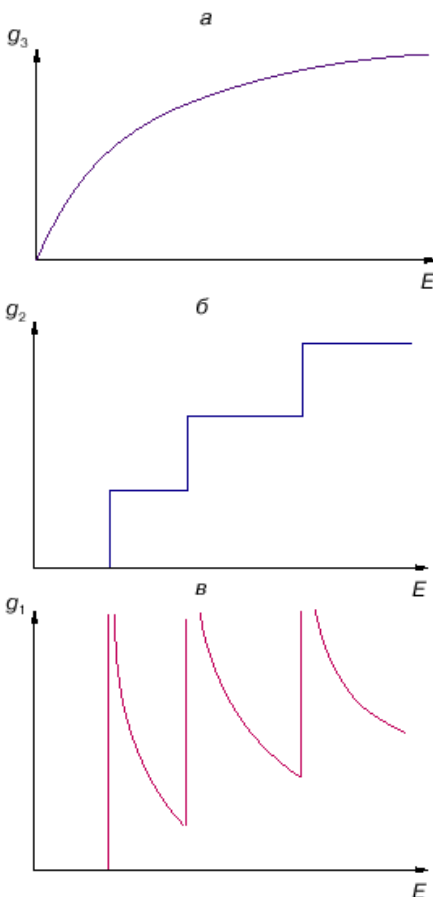


Рис. 18.1. Густина станів у масивному тривимірному напівпровіднику (а), двовимірних електронних структурах – квантових ямах (б) та одновимірних структурах – квантових нитках (в).

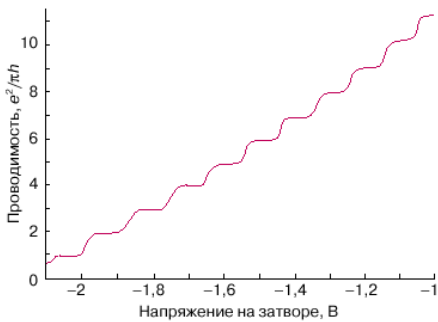


Рис. 18.2. Експериментальна залежність провідності квантової нитки (точкового контакту) від напруги на затворі, що визначає концентрацію носіїв.

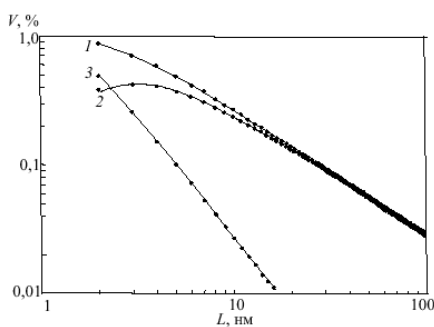


Рис. 18.3. Залежність загальної частки поверхонь розділу (1), частки міжзеренних меж (2) і потрійних стиків (3) у структурі наноматеріалів від розміру зерен (товщина меж зерен прийнята рівною 1 нм).

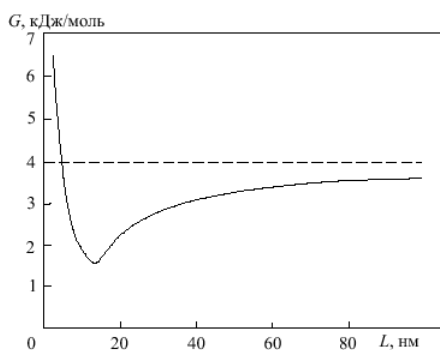


Рис. 18.4. Залежність молярної вільної енергії Гіббса для бінарного нанокристалічного сплаву з середньою концентрацією 5% (ат.) від розміру кристалітів (для звичайного полікристала ($G = 4$ кДж/моль ($T = 600$ K)).

горизонтальних сходинок (рис. 18.1,б), а в квантових нитках необмежено росте кожного разу, коли ми наближаємося зверху до енергії чергового квантового рівня (рис. 18.1, в).

Якщо температура достатньо низька, то електрони заповнюють стани з найменшими енергіями. У кожному стані, згідно принципу Паулі, може розташовуватися по два електрони з протилежними спінами. Легко показати, що поки число електронів на одиницю довжини нитки n менше, ніж $2\sqrt{2m(E_2 - E_1)}/p\hbar$, всі вони розташовуються на першому квантовому рівні і провідність нитки $S = e^2/p\hbar$. Як тільки концентрація перевищить вказане значення, частина електронів буде вимушена розміститися на другому рівні і в провідності з'явиться додатковий член $e^2/p\hbar$, що описує внесок цього рівня. Іншими словами, провідність стрибком зростає. Такі скачки відбуватимуться всякий раз, коли електрони починають заповнювати черговий рівень. В цілому залежність $\sigma(n)$ повинна мати ступінчастий вигляд, причому висота цих сходинок рівна універсальній величині $e^2/p\hbar$ (рис. 18.2).

Практично всі типи наноматеріалів, за винятком супрамолекулярних, через умови отримання і особливості структури є нерівноважними. У найзагальнішому вигляді відхилення від рівноваги і, відповідно, надмірна вільна енергія Гіббса, можуть бути пов'язані з характерною для наноматеріалів великою кількістю поверхонь розділу (міжзеренні і міжфазні межі, потрійні стики), наявністю нерівноважних фаз міжфазних сегрегацій, залишкових напруг і підвищеного вмісту дефектів кристалічної будови. На рис. 18.3 показана зміна загальної частки (поверхонь розділу, а також власне міжзеренних меж і потрійних стиків) залежно від розміру зерен.

Цілком очевидно, що при термічних діях, а також у силових полях (радіаційних, деформаційних і ін.) неминучі релаксаційні, сегрегаційні і гомогенізаційні процеси рекристалізації, а також явища розпаду, фазових перетворень, спікання і запливання нанопор (нанокاپілярів), аморфізації і кристалізації. Все це повинно позначатися на фізико-хімічних, фізико-механічних і інших властивостях, впливаючи тим самим на експлуатаційні ресурси наноматеріалів і визначає важливість вивчення їх стабільності. Це, як і дослідження природи розмірних ефектів, одна з найважливіших і порівняно маловивчених проблем наноструктурного матеріалознавства.

Необхідно також мати на увазі, що у разі легованих наноструктур вплив розміру кристалітів на інтегральну вільну енергію може бути немонотонним (рис. 18.4). Такий хід залежності $G = f(L)$ (рис. 18.4) означає, що через існування мінімуму при $L_{крит}$ ріст зерен в інтервалі

$L < L_{\text{крит}}$ стає термодинамічно невідповідним.

Слід також звернути увагу на те, що, змінюючи зарядність численних міжзеренних областей, можна впливати на електронну структуру в цілому, тобто на магнітні, оптичні, електричні і інші властивості, а також на взаємну розчинність і утворення нових фаз.

Таким чином, розмірні ефекти у наноструктурах супроводжуються квантовими ефектами, оскільки розмір кристалітів стає сумірним з довжиною дебройлівської хвилі λ_B .

Для металів $\lambda_B \approx 0,1 - 1$ нм, а для напівметалів і вузькощілинних напівпровідників $\lambda_B \sim 100$ нм. Отже, первинні концепції одержують нові цікаві імпульси для розвитку: виявлення ролі розмірних і квантових ефектів в модифікації властивостей наноматеріалів.

18.2. Методи отримання наноструктурованих матеріалів

На даний час вже розроблено і ефективно функціонують цілий ряд технологій і методик, що забезпечують отримання наноструктурованих матеріалів, які використовуються як у чисто академічному плані, так і на практиці. Ми звернемо увагу тільки на окремих способах формування напівпровідникових квантово-розмірних структур.

18.2.1. Літографія. Найочевидніше отримати наноструктуру можна через безпосереднє "вирізування" вузької смужки за допомогою літографічної техніки (рис. 18.5, а). При цьому для отримання електронних ниток шириною в сотні ангстрем, де квантування енергій електронів буде помітним, необхідно робити смужки саме такої ширини, що вимагає літографічної техніки надвисокої роздільної здатності. Річ у тому, що на бічних гранях витравленої смужки, як і на вільній поверхні напівпровідника, утворюються поверхневі стани, що створюють, як правило, шар збіднення. Цей шар викликає додаткове звуження провідного каналу, внаслідок чого квантові ефекти можна спостерігати і в смужках більшої ширини — порядку десятої частки мікрона.

Можна поступити і інакше. Поверхню напівпровідникової структури покривають металевим електродом, що створює з напівпровідником контакт Шотткі, який має вузьку щілину (рис. 18.5, б). Якщо гетерограниця знаходиться достатньо близько від поверхні, в шарі збіднення, то двовимірні електрони на межі відсутні усюди, окрім вузької області під щілиною. Такий тип одновимірної структури володіє додатковою перевагою: міняючи напругу на затворі, ми можемо управляти ефективною шириною квантової нитки і концентрацією носіїв в ній.

18.2.2. Молекулярно-променева епітаксія. Метод молекулярно-променевої епітаксії (molecular-beam epitaxy, MBE) призначений для вирощування

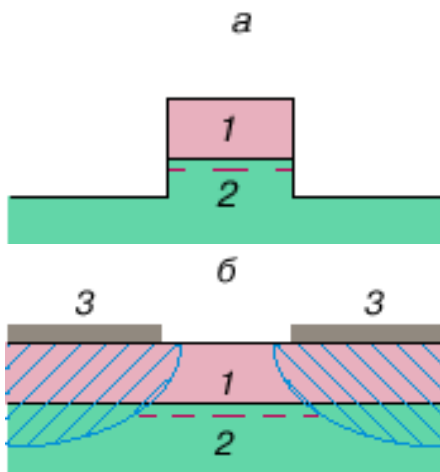


Рис. 18.5. Напівпровідникові гетероструктури з квантовими нитками, одержані за допомогою субмікронної літографії за рахунок витравлення вузької смужки з самої структури (а) або щілини в затворі Шотткі (б): 1 – напівпровідник з широкою забороненою зоною (наприклад, AlGaAs), 2 – напівпровідник з вузькою забороненою зоною (наприклад, GaAs), 3 – металевий затвор. Вузький електронний канал, що утворюється поблизу гетерограниці, показаний штриховою лінією. Заштриховані області збіднення електронами.

кристалічних структур у надвисокому вакуумі за допомогою пучків атомів або молекул, що є компонентами сполуки, що росте. Таким чином, МВЕ є удосконаленням звичного способу напилення плівок випаровуванням у вакуумі. Принципи технології МВЕ формувалися поступово. У 1964 році (R.V. Schoolar, J.N. Zemel) за допомогою молекулярних пучків були одержані досконалі епітаксійні плівки PbS на монокристалічній підкладці NaCl. У 1968 році (J.E. Davey, T. Pankey, J.R. Arthur) в умовах високого вакууму були вирощені епітаксійні плівки GaAs на монокристалічних підкладках арсеніду галію. У цьому ж році (1968) автором огляду успішно були проведені технологічні роботи із отриманням монокристалічних плівок халькогенідів олова відкритим випаровуванням у вакуумі. Ці дослідження стали основою подальшого прогресу по вирощуванню досконалих плівок GaAs і інших напівпровідникових сполук A_3B_5 , A_4B_6 і A_2B_6 методом МВЕ. Новий етап в розвитку і вдосконаленні МВЕ почався в 70-ті роки ХХ ст. і пов'язаний із створенням і промисловим виробництвом відповідного вакуумного устаткування.

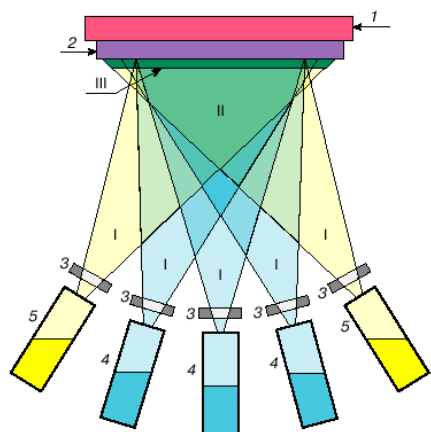


Рис. 18.6. Схема методу МВЕ для отримання плівок легованих потрійних сполук. Вся установка розміщена в камері надвисокого вакууму. 1 – блок нагріву; 2 – підкладка; 3 – заслінка окремої комірки; 4 – ефузійні комірки основних компонентів; 5 – ефузійні комірки легуючих домішок.

Атомні або молекулярні пучки створюються в ефузійних (ефузія – повільне витікання газів через малі отвори, досліджене в 1911 році датським фізиком М. Кнудсеном) комірках при достатньо високій температурі і направляються у напрямку нагрітої до необхідної температури монокристалічної підкладки (рис. 18.6). Атоми в пучках рухаються за інерцією в надвисокому вакуумі, не стикаючись один з одним або будь-якими іншими атомами. Такий балістичний характер транспорту компонентів напівпровідникової сполуки забезпечує створення однорідних метастабільних твердих розчинів навіть у таких системах, які в умовах, близьких до термодинамічної рівноваги (саме така ситуація має місце при традиційних способах вирощування кристалів), мають область незмішуваності. МВЕ забезпечує епітаксійний ріст тонких плівок напівпровідникових сполук за рахунок реакцій між компонентами атомних або молекулярних пучків з поверхнею підкладки. Швидкість осадження речовини на підкладку за порядком величини звичайно складає один моноатомний шар у секунду. Отримання якісних структур можливе при використанні високочистих джерел компонентів і за умови точного контролю температур підкладки і джерел, що може бути реалізоване лише при комп'ютерному управлінні параметрами процесу росту.

Римськими цифрами на рис. 18.6 позначені області (зони), що визначають основні особливості МВЕ: I – зона генерації молекулярних пучків ефузійними комірками Кнудсена з механічними заслінками; II – зона змішування (перетину) пучків елементів, випаровуваних з різних

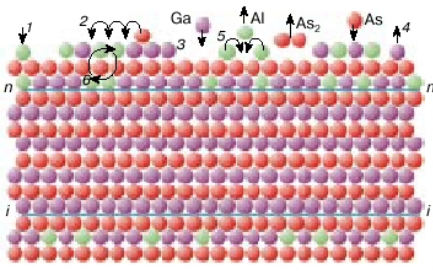


Рис. 18.7. Елементарні процеси у зоні росту в методі МВЕ:

1 – адсорбція атомів із зони змішування на поверхні; 2 – міграція адсорбованих атомів по поверхні; 3 – входження адсорбованих атомів у кристалічну ґратку; 4 – термічна десорбція; 5 – утворення поверхневих зародків; 6 – взаємна дифузія; n-n та i-i – нормальна та інвертована області гетероструктури.

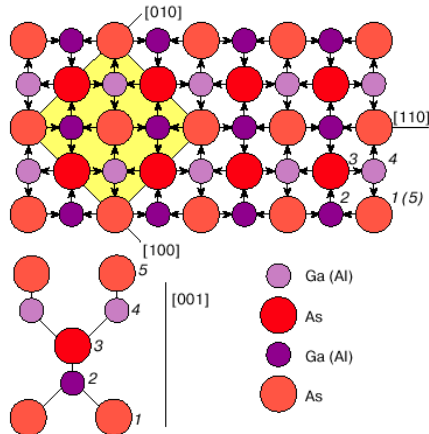


Рис. 18.8. Кристалічна структура цинкової обманки. Вгорі – проекція на площину (001), перпендикулярну осі росту. Виділена грань елементарного куба в площині (001). Цифрами вказані атоми, що заповнюють шар в площині (001), стрілки вказують послідовність шарів.

ефузійних комірок; III — зона кристалізації на підкладці (зона росту).

Зону росту можна умовно розділити на три частини: першою зоною є кристалічна підкладка або черговий моноатомний шар, що виріс; другою – газова суміш компонентів у приповерхневій області; третьою – перехідний шар, геометрія якого і процеси в ньому сильно залежать від вибору умов росту. Отже, якщо необхідно виростити конденсат методом МВЕ, потрібно мати можливість належним чином регулювати структуру і склад перехідного шару. Для вирощування кристалографічно досконалих структур слід так підбирати режим росту, щоб перехідний шар був максимально тонким, тобто моноатомним. Ця умова може бути виконана, якщо потік атомів, які падають на підкладку, близький до потоку атомів, що випаровуються з підкладки. У цьому режимі ріст структури здійснюється шляхом утворення і подальшого росту двовимірних зародків на атомарно-плоскій поверхні.

Епітаксійний ріст у методі МВЕ включає наступні елементарні процеси (рис. 18.7):

- 1) адсорбція (прилипання) падаючих на підкладку атомів або молекул, що становлять вирощувану сполуку;
- 2) міграція (поверхнева дифузія) адсорбованих атомів по поверхні підкладки (їй може передувати дисоціація молекул вирощуваної сполуки);
- 3) вбудовування атомів у кристалічні ґратки підкладки або моноатомний шар гетероструктури, що росте;
- 4) термічна десорбція атомів, що не вбудувалися в кристалічні ґратки;
- 5) утворення і подальший ріст двовимірних зародків кристала на підкладці або на поверхні шару, що росте;
- 6) взаємна дифузія атомів, що вбудувалися в кристалічні ґратки.

Найпоширенішими кристалічними ґратками для сполук A_3B_5 і A_2B_6 є так звана структура цинкової обманки ZnS , показана на рис. 18,8. Якщо поверхня підкладки паралельна одній із граней елементарного куба з індексами Міллера (001), то епітаксійний ріст здійснюється послідовним нарощуванням аніонних і катіонних шарів, як це показано на рис. 18.7, де схематично представлена гетероструктура (ГС) $Al_xGa_{1-x}As - GaAs$, в якій шар арсеніду галію зосереджений між двома шарами твердого розчину алюмінію (з атомною концентрацією x) в арсеніді галію. Алюміній, близький за своїми хімічними властивостями до галію, заміщає атоми галію в ґратках $GaAs$, не призводячи при цьому до будь-якої істотної зміни міжатомних відстаней. Кожен напівпровідник (наприклад, $GaAs$ або $Al_xGa_{1-x}As$) може бути вирощений пошарово при фіксованій швидкості росту, коли температура підкладки забезпечує оптимальну для даної сполуки швидкість поверхневої

дифузії. Оскільки хімічні зв'язки у напівпровідникових сполуках відмінні між собою, то різними будуть і енергії активації поверхневої дифузії катіонів, що входять до складу цих сполук. Тому якість гетеромеж може бути істотно різною залежно від того, яка із сполук при вибраному температурному режимі росте першою. Межі прийнято називати нормальними, якщо компонент з нижчою температурою плавлення росте першим (наприклад, $\text{Al}_x\text{Ga}_{1-x}\text{As}$ на GaAs); для зворотної послідовності використовується термін "інвертована межа". На рис. 18.7 проекції цих двох меж зображені лініями з буквеними індексами n і i відповідно. Щоб одержати гладші і досконаліші гетеромежі, використовується методика переривання росту або методика осадження пульсуючим пучком. Згладжування поверхні протягом часу переривання росту (саме переривання здійснюється механічним перекриттям на деякий проміжок часу заслінок ефузійних комірок) обумовлене поверхневою міграцією і/або сублімацією атомів, адсорбованих на поверхню вирощеного моношару.

Температура підкладки визначає співвідношення між потоками адсорбції і десорбції атомів, що входять до складу структури, що росте. Це співвідношення може бути охарактеризоване коефіцієнтом прилипання атома даного сорту до поверхні, на якій відбувається епітаксійний ріст. Коефіцієнт прилипання показує, яка частина падаючого потоку адсорбується на поверхні. При адсорбції аніонного компоненту, наприклад миш'яку при рості арсеніду галію, важливу роль відіграє те, в якій молекулярній формі (As_2 або As_4) аніонний компонент доставляється на поверхню, де відбуваються дисоціація молекул і подальше вбудовування атомів в кристалічні ґратки.

Крім того, температура підкладки визначає швидкість поверхневої дифузії, що передує вбудовуванню атомів в кристалічні ґратки. Ця температура повинна бути достатньо високою, щоб забезпечити необхідне число ($\sim 10^4$) дифузійних стрибків атома на поверхні, перш ніж він займе своє положення в ґратці. При цьому середнє переміщення атома по поверхні за час t визначається як

$$x = \sqrt{D_s t}, \quad \text{де} \quad D_s = D_{s0} \exp\left(-\frac{E_{sd}}{T}\right) \quad - \quad \text{коефіцієнт}$$

поверхневої дифузії. Передекспоненціальний фактор $D_{s0} = a^2 \nu$, де a — довжина дифузійного стрибка (відстань між сусідніми еквівалентними положеннями атома на поверхні росту), T — температура в енергетичних одиницях, $\nu = 10^{12} \text{ c}^{-1}$ — частота коливань атома на поверхні. Енергія активації поверхневої дифузії E_{sd} в напівпровідниках звичайно складає 1 – 1,5 еВ, тому за одну секунду атом в середньому зміщається на декілька

десятків або сотень міжатомних відстаней при температурі підкладки 600-800 °С. Такий зсув якраз і відповідає приблизно 10^4 дифузійним стрибкам.

Істотне збільшення температури підкладки небажане з двох причин: по-перше, воно може призвести до зменшення коефіцієнтів прилипання; по-друге, до активізації взаємної дифузії, тобто дифузії атомів між шарами. Оскільки ГС є різко неоднорідні (на атомному масштабі) за хімічним складом структури, то з часом за рахунок процесів взаємної дифузії ці структури повинні або переходити в термодинамічно рівноважний стан з однорідним розподілом концентрацій всіх компонентів, або розшаровуватися на фази цілком певного складу. Проте, оскільки енергія активації взаємної дифузії атомів в напівпровідниках звичайно складає 4 – 5 еВ, а передекспоненціальний фактор має той же порядок величини, що і у разі поверхневої дифузії, то очевидно, що в інтервалі температур 600-800 °С цей ефект неймовірно малий, оскільки середній зсув атома в результаті взаємної дифузії навіть за декілька десятків годин виявляється істотно меншим міжатомної відстані. Це твердження тим більше справедливе для кімнатної температури, при якій звичайно працюють прилади, елементами яких є ГС. Вибір і підтримка оптимальної температури росту є однією з найважливіших умов здійснення МВЕ.

При послідовному вирощуванні шарів GaAs і $Ga_{1-x}Al_xAs$, що мають практично однакові параметри кристалічних ґраток, можна виростити досконалі структури, що містять двовимірні електрони. Останніми роками було виявлено, що при понижених температурах росту при моношаровому рості напівпровідника з параметрами кристалічних ґраток, відмінними від параметрів ґраток підкладки, можна одержати на поверхні росту майже однакові за розміром острівці. Острівці осадженого напівпровідника пірамідальної форми практично не містять дефектів, є квантовими точками. Як приклад можна привести осадження InAs (параметр ґратки $a_0 = 0,60584$ нм) на підкладці GaAs ($a_0 = 0,56533$ нм). Такий процес одержав назву самоорганізації квантових точок в процесі росту (рис. 18.9, а).

Рушійною силою утворення острівців, а не плоского шару є прагнення системи до мінімуму енергії. При неспівпаданні параметрів ґраток виникають пружні деформації як острівця, що утворився, так і підкладки. Якщо в процесі росту поверхня залишається плоскою (пунктир на рис. 18.9, а), то енергія пружної деформації зростає, а енергія поверхні не змінюється. При певній товщині (зазвичай декілька моношарів) така ситуація стає енергетично невигідною. Мінімуму енергії системи відповідатиме утворення тривимірних острівців-пірамід

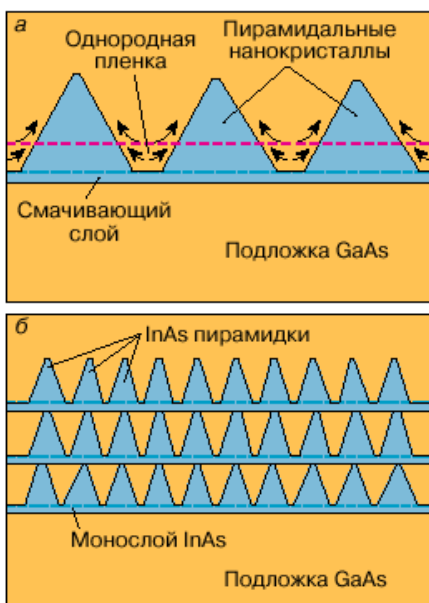


Рис. 18.9. Утворення пірамідальних острівців InAs в результаті трансформації плоских шарів цього напівпровідника (пунктир), вирощуваних на підкладці GaAs з відмінними параметрами ґраток (а); система зв'язаних вертикально квантових точок InAs на підкладці GaAs (б).

(стрілки на рис. 18.9, а) , в яких основа буде сильно деформована (для InAs на поверхні GaAs стиснута), але деформація зменшуватиметься у міру віддалення від підкладки до вершини піраміди. Один шар InAs, званий змочуючим шаром, відділяє піраміди від підкладки GaAs.

Таким чином, рушійною силою утворення квантових точок в процесі самоорганізації є зменшення енергії деформації. Так, якщо на підкладці з арсеніду галію ростити шар арсеніду індію, параметр ґратки якого більший, то виникнуть пружні напруги, що призведуть до росту острівців (квантових точок) InAs на поверхні GaAs. Поверхня покривається пірамідами InAs з розмірами в декілька десятків ангстрем. Виявляється вигіднішим формування острівців в порівнянні з однорідно-напруженою поверхнею. Зростання пірамід відбувається до повного зняття пружної напруги на вершині піраміди.

Утворення погоджено-напружених острівців на поверхні залежить від двох параметрів: розузгодження ґратки і кількості осадженого на поверхню матеріалу. Мінімум енергії відповідає утворенню однакових за розміром і формою острівців. У такий спосіб вдається виростити не тільки моношар пірамід на поверхні арсеніду галію, але і послідовні шари, як це показано на рис. 18.9, б. При цьому вершини пірамід є зародками утворення пірамід в новому шарі – має місце утворення штучних кристалів (див. розділ III цього огляду).

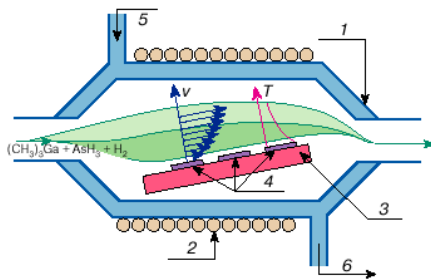


Рис. 18.10. *Схема горизонтального реактора відкритого типу з охолоджуваними стінками для MOVPE-технології наноструктур: 1 – кварцовий корпус, 2 – котушка ВЧ-генератора для нагрівання підкладки, 3 – блок нагрівання, 4 – підкладки, 5 – водяне охолодження (впускання), 6 – водяне охолодження (випуск). Схематично показаний розподіл швидкостей v і температури T в газівому потоці в дифузійному шарі поблизу підкладки (масштаб не дотриманий).*

18.2.3. Осадження плівок з металоорганічних сполук. Методика осадження плівок з металоорганічних (МОС) сполук (metalorganic vapor phase epitaxy, MOVPE), відома також як МОС-гідридна технологія, найістотніше відрізняється від методики МВЕ характером масопереносу: атомні або молекулярні пучки без об'ємної дифузії в МВЕ і газований потік суміші компонентів реакції з дифузиею у напрямі фронту зростання ГС в MOVPE (рис. 18.10).

Ще в 1968 році було показано (Н.М. Manasevit), що газозфазні суміші металоорганічної сполуки триметилгалію $(\text{CH}_3)_3\text{Ga}$ і гідриду миш'яку AsH_3 при здійсненні реакції піролізу (розкладання при підвищеній температурі) в атмосфері молекулярного водню H_2 при температурах 600 – 700°C в реакторі відкритого типу можуть бути використані для вирощування епітаксійних плівок GaAs. Надалі були виконані дослідження, що показали ефективність методики MOVPE для вирощування різних напівпровідникових сполук у вигляді тонких плівок субмікронних розмірів. В даний час більшість сполук A_3B_5 , A_2B_6 і A_4B_6 можуть бути одержані в методиці MOVPE. Загальна проблема, що виникає при вирощуванні різних напівпровідникових структур, полягає в оптимальному виборі металоорганічних сполук для реакцій піролізу в процесі росту.

Металоорганічні сполуки утворюють широкий клас

речовин, що мають хімічні зв'язки “метал – вуглець” або “метал – кисень – вуглець”, а також координаційні сполуки металів і органічних молекул. Сполуки, що представляють інтерес для вирощування напівпровідникових структур, при кімнатній температурі, як правило, є рідинами, хоча деякі із сполук, використовуваних в техніці MOVPE, знаходяться в твердому стані. Ці сполуки можуть бути легко перенесені в зону реакції в потоці газу-носія, наприклад молекулярного водню, що проходить через рідке джерело або над поверхнею твердого джерела. Гідриди, вживані в напівпровідниковій технології, при кімнатній температурі є газами і звичайно використовуються у вигляді розбавленої суміші з H_2 . Металоорганічні і гідридні компоненти змішуються в газовій фазі і піддаються піролізу в потоці атмосфери H_2 через реактор відкритого типу (рис. 18.10), який функціонує при атмосферному або зниженому тиску. Температура піролізу звичайно складає від $600^\circ C$ до $800^\circ C$. Енергія для розігрівання газової суміші підводиться від радіочастотного генератора або потужної лампи. Відбувається розігрівання графітного столика, на якому розміщуються монокристалічні підкладки. Газова суміш поблизу підкладок також нагрівається до високої температури, тоді як стінки камери реактора залишаються холодними, наприклад за рахунок водяного охолодження. Такий режим забезпечує осадження продуктів реакції на підкладках і призводить до відносно малих втрат компонентів реакції на стінках камери. Кристалізація в процесі MOVPE виникає при проходженні газової суміші компонентів реакції в потоці газу-носія поблизу нагрітої підкладки і йде в умовах, істотно ближчих до термодинамічної рівноваги, ніж при МВЕ.

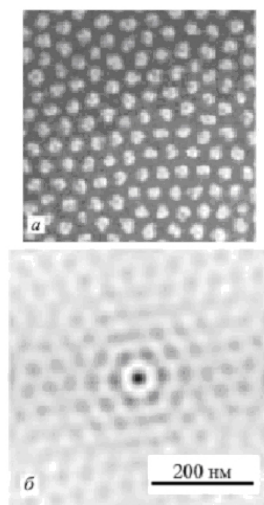
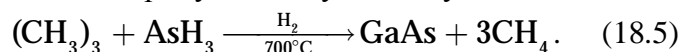
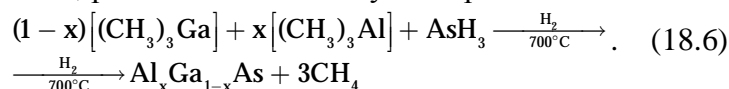


Рис. 18.11. (а) СЕМ-зображення поверхні монокристалічного GaSb при опромінюванні іонами Ar^+ з енергією 420 eV і (б) відповідне двомірне автокореляційне представлення регулярності і гексагонального впорядкування точок.

Найчастіше використовуються метилові або етилові металоорганіки через відносну простоту приготування і легкості піролізу в атмосфері молекулярного водню. Ці сполуки постачають атоми металів з побічними продуктами у вигляді метану або етану. Хімічна реакція в окремому випадку MOVPE росту арсеніду галію з триметилгалієм і гідриду миш'яку може бути записана як



Подібні реакції використовуються для вирощування інших подвійних, а також потрійних і четвертих сполук. Наприклад, ріст $Al_xGa_{1-x}As$ описується рівнянням



В цьому випадку атомна концентрація x алюмінію в твердому розчині $Al_xGa_{1-x}As$ прямо пов'язана з відносним початковим парціальним тиском триметилгалію і триметилалюмінію в газовій фазі.

18.2.4. Іонне бомбардування та імплантація.

Особливо цікавими в плані отримання квантових точок за допомогою іонного розпилювання представляються результати, одержані співробітниками Institute of Semiconductor Electronics (Німеччина). При вертикальному падінні пучка іонів Ar^+ з енергією 420 eВ вони одержали на поверхні (100) GaSb гексагональну кристалічну структуру з квантових точок діаметром 35 нм (рис. 18.11). Слід підкреслити, що описані вище експерименти об'єднуються загальним механізмом, пов'язаним перш за все з іонним розпилюванням поверхні мішені, а також з утворенням на поверхні нових багатокомпонентних сполук.

На Міжнародній конференції з модифікації твердих тіл іонними пучками (ІВММ 98, Амстердам) були представлені результати робіт по формуванню впорядкованої системи нанорозмірних елементів – паралельно розташованих лінійних утворень CoSi_2 на поверхні зразків монокристалічного кремнію, опромінених іонами Co^+ (рис. 18.12). Встановлено, що кластери CoSi_2 ростуть у формі квадрата до тих пір, поки їх розмір не перевищить $e\alpha_0$ (e – основа натурального логарифма, α_0 – розмір, який відповідає мінімальній енергії кластера). Коли це відбувається, одночасно починає збільшуватися довжина кластера і зменшуватися його ширина до значення α_0 (67 нм на Si(100)). Це свідчить про те, що сформовані силіциди перебудовуються для зменшення ширини. Крім того, слід врахувати, що розмір сформованого кластера залежить від енергій міжфазної взаємодії між кластером і поверхнею підкладки, що безпосередньо впливає на значення, яке визначається формулою

$$\alpha_0 = e\phi h \exp(\Gamma / (ch)), \quad (18.7)$$

де c — константа, залежна від тензора напруги острівця, коефіцієнта Пуассона і модуля зсуву підкладки; h і f – висота і краєвий кут кластера відповідно.

Структури Si – Ge з квантовими точками є одним з найцікавіших об'єктів для квантової електроніки і оптоелектроніки. Такі структури вирощуються за допомогою молекулярно-променевої епітаксії і на них, принаймні в макетному варіанті, продемонстровані основні передбачені приладові можливості. Проблема, проте, полягає у тому, що реальна побудова технології навіть з невеликим серійним випуском приладів натрапляє на те, що установки молекулярно-променевої епітаксії є в одиничних екземплярах навіть в найрозвиненіших країнах світу (США, Японія, Англія), а процес як підготовки зразків, так і власне реалізації росту займає тривалий час. У зв'язку з цим достатньо природною альтернативною технологічною можливістю для побудови аналогічних об'єктів, а саме структур Si –

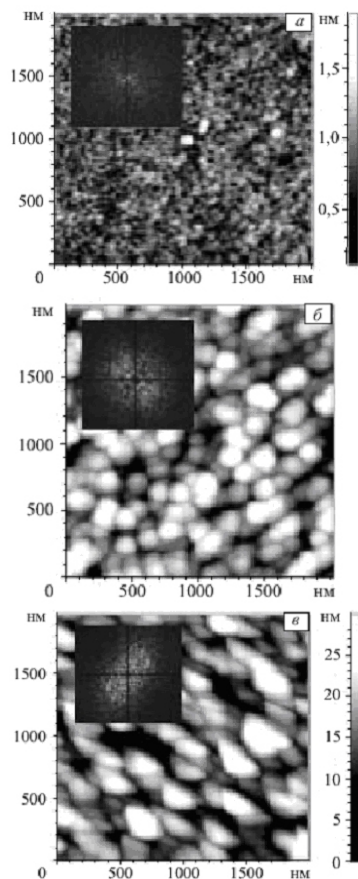


Рис. 18.12. АСМ-топографія у двовимірному вигляді зразка p -Si, опроміненого йонами кобальту ($D = 10^{17} \text{ см}^{-2}$, $E = 150 \text{ кеВ}$, $j = 5 \text{ мкА/см}^2$).

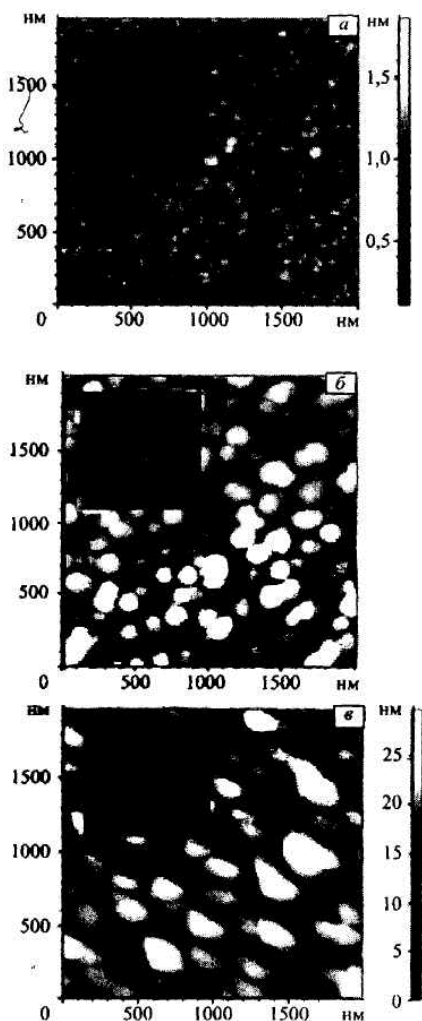


Рис. 18.13. АСМ-топографія у двовимірному вигляді зразка p -Si(111) з питомим опором $\rho = 1$ Ом-см, опроміненого іонами Ge^+ ($D = 10^{17}$ см $^{-2}$, $E = 150$ кеВ, $j = 5$ мкА/см 2): а – без відпалу; б – після пічної відпалу при температурі 980°C протягом 30 хв; в – після двох послідовних пічних відпалів при температурі 980 °С протягом 30 хв (1) і при температурі 1100 °С протягом 30 хв (2). На врізках показані Фур'є-перетворення цих ділянок поверхні.

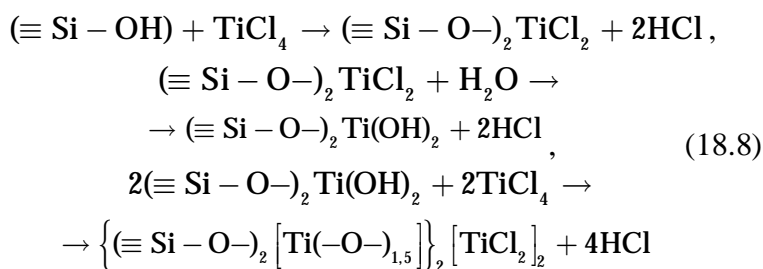
Ge з квантовими точками є іонна імплантація.

При впровадженні іонів Ge^+ в кремнієву підкладку і подальшому дослідженні опроміненої поверхні за допомогою атомно-силового мікроскопа виявилось, що опромінена поверхня стає шорсткою з явно вираженими нанорозмірними неоднородностями (рис. 18.13). Як видно з рис. 18.13, існуючі неоднорідності при певних режимах можуть організовуватися в паралельно розташовані витягнуті утворення, які можуть бути квантовими дротами.

18.2.5. Молекулярне нашарування – метод хімічної зборки.

В основі цих підходів лежать нові уявлення у області хімії надмолекулярних сполук, в рамках якої розроблені прецизійні хімічні способи синтезу твердих речовин – методи хімічної зборки. Одним із них, що одержав найширший розвиток і застосування, є метод молекулярного нашарування (МН). Суть його полягає в реалізації в нерівноважних умовах хімічних реакцій на поверхні твердого тіла між реагентами, що підводяться ззовні, і поверхневими функціональними групами (ФГ) підкладки. Методом МН можна як синтезувати на поверхні твердофазної матриці наноструктури різного хімічного складу (моношари, зокрема багатокомпонентні), так і здійснювати поатомну хімічну збірку поверхневих нано- і мікроструктур шляхом багатократного чергування хімічних реакцій за заданою програмою.

Так, синтез титаноксидного шару на поверхні діоксиду кремнію (біла сажа марки БС-30) здійснюється за рахунок багатократного повторення реакцій:



При цьому товщина утвореного оксидного шару (0,2-2,0 нм) має лінійну залежність від числа циклів МН (1-20).

Дослідження властивостей продуктів, одержаних методом МН, дозволили знайти корисні для практики ефекти (рис. 18.14): розмірно-структурні (рис. 18.14, а-в) – ефект моношару, тобто різка, стрибкоподібна зміна властивостей матриці після нанесення 1 – 4-го моношарів нових структурних одиниць (див. рис. 18.14, а); ефект перекриття підкладки, коли утворюється шар, який фізично екранує поверхню (рис. 18.14, б); ефект взаємного узгодження структури поверхні підкладки і шару, що нарощується (рис. 18.14, в); ефекти

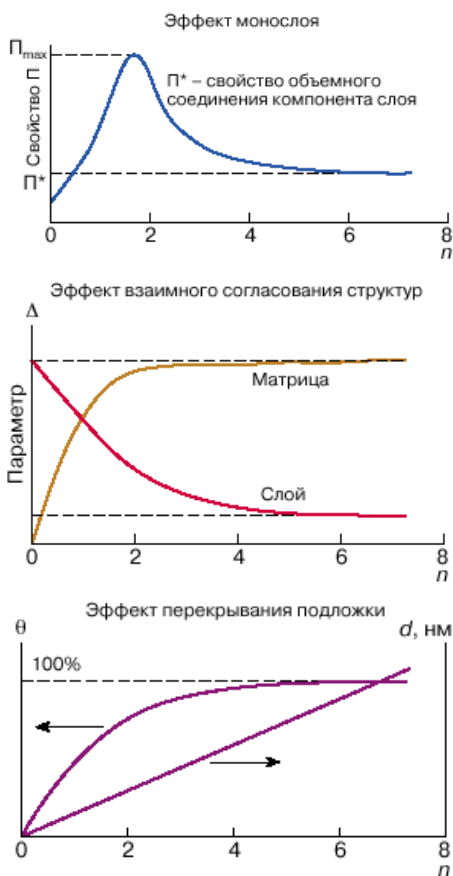


Рис. 18.14. Графічне зображення розмірно-структурних ефектів у процесах МН: n – число циклів МН (або кількість моношарів); θ – ступінь фізичного перекриття поверхні нарощуваним шаром; d – товщина покриття; Δ – структурний параметр.

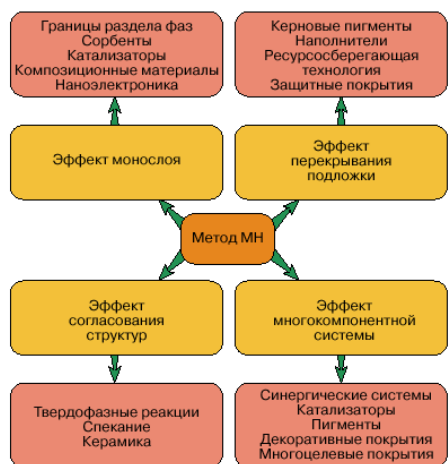


Рис. 18.15. Особливості застосування методу молекулярного нашарування.

багатокомпонентної системи, які виявляються при створенні багатозонних і синергичних низькорозмірних систем.

Можливі області застосування нової прецизійної технології з урахуванням знайдених закономірностей і різних напрямів розвитку твердофазного матеріалознавства представлені на рис. 18.15.

18.2.6. Методи колоїдної хімії. Отримання квантових точок, які характеризуються тим, що виходять нанокристали без внутрішніх пружних напруг, заснований на методах колоїдної хімії. Це один з найстаріших способів, ще М. Фарадей одержував колоїдне золото в 1857 році. Він дозволяє отримувати нанокристали елементів II – IV (наприклад, CdS) або III – V (наприклад, InP, GaP, GaInP₂, GaAs, InAs) груп Періодичної системи елементів сферичної форми розміром від 1 до 5 нм в органічних розчинниках, які полімеризуються при кімнатній або нижчій температурі. Відмінними особливостями цього методу є низька температура (біля 200°C) синтезу колоїдних частинок, можливість широкої зміни концентрації напівпровідникових частинок, невелика концентрація поверхневих дефектів.

Розчин хімічних реагентів, що містять сполуки елементів II і IV груп, вводять в розчинник, що містить молекули, що взаємодіють з поверхнею новостворених наночастинок. Це обмежує їх ріст. Крупніші частинки можна осадити і одержати розчин практично однакових за розміром частинок. В даний час вдається відділити частинки з діаметрами, що відрізняються тільки на декілька відсотків.

Синтез у такий спосіб речовин III – V груп Періодичної системи елементів більш важкий. Не повинно міститися ні повітря, ні води, температура реакції вища, тривалість реакції довшя, доводиться використовувати складніші хімічні сполуки. Найкращі результати одержані в даний час при синтезі InP. Береться In(C₂O₄)₃ і InF₃ або InCl₃. Ці речовини взаємодіють з складними сполуками, що містять фосфор. Реакція йде декілька днів при температурі 270-290°C. Від тривалості реакції залежить розмір одержуваних частинок. У описаному способі частинки покриті зверху шаром молекул з розчину, які можуть бути замінені на інші, наприклад полімерні. Одержані частинки можна вивчати в розчині, у вигляді порошку або поміщати в прозорий полімер або органічне скло.

Існує ще досить поширений спосіб, приготування неорганічного скла, забарвленого нанокристалами сполук II – VI (CdS) і I – VII (CuCl, CuBr, CuI) Періодичної системи елементів Д.І. Менделєєва. Цей спосіб пов'язаний з промисловим отриманням світлофільтрів. Ріст напівпровідникових нанокристалів відбувається при

розпаді пересиченого розчину іонів у склі. При цьому способі виходять стабільні твердотільні стьокла з вкрапленими нанокристаллами. Ріст кристалів в скляній матриці відбувається при температурі 550-700°C. Ця температура звичайно перевищує температуру плавлення об'ємних напівпровідникових кристалів, яка складає 400-500°C. Із зменшенням розміру нанокристала до 1-2 нм температура плавлення знижується до 200-250°C. До недоліків такого росту відносяться широкий розкид розмірів частинок, неможливість впливати на параметри межі розділу скло-нанокристал.

18.2.7. Осадження із газодинамічного потоку пари. Альтернативним до МВЕ є запропонований нами технологічний процес вирощування наноструктур із газодинамічного потоку пари.

Елементи теорії газодинамічного потоку пари. На рис. 18.16 представлена схема реактора для реалізації газодинамічного потоку пари. Профілі газодинамічних змінних розраховуються в безрозмірному вигляді як функції від безрозмірної координати $\xi = (x-x_{кр})/(L-x_{кр})$. Нормування параметрів пари проводиться по їх значенням в початковому газодинамічному перерізі або, що те ж саме, за їх значеннями в критичному перерізі конденсації. В інтервалі $x_0 \leq x \leq x_{кр}$ параметри пари, яка рухається в циліндрі не змінюється.

Задаючи різні значення ξ (від 0 до 1), можна обчислити відповідні їм значення числа Маха.

$$M(\xi) = 1 + 1,91\xi - 1,325\xi^2 + 1,48\xi^3 + \dots \quad (18.9)$$

Обчислення коефіцієнта конденсації α , нормованих значень густини (ρ/ρ_1), температури (T/T_1), витрати (G/G_1), швидкості потоку (u/u_1) пари, швидкості результуючої конденсації ($\omega^* = \omega/\omega_1$), ступеня пересичення парової фази (ψ) проводимо за наступними виразами:

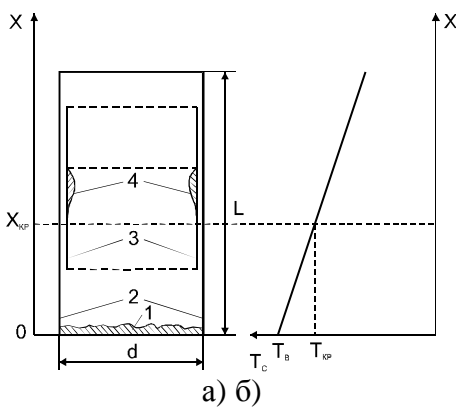


Рис. 18.16. Схема реалізації газодинамічного потоку пари у квазізамкнутому об'ємі (а) і розподіл температури T_c вздовж стінки камери (б): 1 – наважка для випаровування; 2 – стінки камери; 3 – підкладка; 4 – конденсат.

$$a = 1 - \left(\frac{1 + 0,2M^2}{1,2} \right)^3 \exp\left(-\frac{6,88x}{1 - 0,242x} \right); \quad (18.10)$$

$$\frac{r}{r_1} = \left(\frac{1,2}{1 + 0,2M^2} \right)^{2,5}; \quad (18.11)$$

$$\frac{p}{p_1} = \left(\frac{1,2}{1 + 0,2M^2} \right)^{3,5}; \quad (18.12)$$

$$\frac{T}{T_1} = \frac{1,2}{1 + 0,2M^2}; \quad (18.13)$$

$$\frac{G}{G_1} = M \left(\frac{1,2}{1 + 0,2M^2} \right)^3; \quad (18.14)$$

$$\frac{u}{u_1} = M \sqrt{\frac{1,2}{1 + 0,2M^2}}; \quad (18.15)$$

$$w^* = \frac{w}{w_1} = 1,6 \left[\left(\frac{1,2}{1+0,2M^2} \right) - \exp\left(-\frac{6,68}{1-0,242x} \right) \right]; \quad (18.16)$$

$$y = \frac{r(T)}{r_H(T_C)}. \quad (18.17)$$

Нормуючий параметр ω_1 являє собою потік частинок пари, що падає на стінку в початковому газодинамічному перерізі:

$$w_1 = \frac{u_1 r_1 d}{4r_{me}(L-x_{кр})}, \quad (18.18)$$

де ρ_{me} – густина масивного матеріалу для телуриду кадмію. У виразі для ступеня пересичення парової фази чисельник дорівнює густині пари в деякому перерізі $\rho(T)$, а знаменник дорівнює густині насиченої пари $\rho_H(T_C)$ при температурі стінки камери в цьому ж перерізі.

Результати комп'ютерного розрахунку залежностей профілів термодинамічних параметрів пари (ψ , α , ω^*) для ізолюваної витратної дії від безрозмірної координати ξ чи T_{II} для $x_{кр}$ і градієнта температури вздовж стінок $\frac{dT_C}{dx}$ для телуриду олова представлено на рис. 18.17. Межі зміни технологічних параметрів:

$$900 \text{ K} \leq T_B \leq 1400 \text{ K}; \quad \frac{dT_C}{dx} = 5 \dots 80 \text{ K/cm}; \quad L = 8 \text{ см}; \quad d = 3 \text{ см}.$$

Аналіз діаграм дозволяє зробити ряд важливих висновків про характер зміни параметрів пари, а також оптимальних режимів отримання плівок. Так, довжина ділянки конденсуючої підкладки $x_{кр}$, на якій відсутній шар конденсату, зростає як при збільшенні температури випаровування T_B , так і зменшенні градієнта температури $\frac{dT_C}{dx}$.

Зменшення абсолютного значення градієнта температури сприяє більш різкому зростанню $x_{кр}$, ніж збільшення температури випаровування. При достатньо малих значеннях градієнта $\frac{dT_C}{dx} = 5 \text{ K/cm}$ для

$T_B = 900 \dots 1400 \text{ K}$ відстань критичного перерізу перевищує довжину камери ($x_{кр} > L$) і рівноважна конденсація матеріалу проходить за межами випарної камери.

Коефіцієнт конденсації α , як і швидкість конденсації ω^* та пересичення ψ , залежить, в першу чергу, від температури випаровування T_B , градієнта температури вздовж стінок $\frac{dT_C}{dx}$, а також визначається місцем

розміщення конденсуючої підкладки ξ . На ділянці $x = x_{кр}$, $\xi = 0$ конденсація пари відсутня ($\alpha = 0$). При $0,0 \leq x \leq 0,4$ спостерігається різке зростання коефіцієнта конденсації $0,0 \leq \alpha \leq 0,9$. Для безрозмірної координати $0,0 \leq x \leq 1,0$ внесок ревіпарування зменшується і коефіцієнт

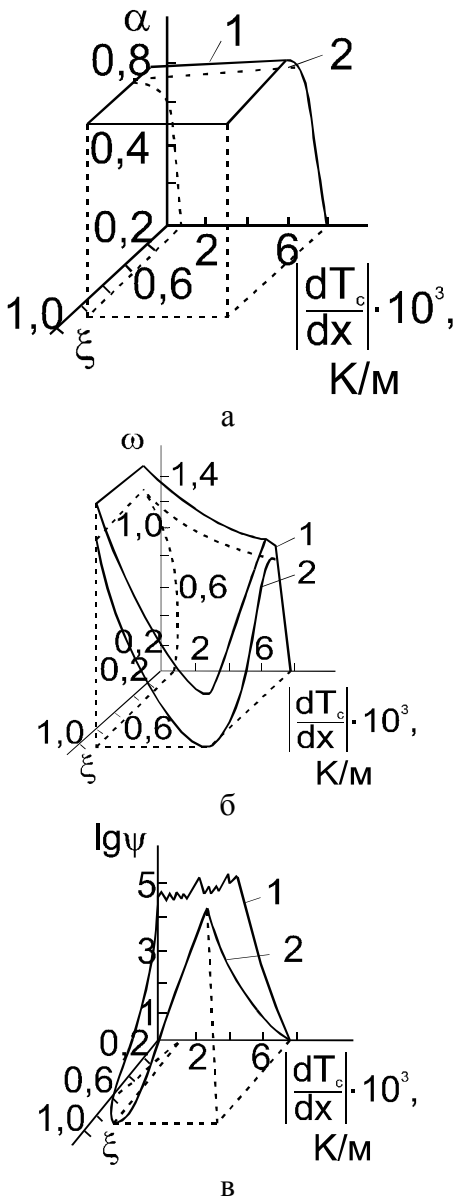


Рис. 18.17. Залежність розрахованих значень коефіцієнта конденсації α , результуючої швидкості конденсації ω^* та ступеня пересичення парової фази ψ SnTe від градієнта температури $\frac{dT_C}{dx}$ та безрозмірної координати ξ (температура випарування T_B , К: 1 – 900; 2 – 1200; $L = 8$ см, $d = 3$ см).

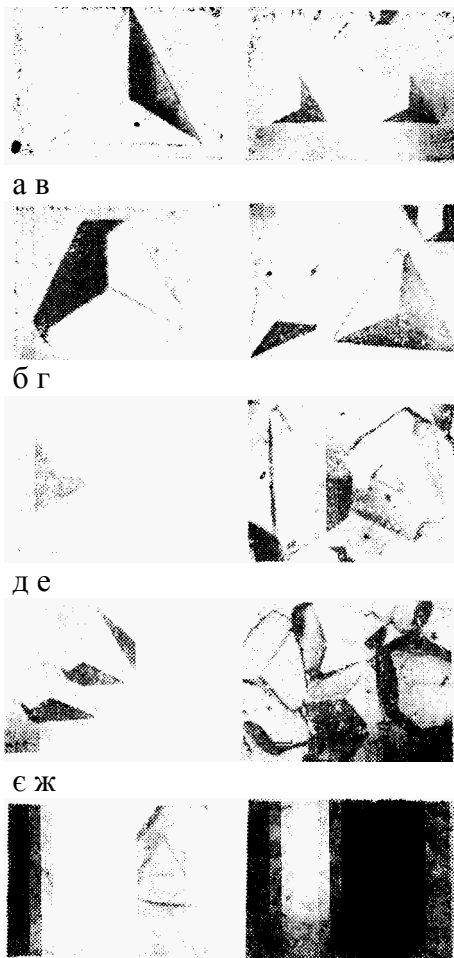
наближається до одиниці ($a \approx 1,0$) (рис. 18.17,а).

Результуюча швидкість конденсації пари ω^* на стінки камери залежить від технологічних параметрів ξ , $\frac{dT_C}{dx}$,

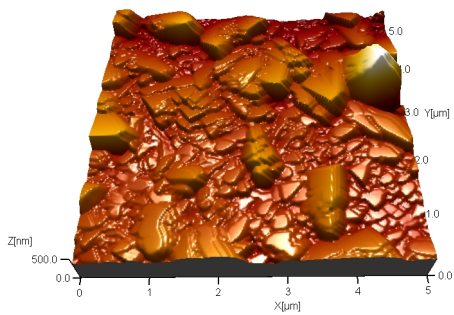
T_{II} . Початкове збільшення нормованої швидкості конденсації обумовлене деяким зменшенням реєвипарування пари (рис. 18.17,б). Подальший більш плавний спад ω^* визначається зменшенням тиску пари в камері. Пара може настільки розрідитися, що газодинамічний потік перетвориться в молекулярний. Положення максимуму швидкості конденсації відповідає умовам синтезу плівок найбільшої товщини. Зменшення $\frac{dT_C}{dx}$ призводить до зсуву ξ_{max} , що відповідає максимальним значенням швидкості конденсації у області великих значень безрозмірної координати. Так само діє збільшення температури випаровування (рис. 18.17).

Структура і процеси росту нанокристалів телуриду олова. На підкладках із сколів слюди, розташованих у напрямі руху пари, у області критичного перерізу ($x = x_{кр}$) при $y = 1,8 \dots 6,2$; $a = 0,1 \dots 0,2$; $w^* = 0,2 \dots 0,4$, не дивлячись на наявність переважних напрямів газового потоку, відбувається ріст окремих структурно-досконалих тривимірних частинок-тетраєдрів (рис. 18.18, а,б,в,г). В умовах зростання, близьких до нормальних, не спостерігається викривлення ребер, які обмежують форми кристалів; грані, ребра і вершини переміщуються узгоджено. Поверхнева густина кристалографічних форм складає $2 \cdot 10^6 \text{ м}^{-2}$, що у $10^7 \dots 10^9$ разів менше густини зародків при отриманні плівок відкритим випаровуванням у вакуумі. На гранях $\{111\}$ кристалів-зародків при сильних збільшеннях помітна шарова будова. У ряді випадків спостерігається також утворення пірамід $3m$ симетрії, але з усіченими вершинами (рис. 18.18, б).

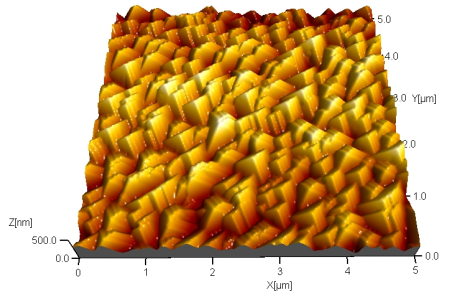
При віддаленні від поверхні випарника вздовж осі камери і зростанні пересичення пари значно збільшується густина кристалів-зародків, зменшуються їх лінійні розміри (рис. 18.18, є, ж). У цих умовах кристалізації змінюється топологія пірамід росту, їх симетрія знижується від $3m$ до m і до 1. Найчастіше виявляється трансформація основ пірамід. Причиною спостережуваних явищ у групах пірамід може бути, зокрема, те, що вони закривають одна одну від напрямлених потоків пари, яка підживлює ріст кристалів. При значних пересиченнях $\psi > 10^2$; $a = 0,9 \dots 1,0$; $w^* = 0,9$ симетрія пірамід ще більше понижується (рис. 18.18, ж). У більшості випадків вони втрачають тригональну форму. Переважаючим механізмом росту є розростання і зрощення тривимірних острівців первинних зародків. Зрощення острівців без злиття частинок викликає появу дефектів і пор на межах зрощення. Слід зазначити анізотропію спікання зародків.



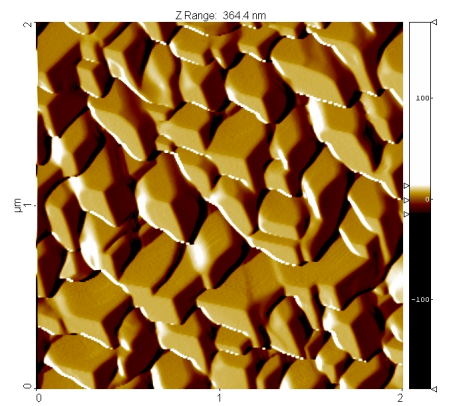
з и
Рис. 18.18. Кристалічна структура SnTe, осадженого на слюді, розміщеній паралельно потоку пари при $\frac{dT_C}{dx} = -5 \cdot 10^3 \text{ К/м}$ і пересиченні пари ψ : а, б – 2, 4; в, г, д, е – 12; є – 34, ж – $2 \cdot 10^2$; з, и – $3 \cdot 10^4$.



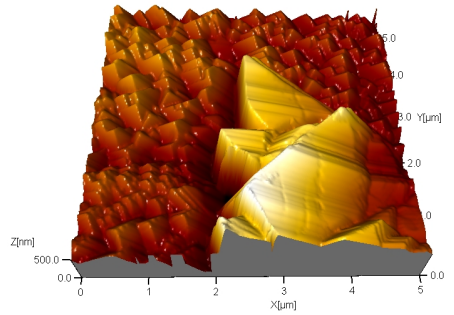
а



б



в



г

Рис. 18.19. Топологія поверхні нанокристалічних структур PbTe на сколах (0001) слюди – мусковіт, вироцених при температурах T_n , К: 353 (а), 380 (б), 408 (в), 633 (г).

Легко зростаються тетраедри, розташовані вершинами один до одного в напрямку $\langle 112 \rangle$. Значно важче спікаються двійникові кристали з протилежною орієнтацією. При цьому утворюються канали вздовж напрямів $\langle 110 \rangle$. Округлення кутів, що відбувається в процесі зростання, полегшує процес спікання.

Після утворення суцільної плівки на її поверхні виникають нові центри зародження (рис. 18.18, і, и). У ряді випадків на поверхні можна спостерігати своєрідні фігури зростання, що є усіченими тетраедрами або їх комплексами, а також кристали з гексагональною симетрією. У таких епітаксійних плівках виникає перехідна область порядку середніх розмірів острівців, що утворюються з окремих зародків до моменту їх спікання в суцільний шар. При автоепітаксії спостерігається «проростання» дислокацій, двійникових меж, дефектів упаковки з нижнього шару у верхній, а також прояв принципу «еволюційної селекції». Ростуть швидше кристали переважних орієнтацій.

Результати, приведені для телуриду олова, відображають загальну закономірність процесів зростання епітаксійних плівок, отриманих в квазізамкнутому об'ємі з температурно-градієнтними стінками.

Топологія поверхні і процеси росту нанокристалічних структур телуриду свинцю. Деякі результати дослідження морфології поверхонь наноструктур PbTe методом АСМ та її розрахунки представлено на рис. 18.19. Аналіз цих результатів дає можливість встановити певні закономірності у формуванні епітаксійних наноструктур у залежності від температури росту (рис. 18.19) та їх товщини. Так, зокрема, низькі температури осадження $T_n = 353$ К (рис. 18.19, а) сприяють формуванню нанокристалів із переважанням швидкості росту у тангенціальному напрямку до поверхні підкладки.

На це вказує те, що їх лінійні розміри у площині підкладки по азимуту (0,2-1,4) мкм значно переважають розміри у нормальному напрямку – висота нанокристалів ~ 100 нм. Підвищення температури росту до $T_n = (380-408)$ К призводить до формування більш однорідних нанокристалів як по формі (рис. 18.19, б, в), так і за лінійними розмірами в азимутальному та нормальному напрямках до поверхні підкладки. Хоча лінійні розміри кристалів у площині підкладки (0,1-0,7) мкм для цих умов осадження значно перевищують їх висоту (100-350) нм.

Подальше значне підвищення температури осадження до $T_n = 633$ К призводить до росту на фоні добре сформованих кристалів діаметром (0,2-0,8) мкм і висотою (100-400) нм окремих «гігантів» із (1,6-1,8) мкм в основі і висотою до (300-1200) нм (рис. 18.19, г).

Зауважимо, що із підвищенням температури росту T_n шорсткість поверхні нанокристалічних структур зростає і

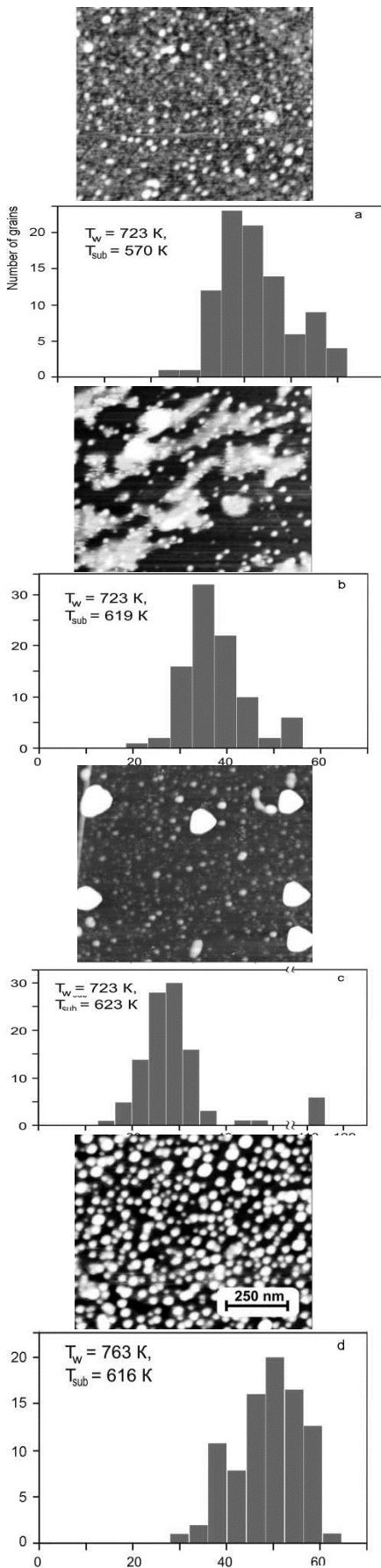


Рис. 18.20. АСМ процесів росту PbTe/VaF₂ N1 у методі гарячої стінки від температури підкладки і гістограми залежності діаметра наночастинок.

складає ~ 33 нм (353 К), ~ 49 нм (383 К) і ~ 86 нм (633 К) відповідно.

Цікавою є залежність топології епітаксійних структур від їх товщини. Для тонких плівок характерним є формування наноструктур із округленими гранями і без чіткої огранки, висота яких сягає до ~ 200 нм, а інші лінійні розміри знаходяться в межах (0,1-0,6) мкм.

Із збільшення часу осадження і товщини конденсату до (5,4-6,5) мкм має місце ріст добре сформованих нанокристалів різної висоти (50-400) нм з лінійними розмірами в основі до 1 мкм.

Таким чином, описаний характер росту нанокристалів на слюді характерний для механізму епітаксії Фольмера-Вебера. Трьохвимірні кристали зароджуються при незначному пересиченні, коли адшар надзвичайно розріджений, що характерно для слабкої адгезії, яка обумовлює орієнтаційно виражений зв'язок кристала-конденсата із підкладкою. Після утворення (спікання) суцільного шару із нанокристалів виникають нові центри зародження.

Утворення тетраедрів із нахиленими до поверхні підкладки гранями (рис. 18.19, б,в) має також енергетичне підґрунтя – зменшення поверхневої енергії меж. Це пов'язано з тим, що їх поверхня «покрита» іонними рядами із щільною упаковкою, які не несуть електричного заряду (особливість напрямків <100> у кристалах із структурою типу NaCl). Кристалографічно такі межі добре описуються моделлю за Бренденом.

Результати дослідження процесів формування нанокристалів PbTe на свіжих сколах (111) кристалів VaF₂ представлені на рис 20. Встановлено, що температура підкладки має визначальний ефект на механізмі їх формування. При низькій температурі підкладки (~ 570 К), ріст наноострівків подібний до тривимірного утворення зародків згідно механізму Вольмера-Вебера, де практично весь матеріал переходить у наноострівки (рис. 18.20). Збільшення температури приблизно до 619 К викликає виникнення двомірних фрагментів, які можуть покрити значну область підкладки. В даному випадку наноострівки з'являються як на чистій підкладці, так і на плато (рис. 18.20, б). Тут, ймовірно, потрібно розглядати комбінацію механізмів росту Вольмера-Вебера і утворення псевдоморфного двовимірного шару, які супроводжуються зародкоутворенням наноострівків згідно механізму Странського-Крastoнова. Подальше відносно мале збільшення T_n аж до температур близько 623 К визначає утворення аномально великих наноострівків з явним кристалографічним гранеутворенням (трикутна піраміда) (рис. 18.20, в). Підвищення температури стінки до 763 К при середній температурі підкладки близько 616 К призводить до істотного збільшення щільності і розмірів наноострівків і

відсутності двовимірних фрагментів плато (рис. 18.20, г). Іншими словами механізм зростання nanoострівків повертається знову до механізму Вольмера-Вебера.

Процеси зародкоутворення, поверхневої дифузії і міграція центрів утворення зародків на поверхні росту можна прослідкувати поетапно з використанням різного часу осадження. На початковому етапі осадження, є багато (понад $350 \mu\text{m}^2$) малих nanoострівків діаметром 20-40 нм і висотою до 2 нм. Подальше осадження матеріалу, поверхнева дифузія і мала міграція nanoострівків приводить до істотного зменшення густини nanoострівків (близько $200 \mu\text{m}^2$) і нерегулярного збільшення їх розмірів. Як згадано вище, на цьому етапі появляються великі пірамідальні nanoострівки.

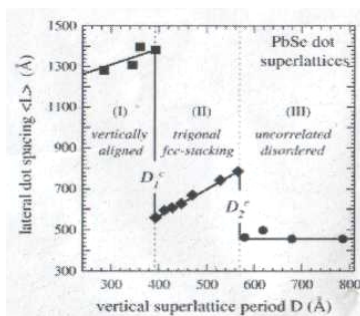


Рис. 18.21. Залежність середньої відстані між квантовими точками (у площині шару) у надрешітках квантових точок PbSe від товщини шару $Pb_{1-x}Eu_xTe$ між шарами PbSe. Видно, три режими: I – вертикальне впорядкування, II – утворення ГЦК ґратки, III – некорельований ріст.

18.3. Штучні атоми і кристали

Квантові точки іноді образно називають "штучними атомами". Раз є штучні атоми, чому б не спробувати "зібрати" з них штучні кристали? Технологія вирощування напівпровідникових квантових точок істотно полегшує цю задачу. Як вже було відзначено, спонтанне формування масиву квантових точок відбувається при нанесенні дуже тонкого (товщиною в декілька атомних шарів) шару матеріалу B на поверхню матеріалу A у разі, коли постійні кристалічних ґраток матеріалів розрізняються на декілька відсотків. При зростанні першого мономолекулярного шару сполуки B на поверхні півки із сполуки A через відмінність сталих кристалічних ґраток матеріалів виникають пружні напруги, при продовженні осадження речовини вони збільшуються і стає вигідним вже не рівномірний розподіл речовини по поверхні (тобто не продовження зростання суцільного шару), а утворення окремих "крапель" речовини (квантових точок) на поверхні першого шару (його називають "змочуючим шаром") – це сприяє мінімізації напруг в утвореній структурі. Тобто ми спочатку одержуємо не окремі "атоми", а відразу цілий шар "штучних атомів" ("атомну площину"). Наростивши зверху на утворену "атомну площину" шар матеріалу A певної товщини (яку, звичайно, можна варіювати), можна знову повторити процедуру – виростити наступну площину штучних атомів і т.д. (утворені структури по аналогії із звичними надрешітками часто називають "надрешітками квантових точок").

За рахунок підбору оптимальних умов росту (на процес формування квантових точок впливає безліч чинників – температура підкладки, швидкість росту, кількість облягаючого матеріалу, властивості використовуваних матеріалів, кристалічна досконалість підкладки і т.д.) досягнуто великих успіхів у вирішенні задачі формування однорідних масивів квантових точок. Але, мабуть, ще більш інтригуючими виявилися

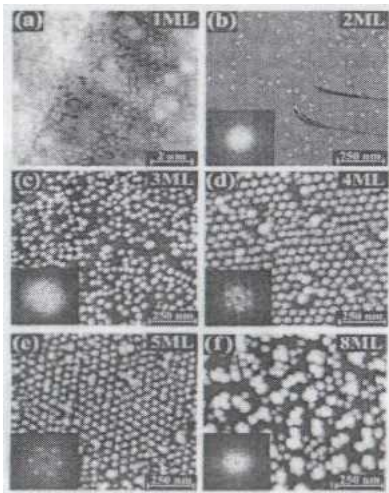


Рис. 18.22. Одержані за допомогою атомно-силової мікроскопії зображення верхнього (незарощеного) шару квантових точок PbSe при різних номінальних ростових товщинах шару PbSe (від одного до восьми моношарів) для надрешітки квантових точок (100 періодів); квантові точки вирощувалися при температурі 360 С. На вставці результат математичної обробки (спектр потужності) одержаних зображень, видно появу шести чітких піків при появі впорядкування ґратки розташування квантових крапок в площині шару (гексагональна ґратка).

результати, одержані при зростанні багатошарових структур квантових точок. Були знайдені такі цікаві ефекти, як поява спонтанного впорядкування в розташуванні квантових точок не тільки у вертикальному напрямі (напрямі росту), але і в площині шару, а також зменшення розкиду розмірів квантових точок.

Так, зокрема, досліджено вплив товщини шару матеріалу, з якого утворюються нанорозмірні острівці (PbSe в даному випадку) і температури підкладки, при якій відбувається ріст структури, на формування впорядкованих масивів квантових точок в системі PbSe/Pb_{1-x}Eu_xTe. Показано, що в структурах може реалізуватися декілька різних ситуацій (рис. 18.21): неврегульоване розташування квантових точок; впорядковане розташування квантових точок, відповідне гранецентрованим кубічним ґраткам; наявність тільки вертикальної кореляції в розташуванні квантових точок. Причому тривимірне впорядковане розташування "штучних атомів" має місце лише в досить вузькому діапазоні номінальної ростової товщини шару PbSe (номінальна товщина шару відповідає товщині однорідного шару речовини, яка б вийшла, якби матеріал розподілювався по поверхні рівномірно) і температур (див. рис. 18.22 і 18.23).

Можна також бачити, при появі тривимірного впорядкування зменшується розкид розмірів квантових точок (рис. 18.24). Це пов'язано з тим, що після вирощування першого шару квантових точок кристалічні ґратки матеріалу, яким він зарощується шар квантових точок, деформовані, причому, внаслідок анізотропії пружних властивостей кристалічних ґраток, напруги розподілені в матеріалі неоднорідно. Ці напруги і визначають переважні місця для формування квантових точок при рості наступного шару квантових точок – в місцях, де напруги мінімальні. Зокрема, для системи PbSe/Pb_{1-x}Eu_xTe в певному діапазоні розмірів квантових точок існує три переважні позиції (зміщені щодо тієї, що знаходиться під ними квантової точки в площині, перпендикулярній напрямку росту) для утворення квантових точок при рості чергового шару. Результатом такої кореляції в розташуванні квантових точок в сусідніх шарах і в площині шару і є формування гранецентрованої кубічної ґратки (у даній системі) і, відповідно, спостережувана на рис. 18.22 і 18.23 гексагональна двовимірні ґратка квантових точок.

Розрахунки показують, що існують нижній і верхній критичний розміри квантових точок, при яких спостерігається тривимірне впорядкування масивів квантових точок. При малій номінальній ростовій товщині шару квантові точки або взагалі не утворюються (рис. 18.22,а), або утворюються квантові точки з розміром нижче за критичний (рис. 18.22,в, 18.22,с), і створювані

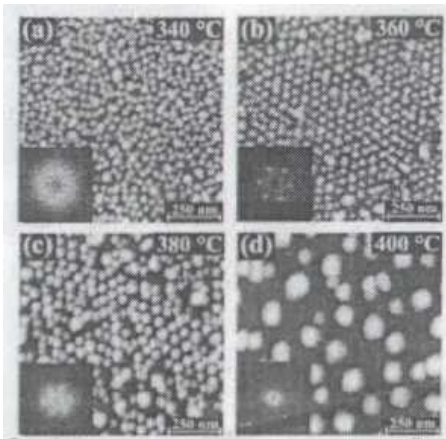


Рис. 18.23. Одержані за допомогою атомно-силової мікроскопії зображення верхнього (незарощеного) шару квантових точок PbSe для надрешітки квантових точок (30 періодів) при номінальній товщині шару PbSe п'ять мономолекулярних шарів для ростових температур від 340°C до 400°C. На вставці результат математичної обробки (спектр потужності) одержаних зображень, видно появу шести чітких піків при появі впорядкування ґратки розташування квантових точок в площині шару (гексагональна ґратка).

ними напруги виявляються дуже малі, щоб "прив'язати" квантові точки в наступному шарі до певних місць. Якщо ж розміри квантових точок перевищують верхній критичний розмір, то, в результаті збільшення розмірів квантових точок і зменшення відстані між місцями на поверхні, де напруги мінімальні, ріст декількох квантових точок в сусідніх мінімумах опиняється вже неможливим і залишається тільки вертикальне впорядкування – точка над точкою (цій ситуації відповідає рис. 18.22 г). Залежність розміру квантових точок, що утворюються, від температури підкладки приводить до схожої поведінки.

Одна з найважливіших задач пов'язана із створенням напівпровідникових лазерів на квантових точках, де використання багатошарових структур і зменшення розкиду розмірів квантових точок дозволяють збільшити ефективність роботи лазера (за рахунок збільшення числа "працюючих" штучних атомів). Крім того, за рахунок наявності зв'язку між шарами квантових точок іноді вдається отримати із структур те, що за звичайних умов одержати не вдається. За рахунок росту двошарових структур квантових точок, в яких перший шар є "приманкою" для другого, і відповідного підбору технологічних параметрів при рості кожного шару вдалося зсунути лінію люмінесценції квантових InAs/GaAs точок в довгохвильову область до 1,4 мкм при кімнатній температурі при збереженні досить високої ефективності люмінесценції (довжини хвиль 13 мкм і 1.55 мкм оптимальні для роботи з оптоволоконними системами передачі інформації).

18.4. Елементи наноелектроніки

Розроблені останніми роками наноелектронні елементи по своїй мініатюрності, швидкодії і споживаній потужності складають серйозну конкуренцію традиційним напівпровідниковим транзисторам і інтегральним мікросхемам на їх основі як головним елементам інформаційних систем. Вже сьогодні техніка впритул наблизилася до теоретичної можливості запам'ятовувати і передавати 1 біт інформації (0 і 1) за допомогою одного електрона, локалізація якого в просторі може бути задана одним атомом.

Напівпровідникові квантові точки, тобто кристали з розмірами в декілька десятків нанометрів, по суті є гігантськими молекулами (що складаються з $10^3 - 10^5$ атомів), сформованими із звичайних напівпровідникових матеріалів, таких, як InAs, Si, InP, CdSe і т.д. Вони істотно менші за традиційні для сучасної електроніки структури з характерними розмірами від мікрона і більше. У багатьох випадках властивості нанокристалів відповідають атомам. Наприклад, нижчий по енергії стан електрона в нанокристалі відповідає s-стану електрона в атомі, а

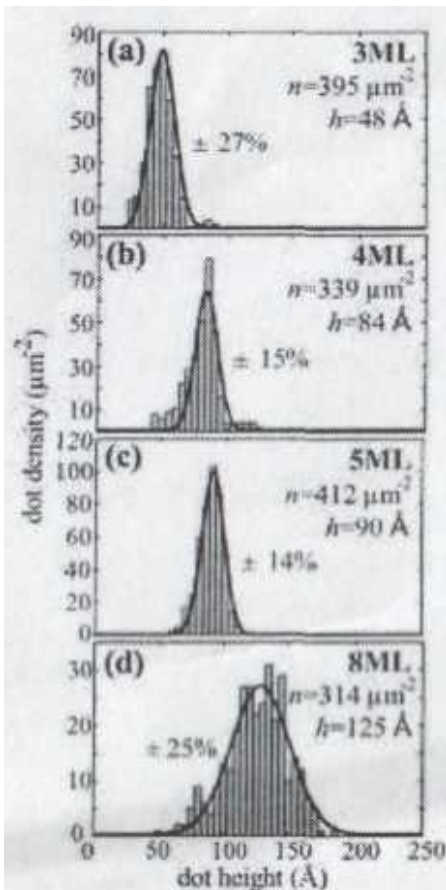


Рис. 18.24. Розподіл квантових точок по висоті залежно від номінальної ростової товщини шару PbSe.

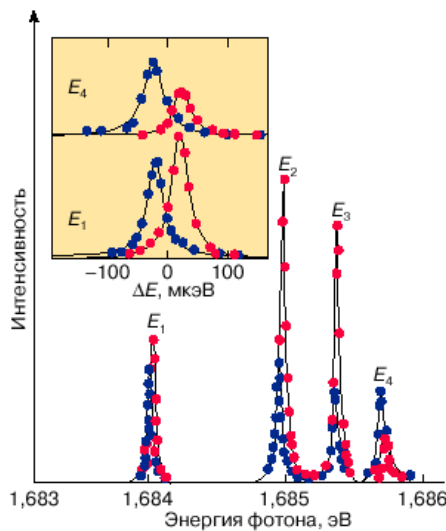


Рис. 18.25. Спектри випромінювання однієї квантової точки. Енергії E_i відповідають переходам між різними рівнями. На вставці показані переходи E_1 і E_4 в більшому масштабі.

наступний – р-стану. Заповнення станів двома або шістьма електронами в квантовій точці найбільш вигідне по енергії, в атомі це відповідає повністю заповненим електронним оболонкам.

18.4.1. Наноласери. Структури з квантовими точками перспективні для створення напівпровідникових лазерів. Електрон в квантовій точці переходить з одного рівня в зоні провідності на іншій у валентній зоні з випуском фотона з енергією $h\nu$, рівній різниці енергій рівнів розмірного квантування. Якщо в структурі з безліччю однакових квантових точок ініціювати узгоджені переходи електронів, то виникає лазерне випромінювання. Електрони на рівні звичайно потрапляють при пропусканні струму через структуру. Оскільки пониження розмірності кристала зменшує ширину зон розмірного квантування, в квантовій точці існують тільки дискретні рівні енергії, як в атомі, це веде до вузьких оптичних спектрів. Як приклад на рис. 18.25 приведений спектр випромінювання від однієї квантової точки GaAs з переходами між чотирма різними рівнями. Окрім надзвичайно вузького спектру випромінювання ще видно поляризацію світла в двох взаємно перпендикулярних площинах.

У структурах напівпровідникових квантових точок відбувається пониження порогового струму лазерної генерації, можливі великі швидкості модуляції випромінювання. Вже одержана лазерна генерація від структур з квантовими точками. Перед ученими зараз стоїть технологічна задача зменшити розкид розмірів квантових точок при рості структур.

18.4.2. Одноелектронні транзистори. Схематичне зображення транзистора з нанокристалом CdSe як активним елементом приведене на рис. 18.26. Звичайно на приготувану структуру осідають нанокристали, і один з них стає активним елементом, як показано на рис. 18.26. Прилад виготовляється на кремнієвій підкладці, на яку можна подати напругу V_g для зміни положення рівня енергії в нанокристалі. Підкладка відокремлена шаром оксиду кремнію SiO_2 від золотих електродів, на які висаджені нанокристали. Один з нанокристалів замикає електроди. Опір такого приладу звичайно складає близько 10 МОм. Для сферичного нанокристала радіусом R , який поміщено в середовище з діелектричною проникністю ϵ , ємність буде $C = \epsilon R$, а потенціал $U = q/C$, де q – електричний заряд. Для R в декілька нанометрів величина ємності складає приблизно 10^{-18} Ф. Якщо помістити в нього один електрон (заряд $1,6 \cdot 10^{19}$ Кл), то потенціал зміниться на $\sim 0,1$ В і збільшується пропорційно $1/R$. Цього потенціалу цілком достатньо, щоб перешкодити руху інших електронів.

На структурі з одиночною квантовою точкою добре демонструються експериментально атомні властивості

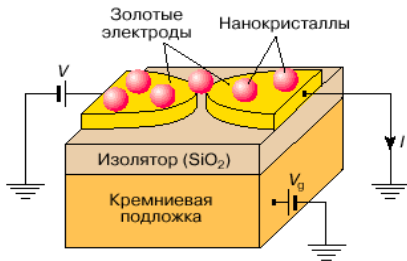


Рис. 18.26. Схематичне зображення одноелектронного транзистора з нанокристалом CdSe як активним елементом.

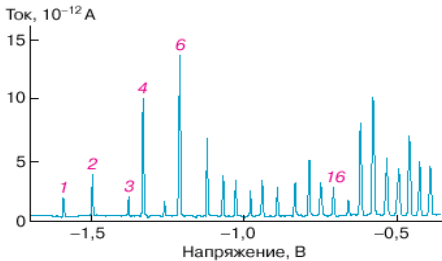


Рис. 18.27. Залежність струму від напруги при послідовному заповненні електронами по одному однієї квантової точки. Числа на піках – число електронів, які пройшли.

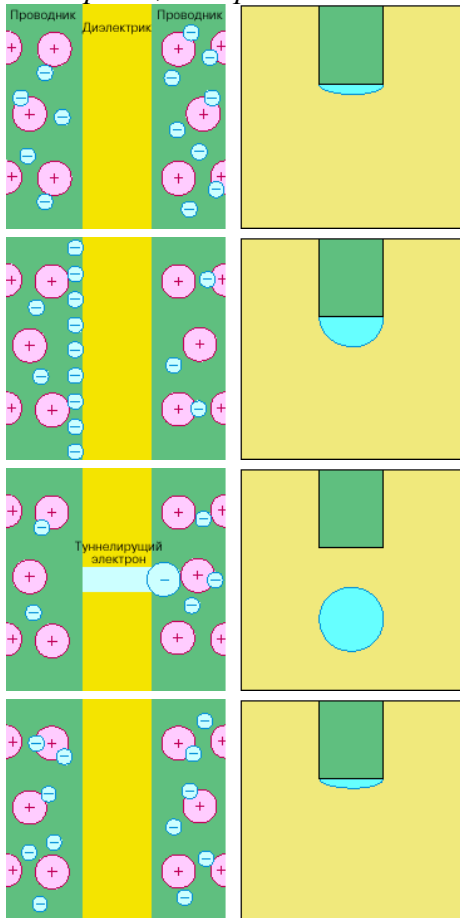


Рис. 18.28. Одноелектронне тунелювання в умовах кулонівської блокади.

квантових точок. Якщо прикласти невелику (декілька мілівольт) напругу V між електродами, то при певній напрузі V_g (див. рис. 18.26) електрон може потрапити в нанокристал. Електрон потраплятиме на рівень в квантовій точці, якщо його енергія співпадає з енергією рівня. Число електронів в квантовій точці можна змінювати поодиночі. При цьому спостерігається сплеск струму, що становить 10-12 А. При зміні V_g далі знову відбудеться сплеск струму. Кількість таких сплесків залежить від числа рівнів розмірного квантування і у принципі визначається розмірами нанокристала. Описаний тут прилад поки добре працює тільки при дуже низьких температурах.

На рис. 18.27 приведена залежність струму, що проходить через квантову точку, від напруги V_g , яка змінює положення рівня енергії в точці. Кожен пік відповідає проходженню одного електрона. Спостерігаються дві істотні особливості: відстані між піками не рівні і величини піків різні. Величини піків визначають енергію, тут найбільша енергія відповідає двом і шести електронам, як в атомі при заповненні електронним оболонкам відповідає число електронів два або вісім в атомах гелію і неону, в точці повністю заповненим станам двох рівнів відповідають шість електронів). Взаємодія електронів призводить до того, що їх енергія (положення рівня) залежить від числа електронів, що знаходяться на рівнях розмірного квантування, що приводить до зміни відстані між піками на рис. 18.27. Без взаємодії рівні повинні бути розташовані на однаковій відстані.

З'явилися вже повідомлення про створення одноелектронної пам'яті, що працює при кімнатній температурі. Прилад заснований на транзисторі, в якому один електрон, захоплений в нанокристал, приводить до замикання провідності вузького каналу транзистора.

18.4.3. Одноелектронне тунелювання. Іншим специфічним проявом квантового обмеження є одноелектронне тунелювання в умовах кулонівської блокади (рис. 18.28). Щоб пояснити цей термін, розглянемо ілюстрований рис. 18.28 приклад проходження електронем структури метал – діелектрик – метал. Як наочна ілюстрація паралельно проводиться аналогія з краплею, що відривається від краю трубки. Первинно межа розділу між металом і діелектриком електрично нейтральна. При поданні до металевих областей потенціалу на цій межі починає накопичуватися заряд. Це продовжується до тих пір, поки його величина не виявиться достатньою для відриву і тунелювання через діелектрик одного електрона. Після акту тунелювання система повертається в первинний стан. При збереженні зовнішньої прикладеної напруги все повторюється знову.

Таким чином, перенесення заряду в такій структурі здійснюється порціями, рівними заряду одного електрона. Процес же накопичення заряду і відриву електрона від межі металу з діелектриком визначається балансом сил кулонівської взаємодії цього електрона з іншими рухомими і нерухомими зарядами в металі.

Транзистори на резонансному тунелюванні являють собою двохбар'єрний діод на квантових ямах, у якого потенціал ям і відповідні резонансні умови контролюються третім електродом. Ці транзистори мають частоти перемикання порядку 10^{12} Гц, що в 100 – 1000 разів вище, ніж у найкращих кремнієвих транзисторів з сучасних інтегральних мікросхем. Є пропозиції по створенню на таких транзисторах елементів статичної пам'яті і інших елементів для обчислювальних систем.

Лекція 19

СОНЯЧНІ ЕЛЕМЕНТИ НА ОСНОВІ СЕНСИБІЛІЗОВАНИХ БАРВНИКАМИ НАПІВПРОВІДНИКІВ

План

Вступ

1. Принцип дії барвникових-сенсibilізованих сонячних елементів
 2. Матеріали для барвникових-сенсibilізованих сонячних елементів
 3. Барвник – сенсibilізатор
- Висновки

Література

19.1. Б.К. Остафійчук, І.Ф. Миронюк, В.О. Коцюбинський, В.Д. Глуханюк, В.Л. Челядин, Н.І. Нагірна Сонячні елементи на основі сенсibilізованих барвниками напівпровідників (огляд) // Фізика і хімія твердого тіла. Т. 9, № 1 (2008) с. 11-18.

Вступ

Швидкий розвиток фотовольтаїки, як окремого прикладного розділу фізики, визначається нагальною необхідністю розробки альтернативних джерел енергетичного забезпечення. Надзвичайно важливим питанням є екологічна чистота енергетичних технологій, що зумовлює підвищену увагу до відновлювальних джерел енергії, зокрема сонячної.

Не дивлячись на те, що сонячна енергетика має величезний потенціал і найширші перспективи, на сьогодні <1% виробленої у світі електроенергії є «сонячного» походження.

Основою найпоширеніших на сьогодні комерційних фотоелектричних пристроїв є твердотільні монокристалічні кремнієві сонячні елементи з р-п-переходами. Добре відпрацьована технологія отримання та обробки монокристалічного кремнію дозволяє очікувати збереження ключових позицій для сонячних елементів на його основі в найближчому майбутньому. Основними конкурентами для монокристалічних кремнієвих елементів донедавна були високоефективні фотоперетворювачі базовані на твердотільних багат шарових структурах GaInP, GaAs, полікристалічні плівки аморфного кремнію (a-Si), селеніди міді та індію, телурид кадмію. Досягнуто значних успіхів у підвищенні коефіцієнта корисної дії сонячних елементів, зокрема для лабораторних зразків тандемного елемента на базі GaInP/GaAs/Ge перейдено бар'єр 30%. Порівняльна діаграма максимальних ефективностей сонячних елементів з використанням різних матеріалів та технологій представлена на рис. 19.1.

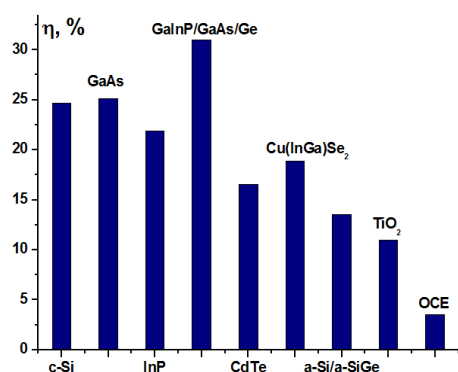


Рис. 19.1. Максимальні коефіцієнти корисної дії η сонячних елементів, що базуються на різних матеріалах та технологіях.

Широкому застосуванню фотоелектричних перетворювачів перешкоджають, в першу чергу, висока

вартість технологій отримання монокристалічних плівок кремнію та полікомпонентних багатошарових напівпровідникових структур. Комерційні елементи на основі монокристалічного кремнію володіють ККД на рівні 15-20%, теоретична межа для матеріалів a-Si:H становить 16%.

Технічну і економічну альтернативу існуючим фотовольтаїчним пристроям можуть в найближчому майбутньому скласти елементи, принцип дії яких ґрунтується на фотоелектрохімічних методах фотон-електричного перетворення. Принципово нові концепти, такі як органічні сонячні елементи (ОСЕ) та барвникові-сенсibilізовані сонячні елементи (СБСЕ) завойовуватимуть ринок, оскільки їх технології володіють перспективою значного здешевлення та спрощення, і перебувають на стадії фундаментальної розробки. Зокрема, технологія СБСЕ стрімко прогресує і близька до комерціалізації.

19.1. Принцип дії СБСЕ

СБСЕ представляють собою наноінженерний пристрій в якому, завдяки вдалому поєднанню властивостей матеріалів, вдалося реалізувати фотовольтаїчне перетворення, багато в чому близьке до фотосинтезу. Застосування недорогих технологій та матеріалів відкриває можливість широкого виробництва та охоплення ринку. Величезний прогрес, досягнутий останнім часом у розробці нанодисперсних матеріалів заданої морфології, дозволяє прогнозувати збільшення ефективності та підвищення стабільності характеристик СБСЕ.

Ідея фотоелектрохімічного елемента на основі сенсibilізованого нанодисперсного діоксиду титану була запатентована в 1990 році, і розвинута в серії робіт. В якості робочого матеріалу для СБСЕ застосовується сенсibilізований барвником нанодисперсний оксидний напівпровідник; поглинання випромінювання та транспортування фотогенерованих носіїв відбувається у різних функціональних середовищах, що дозволяє гнучко керувати цими процесами.

Схема компонування СБСЕ та діаграма електронних переходів в ньому представлені відповідно на рис. 19.2 та рис. 19.4. На рис. 19.3 представлено електронно-мікроскопічне зображення плівки TiO_2 , її наноструктури та схема хемосорбції барвника.

Оксидна плівка нанокристалічної морфології (питома поверхня 60-150 м²/г, типова пористість 50-65%, середній розмір частинок 10-30 нм, середній розмір пор 15 нм), покрита моношаром органічного барвника, характеризується широкою областю поглинання (від УФ до ближньої ІЧ областей спектру) і дозволяє захопити 70-90% фотонів сонячного випромінювання.

Поглимаючи фотон, молекула барвника-сенси-

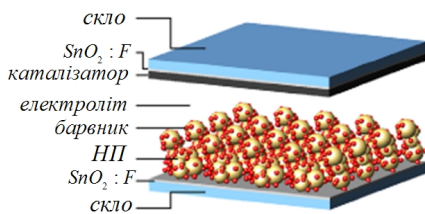


Рис. 19.2. Схема компонування СБСЕ.

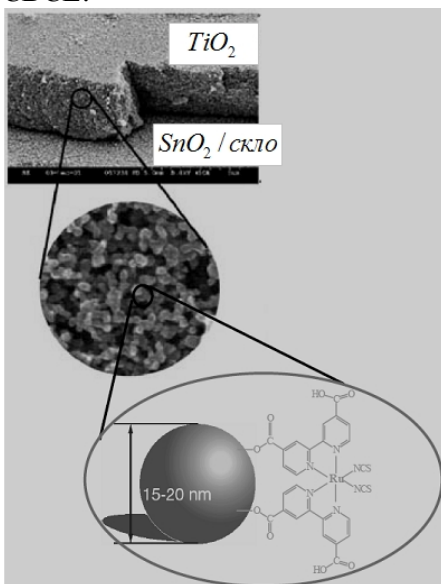


Рис. 19.3. Зображення плівки TiO_2 (трансмісійна електронна мікроскопія) з демонстрацією морфології частинок та схемою хемосорбції поверхню TiO_2 молекули біпіридилного рутенієвого барвника.

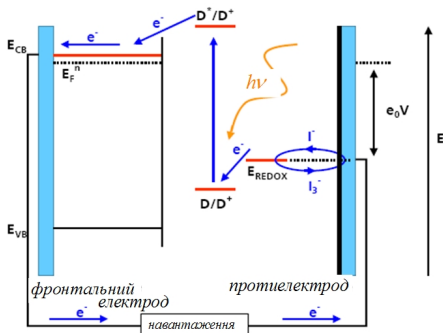
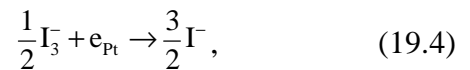
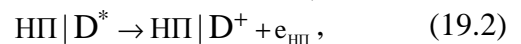


Рис. 19.4. Енергетична діаграма СБСЕ.

білізатора, хемосорбованого на поверхні нанодисперсного широкозонного оксидного напівпровідника (НП) n-типу, переходить в збуджений стан, що відповідає переносу електрона з найвищої заповненої молекулярної орбіталі на найнижчу незаповнену орбіталь. Релаксація збудженого стану молекули барвника відбувається через інжекцію електрона в зону провідності НП, який виступає в ролі транспортного середовища. З НП електрони переходять на колектор зарядів – фронтальний електрод, яким служить скляна пластинка, покрита провідним прозорим оксидом ($\text{SnO}_2\text{: F}$, $\text{SnO}_2\text{: Sb}$, In_2O_3). Омичний контакт між провідним шаром на внутрішній поверхні фронтального електрода та мезопористою плівкою НП досягається шляхом прожарювання при температурі 350-450 °С.

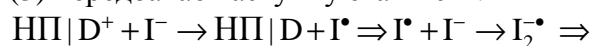
Перенос заряду в зоні провідності НП відбувається шляхом дифузії електронів до фронтального електрода. Окислені молекули барвника відновлюються шляхом переходу електрона з електроліту. Реакція відбувається поблизу протиелектрода, покритого шаром каталізатора (Pt, графіт), на який через зовнішнє навантаження, замикаючи коло, переходить електрон з фронтального електрода. Різниця потенціалів, що генерується при освітленні дорівнює різниці між рівнем Фермі для нанодисперсного оксиду і редокс-потенціалом електроліту. Найпоширеніший тип електроліту – органічний розчинник, що містить редокс-пари йодиту та трийодиту I^-/I_3^- .

Таким чином, робочий цикл СБСЕ відбувається без будь-яких необоротних хімічних реакцій і може бути описаний наступними рівняннями:



де $\text{НП}|\text{D}$, $\text{НП}|\text{D}^*$ та $\text{НП}|\text{D}^+$ позначено хемосорбовану на поверхні НП молекулу барвника, відповідно, у основному, збудженому та окисленому станах; $e_{\text{НП}}$ та e_{Pt} позначено електрони провідності плівки TiO_2 , та каталізатора – протиелектрода. Концентраційна рівновага іонів I^-/I_3^- відновлюється дифузійним шляхом – негативний заряд I^- мігрує від каталізатора до фронтального електрода і, відновлюючи молекулу барвника D^+ , замикає електричне коло.

Механізм двоелектронного відновлення молекул барвника (3) передбачає наступну етапність:



Утворення радикалу $\text{I}_2^{\bullet -}$ є термодинамічно вигідніше

порівняно з реакцією утворення молекули I_2 , за умов наявності значної кількості пар (I, I^-) та (D^+, I^-) на границі розділу барвник – електроліт.

Можливий варіант заміни розчину електроліту широкозонним неорганічним напівпровідником р-типу (CuI чи CuSCN), або ж іншим твердотільним середовищем з дірковим типом електропровідності (аморфні органічні арилами́ни, їх похідні). В пристроях такого типу збуджені молекули барвника окислюються, інжектуючи електрони в напівпровідник n-типу і відновлюються шляхом переходу дірки в матеріал з р-типом провідності.

Процес генерації струму супроводжується наступними втратами:

- збуджений стан молекули може релаксувати в основний без інжекції електрона в НП. Цей механізм є малоймовірним, оскільки час життя збудженого стану в 10^3 разів перевищує час необхідний для інжекції.
- електрон з зони провідності НП може рекомбінувати з окисленою молекулою барвника, перед тим, як барвник окислюється електролітом – $NP|D^+ + e_{nnp} \rightarrow NP|D$, проте реакція відновлення з електролітом відбувається в близько 10^3 разів швидше.
- найбільші втрати в СБСЕ зумовлюються рекомбінацією електронів в НП з дірками електроліту, тобто $I_3^- + 2e_{nnp} \rightarrow 3I^-$.

Електронний перехід дифузійним шляхом в НП і відновлення I_3^- відбувається одночасно, тому кількість іонів I_3^- в електроліті повинна бути мінімальною, проте достатньою для наявності достатньої кількості партнерів в актах рекомбінації для електронів поблизу протиелектрода. Проте, опір зарядового переходу поблизу протиелектрода зменшений на багато порядків впливом каталізатора, в той час, як зарядовий перехід НП/електроліт для оксиду та редоксних пар I^-/I_3^- є значно важчий з енергетичної точки зору.

19.2. Матеріали для СБСЕ

На сьогодні в СБСЕ на сьогодні застосовуються такі широкозонні напівпровідники як TiO_2 , ZnO , SnO_2 , Nb_2O_5 ; окрім того ведеться широкий пошук композитних матеріалів на основі цих оксидів, а також Al_2O_3 , SiO_2 , вуглецевих наноструктур.

Перевагами діоксиду титану в порівнянні з іншими матеріалами, застосованими в СБСЕ, є хімічна стійкість, нетоксичність, біосумісність, низька вартість. Його особливістю окрім того є значна фотоактивність та яскраво виражена залежність електричних властивостей від морфології поверхні і типу кристалічного впорядкування. В природі діоксид титану кристалізується в трьох поліморфних модифікаціях: анатаз, рутил та брукіт, які є широкозонними напівпровідниками з

валентною зоною, сформованою 2p станами кисню і зоною провідності, що містить 3d стани титану. Завдяки поєднанню електричних та оптичних властивостей, для фотовольтаїки найбільший інтерес представляє саме анатаз, метастабільна фаза, яка необоротно перетворюється в рутил при нагріванні вище критичної температура (400-1200°C, залежно від розміру та морфології частинок, складу атмосфери, типу та концентрації домішок). Дефекти кисневої підгратки (двозарядні кисневі вакансії V_{O}^{**}) викликають збільшення провідності n-типу матеріалу пропорційно до величини втрат кисню. Кисневі вакансії формують як «мілкі» (0 – 200 meV), так і «глибокі» (600 – 750 meV) донорні рівні; для анатазу енергія активації носіїв становить близько 4 меВ. Коефіцієнт дифузії фотоелектронів в анатазі різко залежить від інтенсивності освітленості, змінюючись в межах 10^{-11} - 10^{-9} м²·с⁻¹, причому фотоелектрони як накопичуються на поверхневих станах, так і переходять в зону провідності, причому концентрація носіїв залежить від розмірів наночастинок та товщини плівки.

Ширина забороненої зони для анатазу становить 3,1 – 3,2 еВ і прямозонні електронні переходи збуджуються при поглинання випромінювання з довжиною хвилі < 390 нм. Таким чином, за відсутності сенсibilізатора анатаз поглинає < 4 % випромінювання спектру Сонця. Для СБСЕ поглинання ультрафіолетового випромінювання небажане, оскільки TiO_2 з дірками у валентній зоні є хімічно високоактивним, що унеможливує довготривалу експлуатацію елемента.

Анатазна модифікація TiO_2 отримується золь-гель методом, що базується на гідролізі тетрахлориду чи ізопропоксиду титану. Можливе використання комерційного TiO_2 , зокрема пірогенного анатазу P25 (Degussa AG) з середнім розміром частинок 25 нм. Для попередження агрегатації частинок застосовується ацетилацетон чи оцтова кислота; для забезпечення однорідності товщини при нанесенні плівки використовують неіонні детергенти (Triton X-100, етилцелюлоза).

Найбільш поширений та використовуваний сьогодні метод підготовки фотоелектродів полягає в нанесенні на прозору, провідну підкладку (скло/ $SnO_2:F$, металева фольга, полімерна плівка) колоїдальної пасти TiO_2 , в'язкість якої становить 15-20 мПа·с. Застосовуються два варіанти формування плівки: механічне нанесення та трафаретний друк на підкладці. Типова товщина плівки становить 5-15 мкм, поверхнева густина 1-2 мг/см². Оптимальна товщина плівки TiO_2 становить 10-15 мкм. Після нанесення колоїдного розчину TiO_2 на підкладку, застосовують додаткову обробку плівки тетрахлоридом титану, що сприяє підвищенню величин струму короткого замикання $j_{кз}$ та напруги відкритого кола $U_{вк}$ СБСЕ. Для

досягнення омичного контакту між окремими наночастинками матеріалу та провідною скляною підкладкою застосовують, відпал нанесеного покриття при температурі 400-500 °С.

Крім вищезгаданих методів пропонується осадження розпиленого прекурсора $Ti(OCH(CH_3)_2)_4$ безпосередньо на поверхню підкладки при температурі 450 °С. В роботі для осадження TiO_2 застосовано метод магнетронного розпилення, що дозволило отримати плівки впорядкованої колонарної структури з розвинуеною поверхнею та високими значеннями електропровідності; метод пропонується авторами як можливий промисловий варіант виробництва СБСЕ. Автори зафіксували ефективність конверсії на рівні 4,1% для СБСЕ на основі електрофоретично нанесених плівок прожарених при температурі 150 °С. Активно ведуться розробки технології гнучких СБСЕ, що дозволить зменшити товщину, вагу та розширити області застосування фотоперетворюючих пристроїв. Нещодавно повідомлено про варіанти низько-температурного способу нанесення плівки на полімерні підкладки при температурах нижче 150 °С та застосування механічного тиску для забезпечення омичного контакту між наночастинками TiO_2 . Група розробила високоефективний ($h = 7,2\%$) СБСЕ на основі окисленої титанової фольги в якості фотоанода, що є на сьогодні найвищим показником для СБСЕ на гнучкій основі.

Висока питома поверхня плівки забезпечує можливість хемосорбції значного числа молекул барвника та збільшення коефіцієнта поглинання випромінювання, проте водночас збільшується ймовірність рекомбінаційних процесів переходу електронів з зони провідності TiO_2 до акцепторів в електроліт. Для зменшення ймовірності такого переходу застосовують багатошарові композитні структури, в яких між TiO_2 та барвником формується буферна область, в якості якої виступає оксидний напівпровідник (ZnO , Nb_2O_5 , SrO , $SrTiO_3$, Al_2O_3 , ZrO_2 , SiO_2) з шириною забороненої зони більшою, порівняно з TiO_2 чи шар електропровідного полімерного матеріалу.

Вдосконалення технології формування напів-провідникових електродів є ключовою проблемою для подальшого збільшення ККД СБСЕ і можливості їх широкого комерційного застосування. Основною метою тут є розробка підходів для отримання функціональних матеріалів з однотипними за розміром та морфологією, слабоагрегативними частинками. Плівка оптимальної морфології повинна володіти системою паралельних між собою та перпендикулярних до площини фронтального електрода каналів.

19.3. Барвник – сенсibilізатор

Барвник для успішного застосування в якості фотосенсibilізатора поверхні НП повинен характеризуватися високими значеннями коефіцієнта поглинання

світла в видимій та ближній інфрачервоній областях спектру сонячного випромінювання ($\lambda \leq 920$ нм), близькою до одиниці ефективністю фотон-електронного перетворення, бути радіаційно- і температуростійким та здатним до циклічного оборотного процесу відновлення-окислення ($> 10^7$ - 10^8 циклів окислення-відновлення без деградації). Основна вимога – енергетичні положення найвищої зайнятої та найнижчої незайнятої молекулярних орбіталей сенсibilізатора повинні бути, відповідно, вищими та нижчими порівняно з положеннями дна зони провідності НП та редокс-потенціалу електроліту.

Висока ефективність фотон-електронного перетворення в СБСЕ визначається різницею швидкостей перебігу різних редокс-процесів, які відбуваються при освітленні сенсibilізованого електрода. Надшвидка інжекція фотоелектрона в зону провідності НП відбувається в фемто-пікосекундному діапазоні (10^{-15} - 10^{-12} с). Відновлення окисленого барвника типово характеризується швидкістю перебігу 10^5 - 10^7 с $^{-1}$, що більш ніж в 100 разів перевищує швидкість рекомбінації інжекттованих електронів з окисленими редокс-частинками електроліту і на порядок швидше зворотного переходу електрона безпосередньо до катіону барвника, що веде до ефективного накопичення фотоелектронів в НП (рис. 19.5).

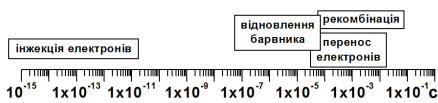


Рис. 19.5. Діаграма тривалості перебігу основних етапів електронного переносу в СБСЕ.

Умовно сенсibilізатори, що застосовуються в СБСЕ можна розділити на дві групи: органічні та неорганічні барвники. Неорганічні барвники включають поліпіридилні комплекси рутенію та осмію, порфірини, фталоціаніни. До органічних можна віднести природні та синтетичні антоціаніни, хлорофіли, які порівняно з неорганічними менш стабільні термально та хімічно.

Найбільш ефективними сенсibilізаторами нанокристалічного TiO_2 , як за величиною квантового виходу так і з міркувань стабільності характеристик, залишаються біпіридилні хелатні комплекси перехідних металів (найчастіше Ru(II) , Ru(III) , Os(III)) центральний іон яких оточений нейтральними і(чи) аніонними лігандами (рис. 19.6). Найбільш розповсюдженою геометрією є формування октаедричного оточення центрального іона. Згідно з теорією кристалічного поля, дана властивість сполуки є наслідком розщеплення d-орбіталей в електростатичному полі катіонів металу та лігандних аніонів. Наявність π -зв'язків необхідно пояснювати з точки зору теорії молекулярних орбіталей, відповідно до якої взаємодія між металом та лігандами описується лінійною комбінацією атомних орбіталей і передбачає електронний обмін між центральним іоном та лігандами. Взаємодія між орбіталями залежить від різниці в їх енергіях та ступеня їх перекриття. Для сполук перехідних металів в утворенні хімічних зв'язків беруть участь s, p і d-орбіталі металу та орбіталі σ - та π -типу,

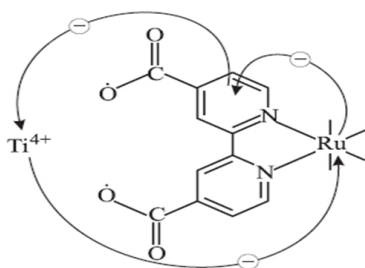


Рис. 19.6. Схема електронних переходів сенсibilізатор $-\text{TiO}_2$ в СБСЕ.

якими здійснюється зв'язок центрального іона з лігандами. Відновленому та окисленому станам молекули барвника відповідають енергії найвищої заповненої (НОМО) та найнижчої незаповненої (LUMO) молекулярних орбіталей. Перевагою рутенієвих комплексів є широка смуга поглинання у видимій області спектру внаслідок фотоіндукованих електронних переходів іон металу/ліганд, що веде до появи збуджених станів молекули барвника час життя яких становить десятки пікосекунд.

Керуючи хелатоутворюючою реакцією приєднання амфідентатних лігандів CN⁻ чи -SCN до іона металу, можна впливати на спектральну чутливість барвника. Вибір периферійних груп, крім того, визначає можливість хемосорбції барвника на поверхні наночастинки TiO₂, що є критично важливим для успішної фотосенсибілізації. Заміщуючі додаткові групи на 4,4'-позиціях біпіридилного кільця утворюють зв'язок молекули барвника з поверхнею НП та служать містками для переходу електронів з донорних рівнів барвника в зону провідності TiO₂. Сенсибілізатор утворює зв'язки з поверхнею напівпровідника через карбоксилатні чи фосфонатні групи, які спонтанно реагують з гідроксильними групами на поверхні НП. Поліпіридилні рутенієві сенсибілізатори діляться на карбоксилатні, фосфонатні та полядерні барвники в залежності від типу адсорбційної зв'язуючої групи. Крім того останні характеризуються наявністю кількох іонів металів.

Для фотосенсибілізації TiO₂ на перших етапах розробки СБСЕ використовувався комплекс Na₂[(NC)Ru(bpy)₂(CN) Ru(dcb)₂(NC)Ru(bpy)₂CN)] (dcb позначає 2,2'-біпіридил 4,4'-дікарбоксильну кислоту, bpy – 2,2'біпіридилні ліганди.

Найкращі результати при застосуванні в СБСЕ продемонстрував карбоксилатний поліпіридилний барвник N3 – комплекс цис-дитіоціанато-біс (2,2'-біпіридил-4,4'-дікарбоксилат) Ru(II) та його тетрабутил-амонієва сіль N719 – «black dye» (спектральна залежність коефіцієнта фотон-електронного перетворення β представлено на рис. 19.7) – ККД СБСЕ перевищив значення 11%.

Даний барвник володіє високою стабільністю, забезпечуючи мінімум 10⁸ циклів окислення – відновлення без деградації властивостей. Хемосорбція на поверхні TiO₂ забезпечується карбоксилатними групами, в той час як NCS-групи забезпечують широку смугу поглинання у видимій області спектру що володіє максимумом на 540 нм та простягається в довгохвильову область до 750 нм.

При підборі барвника важливим фактором є розчинність барвника в водних та органічних електролітах. Недоліком карбоксилатних поліпіридилні рутенієвих барвників є легка десорбція з поверхні

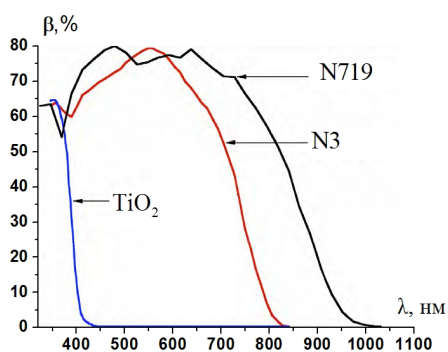


Рис. 19.7. Спектральна залежність коефіцієнта фотон-електронного перетворення β.

напівпровідника в водних розчинах при $pH < 5$, в той час як хемосорбція фосфонатних барвників є стійкішою, проте їх будова ускладнює електронну інжекцію в зону провідності TiO_2 , тому значна кількість досліджень присвячена модифікації електронної структури фосфонатних барвників. Авторами отримано фосфонатний поліпіридинний рутенієвий комплекс з часом життя збудженого стану 15 нс та з коефіцієнтом фотон-електронного перетворення близьким до 70%. Максимальну ефективність конверсії для СБСЕ з використанням фосфонатних поліпіридинних рутенієвих комплексів (8%) отримано в.

Використання полядерних комплексів дозволяє підвищити коефіцієнт поглинання, проте порівнянно більший розмір комплексу ускладнює хемосорбцію на нанопористій поверхні напівпровідника і, як наслідок, зменшується коефіцієнт фотон-електронного перетворення.

Велику увагу в якості сенсibilізатора викликають порфірини і фталоціаніни. Основними проблемами тут виступають незручність енергетичної позиції найнижчої незаповненої молекулярної орбіталі, яка є надто низькою для електронного переходу до зони провідності TiO_2 . Інша стратегія, спрямована на розширення області оптичного поглинання полягає в комбінації кількох барвників – порфіринів і фталоціанінів. Ефективність процесу фотон-електронного перетворення при застосуванні в якості сенсibilізатора фталоціанінів (Pc) порівняно невисока, крім того висока ймовірність фотокорозії напівпровідника. Дослідження фотоактивності матеріалу TiO_2/Pc (PcMg, PcZn, PcAl(Cl), PcTi(O), PcCo, PcFe) показали, що основну роль відіграє морфологія поверхні, проте загалом $J_{кз}$ не перевищує 2-4 mA/cm^2 . Найкращі результати при застосуванні фталоціанінового барвника були отримані авторами, які досягнули значення струму короткого замикання $J_{кз}$ близько 10 mA/cm^2 при випромінюванні потужністю 100 mW/cm^2 застосовуючи карбоксилатні чи сульфонатні фталоціанінові комплекси Zn та Al(OH).

Перевагами застосування в якості сенсibilізатора органічних барвників є високий коефіцієнт поглинання внаслідок міжмолекулярних $\pi-\pi^*$ -переходів та дешевизна (відсутність дорогих компонентів типу Ru). Цей напрямок швидко розвивається і ефективність конверсії для СБСЕ з застосуванням органічних барвників вже порівняна з елементами, побудованих з використанням поліпіридинних рутенієвих сенсibilізаторів. Типовий органічний сенсibilізатор має загальну структуру: «донор – π -зв'язок – акцептор», його область поглинання зміщена до червоної частини спектру, що зумовлює підвищення поглинання світла і, відповідно, струму короткого замикання комірки.

Авторами розроблено сенсibilізатори на основі похідних кумарину і досягнуто ККД на рівні 7,6%. Групи

дослідників використовуючи барвник індолін досягнули значень η на рівні, відповідно, 8,0 та 9,0 %.

Пропонується застосування інших органічних барвників, зокрема хлорофілу, застосування якого дозволило отримати значення ККД 2,6%.

В роботі в якості органічного барвника-сенсibilізатора застосовано натуральний антоціаніновий барвник ціанін, абсорбційні властивості якого можна змінювати, зсуваючи спектр поглинання в довгохвильову сторону шляхом утворення комплексів з поліфенолами, пектинами та іонами Al^{3+} , Fe^{3+} , Cr^{3+} , Ti^{4+} , Sn^{4+} . Адсорбція ціаніну поверхнею TiO_2 – реакція заміщення групи ОН поверхневого іону Ti^{4+} на молекулу ціаніну, що втрачає протон; в результаті утворюється H_2O та стійкий комплекс. Згідно з квантово-механічними розрахунками для комплексу ціанін- Ti^{4+} електрони найвищої заповненої молекулярної орбіталі локалізовані на хромофорній частині комплексу, в той час як електронна густина найнижчої незаповненої молекулярної орбіталі локалізована поблизу іону металу.

Процеси електронного переносу в СБСЕ активно вивчаються методами фотоелектрохімії, імпедансної та оптичної спектроскопії. Відновлення барвника та зворотній електронний перехід НП / барвник досліджуються методами піко- та фемтосекундної лазерної спектроскопії.

ДЛЯ ПОДАТОК

ДЛЯ ПОДАТОК

Навчальне видання

ФРЕЙК Дмитро Михайлович
ЧОБАНЮК Володимир Михайлович
ГОТРА Зенон Юрійович
та інші

Фізика процесів у напівпровідниках та елементах електроніки

Курс лекцій

Редактор *Фреїк Д.М.*
Технічний редактор *Дзундза Б.С.*
Комп'ютерна верстка *Ткачук А.В.*
Дизайн обкладинки *Туровська Л.В.*

Підписано до друку 25.02.10. Формат 60×⁸⁴/₈.
Папір офсетний. Гарнітура „Times New Roman”.
Умовних аркушів 30,8. Тираж 200. Зам. 148.

ISBN 978-966-640-275-5

Видавець
Прикарпатський національний університет імені Василя Стефаника
76000, м. Івано-Франківськ, вул. С. Бандери, 1, тел. 71-56-22.
Свідоцтво суб'єкта видавничої справи
Серія ДК 2718 від 12,12,2006