

Розділ 13

ЯВИЩА НА ПОВЕРХНІ НАПІВПРОВІДНИКІВ

До поверхневих явищ належать такі фізичні процеси, що відбуваються на межі розподілу між твердим тілом та оточуючим середовищем. Явища на межі розподілу **напівпровідник-вакуум** та **напівпровідник-газове середовище** становлять великий клас фізичних і фізико-хімічних ефектів, що широко використовуються у сучасних технологіях, та є об'єктом прискіпливої уваги фахівців у сучасній фізики напівпровідників. У розд. 8 було з'ясовано, що порушення періодичності структури кристалу, що є наслідком наявності домішок і дефектів, спричиняє появу у забороненій зоні напівпровідника додаткових енергетичних рівнів, тобто порушення періодичності структури є причиною модифікації енергетичної структури кристалу. Але ж поверхня є порушенням періодичності структури кристалу, наявність якої також має спричиняти зміни в енергетичній структурі напівпровідника. Таким чином, на поверхні напівпровідника відбуваються специфічні явища, що породжуються зміною енергетичної структури електронних станів. Саме такі явища розглядаються у цьому розділі.

■ 13.1. Поверхневі стани

Який внесок до формування енергетичних спектрів дає наявність межі розподілу **кристал-зовнішнє середовище**? Розглянемо одновимірну модель потенціального поля із межею розподілу при $x = 0$. Припустимо, що ліворуч границі напівнескінченного одновимірного кристала значення потенціалу є сталою величиною, а праворуч на електрон діє кристалічний періодичний потенціал

$$U(x) = \begin{cases} U_0, & x < 0 \\ U(x) = U(x + a), & x > 0 \end{cases} \quad (13.1)$$

Рівняння Шредингера у цьому випадку мають вигляд

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi_1(x)}{dx^2} + U_0 \psi_1(x) = E \psi_1(x), \quad x < 0, \quad (13.2)$$

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi_2(x)}{dx^2} + U(x) \psi_2(x) = E \psi_2(x), \quad x > 0. \quad (13.3)$$

При $E < U_0$ розв'язок рівняння (13.2) можна записати у вигляді

$$\psi_1(x) = Ae^{\eta x} + Be^{-\eta x}, \quad \eta = \sqrt{\frac{2m(U_0 - E)}{\hbar}}. \quad (13.4)$$

Зрозуміло, що при $x \rightarrow -\infty$ другий доданок у (13.4) необмежено зростає. Тому коефіцієнт B варто занулити. Можна вважати, що вплив границі на рух електрона за великих позитивних x буде слабким, тому розв'язок рівняння (13.3) матиме вигляд функції Блоха

$$\psi_2(x) = e^{\pm i k x} \varphi_{\pm k}(x). \quad (13.5)$$

У необмеженому кристалі хвильовий вектор є величиною дійсною, оскільки в іншому разі блохівська функція буде необмеженою на нескінченності. Але за наявності границі значення аргументу $x < 0$ не реалізуються – умови дійсності вектора k можуть бути зняті. Таким чином, необхідно припустити, що хвильовий вектор є комплексною величиною, тобто

$$k = k' + i k''. \quad (13.6)$$

Тоді формула (13.5) набуде вигляду

$$\psi_2(x) = Ce^{i(k' + i k'')x} \varphi_k(x) + De^{-i(k' + i k'')x} \varphi_{-k}(x) \quad (13.7)$$

Якщо коефіцієнт $D \neq 0$, то при $x \rightarrow \infty$ $\psi_2 \rightarrow \infty$, тобто для отримання фізично визначеного розв'язку необхідно покласти $D = 0$. Для визначення коефіцієнтів A та C запишемо умови неперервності ψ та ψ'

$$\psi_1(0) = \psi_2(0) \Rightarrow A = C \varphi_k(0), \quad (13.8)$$

$$\psi_1'(0) = \psi_2'(0) \Rightarrow \eta \cdot A = C [i k \varphi_k(0) + \varphi_k'(0)]. \quad (13.9)$$

Звідси маємо

$$\eta \varphi_k(0) = i k \varphi_k(0) + \varphi_k'(0). \quad (13.10)$$

Це рівняння дає

$$\eta = i k + \frac{\varphi_k'(0)}{\varphi_k(0)} = i k + [\ln \varphi_k(0)]'. \quad (13.11)$$

Логарифм хвильової функції є комплексним, тому позначимо

$$[\ln \varphi_k(0)]' = \lambda + i \mu. \quad (13.12)$$

Тоді із (13.11) маємо

$$\eta = ik' - k'' + \lambda + i\mu, \quad (13.13)$$

або $\eta = \lambda - k'', k' = -\mu. \quad (13.14)$

Використовуючи визначення η (формула (13.4)), запишемо для енергії електрона в обмеженому кристалі

$$E(k) = U_0 - \frac{\hbar^2(k'' - \lambda)^2}{2m}. \quad (13.15)$$

При $k'' = 0$ енергія $E(k) = E(k')$ описує деяку зонну структуру, оскільки $\lambda = \text{Re}[\ln \varphi_k(0)]'$ є функцією хвильового вектора, тобто k' . При $k'' \neq 0$ енергія виявляється квадратичною функцією k'' за будь-якого значення λ . Це означає що при $k'' \neq 0$ з'являються такі енергетичні рівні, яких не було в необмеженому кристалі. Цим станам відповідають хвильові функції

$$\psi_1(x) = Ce^{ik'x} \varphi_k(x) e^{-k''x}, \quad x > 0, \quad (13.16)$$

$$\psi_2(x) = Ae^{\eta x}, \quad x < 0, \quad (13.17)$$

що як експоненти спадають по обидва боки границі, тобто ці стани локалізовані у пограничному шарі із товщиною $\sim 1/k''$. Ці стани називаються поверхневими, а додаткові енергетичні рівні – поверхневими таммовськими рівнями. Рівні, що виникають у результаті обмеження кристала, мають перебувати у забороненій зоні. Це дуже просто можна пояснити, розглядаючи наявність поверхні (тобто границі, де обривається періодична кристалічна структура) в одновимірному кристалі із взаємодією між найближчими сусідами, яка утворюється дефектом кристалічної структури – вакансією (рис. 13.1).



Рис. 13.1. Дефект (вакансія) в одновимірному ланцюжку атомів як утворення поверхні

Дійсно, наявність вакансії з одного боку можна розглядати як дефект (перехід від рис. 13.1 а до 13.1 б), що спричиняє виникнення локального потенціалу, який визначатиме додаткове розсіювання електронів у дефектному кристалі. А наявність дефекту (як показано у розд. 8) спричиняє виникнення додаткового енергетичного рівня у забороненій зоні. З іншого боку, наявність вакансії можна розглядати як виникнення межі розподілу **кристал-зовнішнє середовище** (на рис. 13.1 б такі межі розподілу подано штриховою лінією). Такими чином, поверхню можна розглядати як один дефект (у тривимірному випадку – це двовимірний дефект), наявність якого спричиняє виникнення енергетичного рівня (рівнів) у забороненій зоні.

Шоклі розглянув модель одновимірного ланцюжка із восьми атомів, вважаючи, що на краю ланцюжка хід потенціалу не змінюється. Він обчислив енергетичні рівні атомного ланцюжка за різних значень міжатомних відстаней. З'ясувалося, що за міжатомних відстаней, менших від критичної, в енергетичній структурі ланцюжка відбувається розщеплення рівнів у смуги, розділені забороненими зонами. Усередині кожної забороненої зони виникають два стани, спричинені порушенням періодичності на двох поверхнях (ланцюжок характеризується двома поверхневими атомами). При цьому у кожній із сусідніх смуг енергії зникає по одному об'ємному стану (рис. 13.2). Такі поверхневі стани називаються *шокліївськими*. Їх можна пов'язувати із виникненням обірваних хімічних зв'язків при утворенні поверхні кристалу. Іншими словами, поверхневі рівні виникають завдяки наявності неспарених електронів, що існують поблизу поверхні через обрив ґратки (утворення поверхні). Таким чином, можна стверджувати, що й таммовські стани, і поверхневі стани Шоклі – два випадки одного явища – перебудови електронної структури кристалу за наявності поверхні. При цьому таммовські стани відповідають моделі квазівільного електрона, а шокліївські – моделі квазізв'язаного електрона.

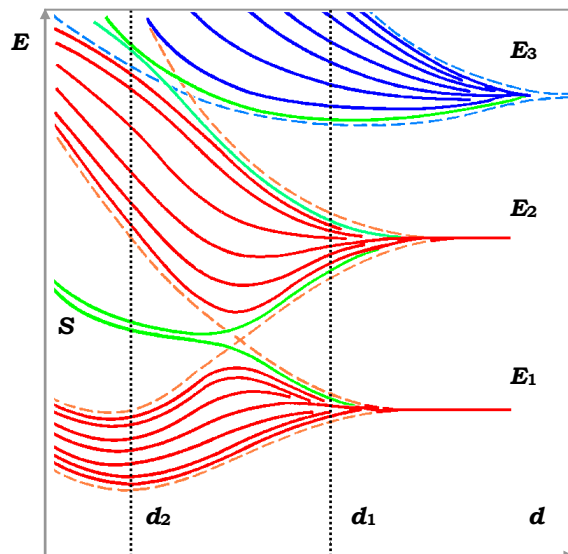


Рис. 13.2. Формування поверхневих станів за Шоклі.
 Червоним позначено валентні зони, синім – зону провідності.
 Зелений колір позначає поверхневі стани

Зрозуміло, що у реальних тривимірних кристалах замість дискретних поверхневих станів виникають поверхневі енергетичні смуги (зони), що

породжуються можливістю руху електронів паралельно поверхні, який характеризується хвильовими векторами \mathbf{k}_{\parallel} . Якщо така зона не повністю заповнена, то виникає можливість поверхневої металічної провідності. Через неідеальні умови поверхневі зони можуть перекриватись. Вони також можуть перекриватись із об'ємними зонами. Більш коректно вважати, що із поверхнею пов'язаний деякий неперервний розподіл поверхневих енергетичних станів. Такі стани найбільш коректно характеризувати їх щільністю. Якщо за деяких значень енергії щільність станів має яскраво виражені максимуми, то іноді говорять про квазідискретні поверхневі рівні. Але подібна ситуація можлива тільки для майже ідеальних – так званих атомарно чистих поверхонь, які можна отримати в умовах вакууму. Технології отримання таких поверхонь є досить складними у реалізації. У реальних умовах поверхня напівпровідника є геть не ідеальною: її зазвичай укривають окисні шари, на неї адсорбуються атоми та радикали, що також створюють додаткові поверхневі стани. При цьому створення поверхневих станів у процесі адсорбції може відбуватись двома різними механізмами. По-перше, адсорбовані частинки можуть утворювати поверхневі хімічні з'єднання – відбувається так звана хімічна адсорбція, тобто адсорбовані атоми чи молекули можуть обмінюватись із поверхнею електронами та, залежно від їхньої природи, віддавати електрон напівпровідникові або, навпаки, приєднувати до себе електрон напівпровідника. Таким чином, у результаті хімічної адсорбції утворюються заряджені поверхневі стани. Існує й інший тип зв'язку атомів або молекул із поверхнею напівпровідника: зв'язок відбувається завдяки квантовим флуктуаціям поля та неоднорідності щільності заряду – це так звані дисперсійні взаємодії або сили Ван-дер-Ваальса. У результаті таких взаємодій частинки із поверхнею обмін електронами між частинкою та поверхнею не відбувається. Така адсорбція називається фізичною. Оскільки приповерхнева область, де відбулась адсорбція, характеризується значно неоднорідним розподілом локального поля, можна вважати цю область поверхневим дефектом, з яким також може бути пов'язаний локалізований електронний стан. Таким чином, в області локалізації фізично адсорбованої частинки завдяки наявності локального поверхневого стану може зосереджуватись електрон з об'єму напівпровідника, викликаючи додаткову неоднорідність у щільності заряду. Отже, можна стверджувати, що у приповерхневій області напівпровідника може існувати вельми складна картина енергетичних станів. На неперервний енергетичний розподіл електронних станів (рис. 13.3 а) можуть нашаровуватись дискретні електронні стани (рис. 13.3 б) і таким чином утворювати складну картину сплесків густини електронних станів на фоні неперервного спектру поверхневих станів (рис. 13.3 в).

Зрозуміло, що зміну електричного потенціалу у приповерхневій області кристалу спричинятиме наявність поверхневих станів, але вона не є єдиною причиною зміни потенціалу. Саме існування межі розподілу є дуже вагомою причиною виникнення додаткового поверхневого заряду, що приводить до принципово відмінних ефектів від ситуації в об'ємі. За локалізації заряду у поверхневих електронних станах (незалежно від їхньої природи) поверхня кристала заряджається. Вільні носії в об'ємі перерозподіляються так, що нейтралізують поверхневий заряд. Зрозуміло, що з цим буде пов'язане виникнення поверхневого потенціалу, який є джерелом електричного поля.

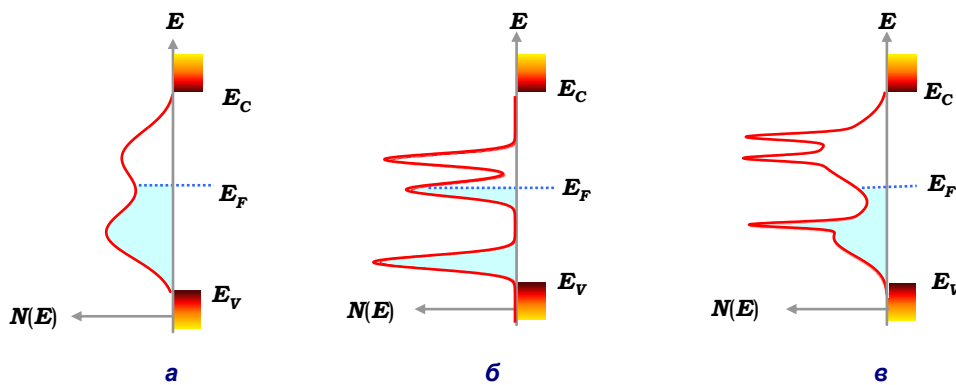


Рис. 13.3. Зонна енергетична схема поверхні:
 а – неперервний розподіл електронних станів; б – дискретні електронні стани;
 в – дискретні поверхневі електронні стани
 на фоні неперервного розподілу поверхневих станів

■ 13.2. Вплив поверхневого потенціалу на енергетичні зони напівпровідника поблизу поверхні

Поверхня є межею розподілу між напівпровідником і зовнішнім середовищем. Контакт є такою самою межею розподілу між двома речовинами. Таким чином, поверхневі та контактні явища мають багато спільних рис, які перш за все зумовлені наявністю додаткового потенціалу. Отже, можна стверджувати, що основою поверхневих і контактних явищ є процеси, що протікають у напівпровіднику під дією електричного поля, який виникає на межі розподілу. Для уявлення про результати дії таких процесів розглянемо поведінку напівпровідника n -типу, що внесений до зовнішнього однорідного електричного поля. Під дією поля у напівпровіднику відбу-

деться перерозподіл носіїв заряду так, що у його приповерхневій області з'являться об'ємний заряд зі щільністю $\rho(\mathbf{r})$ та екрануюче електричне поле $\mathbf{E}(\mathbf{r})$. Це поле заекранує зовнішнє електричне поле, завдяки чому воно зможе проникнути тільки у приповерхневий шар напівпровідника. Від'ємний об'ємний заряд створить електричне поле, напруженість якого буде максимальною на поверхні напівпровідника. Поле спричинить зміну потенціальної енергії електрона на величину $U(\mathbf{r}) = -e\varphi(\mathbf{r})$. Таким чином, зовнішнє електричне поле викличе таке викривлення зон напівпровідника, що

$$E_C(\mathbf{r}) = E_C - U(\mathbf{r}), \quad (13.19)$$

$$E_V(\mathbf{r}) = E_V - U(\mathbf{r}). \quad (13.20)$$

Разом із цим зміняться положення всіх рівнів енергії, у тому числі й рівнів домішок, що розташовані у забороненій зоні. З іншого боку, оскільки напівпровідник перебуває у стані термодинамічної рівноваги, то положення рівня Фермі буде сталим, тобто відстань між рівнем Фермі та краями валентної і зони провідності залежатиме від відстані до межі розподілу. Тепер відстань між зоною провідності та рівнем Фермі визначатиметься співвідношенням

$$\Delta E_C = [E_C - U(\mathbf{r})] - E_F. \quad (13.21)$$

Відстань між валентною зоною та рівнем Фермі

$$\Delta E_V = E_F - [E_V - U(\mathbf{r})]. \quad (13.22)$$

Розглянемо вплив зовнішнього поля на однорідний невідроджений напівпровідник n -типу, що характеризується діелектричною константою ϵ_r . Напруженість електричного поля пов'язана з об'ємним зарядом рівнянням Пуассона

$$\frac{dE}{dx} = \frac{4\pi}{\epsilon_r} \rho(x), \quad (13.23)$$

або

$$\frac{d^2\varphi}{dx^2} = -\frac{4\pi}{\epsilon_r} \rho(x). \quad (13.24)$$

Концентрація в об'ємі напівпровідника становить n_0 й, оскільки напівпровідник невідроджений, то

$$n_0 = N_C e^{-\frac{E_C - E_F}{kT}}, \quad x \rightarrow \infty. \quad (13.25)$$

У приповерхневій області для концентрації носіїв маємо

$$n = N_C e^{-\frac{E_C + U - E_F}{kT}} = n_0 e^{-\frac{U}{kT}}. \quad (13.26)$$

Поблизу поверхні об'ємний заряд визначається позитивними іонами донорної домішки та вільними електронами. Якщо донорна домішка повністю іонізована, то в об'ємі

$$n_0 = N_D^+ = N_D. \quad (13.27)$$

Тоді у приповерхневій області поверхневий заряд

$$\rho = e(N_D - n). \quad (13.28)$$

Із цих рівнянь маємо

$$\rho = e(n_0 - n) = en_0(1 - e^{-U/kT}). \quad (13.29)$$

Якщо вважати, що зони викривлені не значно, тобто $|U| \ll kT$, то можна розкласти експоненту у ряд. Із точністю до першої поправки за U/kT запишемо

$$\rho = \frac{en_0U}{kT} = -\frac{e^2n_0}{kT}\varphi. \quad (13.30)$$

Підставимо цей вираз до правої частини рівняння (13.24)

$$\frac{d^2\varphi}{dx^2} = \frac{4\pi e^2n_0}{\epsilon_r kT}\varphi. \quad (13.31)$$

Позначивши $\frac{4\pi e^2n_0}{\epsilon_r kT} = L_S^{-2}$, маємо із (13.31)

$$\frac{d^2\varphi}{dx^2} - \frac{1}{L_S^2}\varphi = 0. \quad (13.32)$$

Розв'язок цього рівняння шукаємо у вигляді¹

$$\varphi = A_1 e^{-x/L_S} + A_2 e^{x/L_S}. \quad (13.33)$$

Оскільки при $x \rightarrow \infty$ поле відсутнє, коефіцієнт $A_2 = 0$. З іншого боку, при $x = 0$ $\varphi = -\varphi_s$. Тоді

$$\varphi(x) = -\varphi_s e^{-x/L_S}. \quad (13.34)$$

Звідси маємо вирази для електричного поля

$$E(x) = -\frac{d\varphi}{dx} = -\frac{\varphi_s}{L_S} e^{-x/L_S} = -E_s e^{-x/L_S} \quad (13.35)$$

та потенціальної енергії електронів

$$U(x) = -e\varphi(x) = e\varphi_s e^{-x/L_S} = U_s e^{-x/L_S}. \quad (13.36)$$

¹ Згадаємо, що L_S є ні чим іншим, як довжиною екранування Дебая (див. п. 11.4)

Щільність зарядів на поверхні зразка

$$\rho_s = \frac{en_0}{kT} U_s. \quad (13.37)$$

Таким чином, за дії електричного поля на напівпровідник в його приповерхневій області відбувається викривлення зон і зміна концентрації носіїв. При $U_s > 0$ зони викривляються вгору й збільшується концентрація дірок (рис.13.4 а,б). У цьому випадку у напівпровідника n -типу приповерхневий шар збагачується неосновними носіями, а у напівпровідника p -типу – основними носіями. При $U_s < 0$ зони викривляються вниз і збільшується

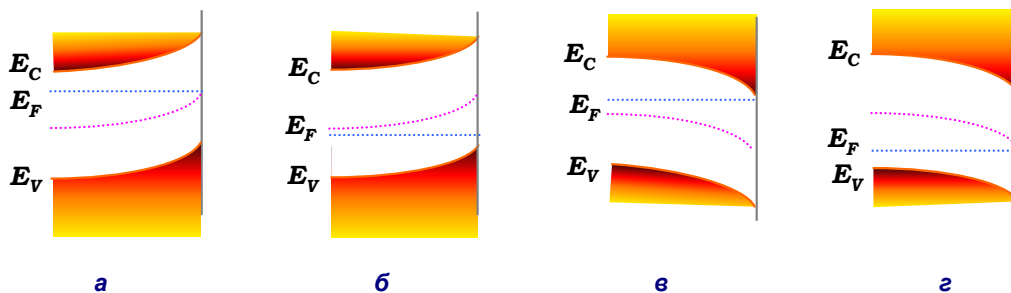


Рис. 13.4 Вигин зон залежно від знаку поверхневого потенціалу для напівпровідників: а, в – n -типу; б, г – p -типу.

Діаграми (а, б) відповідають позитивному значенню потенціалу ($U_s > 0$), (в, г) – негативному значенню потенціалу ($U_s < 0$). Іншими словами, у напівпровідника n -типу приповерхневий шар збагачується основними носіями, а у напівпровідника p -типу – неосновними. Дебаєвська довжина екранування L_S для метала ($n_0 = 10^{22} \text{ см}^{-3}$, $\epsilon_r = 1$) становить величину 10^{-8} см . У напівпровідниках ця величина є набагато більшою, наприклад для германію ($n_0 = 10^{14} \text{ см}^{-3}$, $\epsilon_r = 16$) $L_S = 4 \cdot 10^{-4} \text{ см}$.

■ 13.3. Приповерхнева область просторового заряду

У попередньому розділі ми дійшли висновку, що біля поверхні напівпровідника існує досить товстий шар, де концентрація рівноважних носіїв визначається не об'ємним легуванням, а величиною та знаком поверхневого заряду. Наявність останнього спричиняє виникнення вигину енергетичних зон, що наочно подано за допомогою енергетичної схеми напівпровідника, на якій за віссю ординат відкладено енергію, а за віссю аб-

счис – відстань від поверхні кристала, якій відповідає значення $x = 0$. Вигин енергетичних зон біля поверхні ілюструє зміну потенціальної енергії електронів на поверхні за рахунок заряду, що локалізований на поверхневих рівнях. Нахил краю зони у приповерхневій області характеризує поле в області просторового заряду (ОПЗ).

Розглянемо модель приповерхневої області напівпровідника, наприклад n -типу. Нехай ψ – електростатичний потенціал, що відповідає середині забороненої зони. Оскільки він є функцією координати x , вважатимемо, що

$$\psi(x = 0) = \psi_s, \quad \psi(x \rightarrow \infty) = \psi_b. \quad (13.38)$$

Таким чином, в області, далекій від поверхні, значення потенціалу збігається із таким самим в об'ємному зразку ψ_b , а на поверхні потенціал набуває визначеного значення ψ_s . Аналогічно всі значення параметрів на поверхні зразку позначатимемо індексом s (від *surface* – поверхня), а значення параметрів, що визначаються в області, далекій від поверхні, – індексом b (від *bulk* – об'єм). За нульове значення потенціалу візьмемо дно зони провідності, енергію ж відкладатимемо вгору від дна зони. Для визначеності розглядатимемо невідроджений напівпровідник, де вигин зон не дуже великий, що означає: на поверхні напівпровідник залишається невідродженим. Для з'ясування залежності електростатичного потенціалу від координати розглянемо одновимірне рівняння Пуасона

$$\frac{d^2\psi}{dx^2} = -\frac{4\pi}{\varepsilon}\rho(x), \quad (13.39)$$

де ε – діелектрична константа напівпровідника, $\rho(x)$ – густина об'ємного заряду на відстані x від поверхні. Це рівняння необхідно доповнити граничними умовами (13.38) (вважатимемо зразок напівпровідника нескінченним уздовж вісі Ox) та

$$\left. \frac{d\psi}{dx} \right|_{x \rightarrow \infty} = 0. \quad (13.40)$$

Вважаючи всі домішки іонізованими, для густини об'ємного заряду запишемо

$$\rho(x) = e[p(x) - n(x) + N_D - N_A]. \quad (13.41)$$

Тут $n(x)$ і $p(x)$ – концентрації електронів і дірок у точці x ; N_D і N_A – концентрації донорів та акцепторів, відповідно. Вважатимемо, що напівпровідник перебуває у нерівноважних умовах, коли вільні носії у приповерхневій області задаються не тільки термічним, але й іншими способами генерації. В об'ємі кристалу виконується умова електронейтральності, тобто

$$p_b - n_b + N_D - N_A = 0, \quad (13.42)$$

де p_b і n_b – рівноважні концентрації дірок та електронів в об'ємі, відповідно. Підставляючи із цього рівняння різницю $N_D - N_A$ до (13.41), дістанемо

$$\rho(x) = e[p(x) - p_b - n(x) + n_b]. \quad (13.43)$$

Вважаючи рівень хімічного потенціалу (рівень Фермі) власного напівпровідника таким, що розташований усередині забороненої зони², запишемо для концентрації рівноважних носіїв заряду в об'ємі кристалу

$$n_b = n_i e^{\beta(\psi_b - \Phi_b)}, \quad (13.44)$$

$$p_b = n_i e^{-\beta(\psi_b - \Phi_b)}, \quad (13.45)$$

де $n_i = N_C e^{-\beta\psi_b}$ – концентрація носіїв заряду в об'ємі власного напівпровідника; $\beta = e/kT$. Величина $e\Phi_b = E_F$ є не чим іншим, як рівнем Фермі в об'ємі власного напівпровідника. Аналогічно для нерівноважної концентрації носіїв

$$n(x) = n_i e^{\beta(\psi_b - \Phi_n)}, \quad p(x) = n_i e^{-\beta(\psi_b - \Phi_p)}, \quad (13.46)$$

де $e\Phi_n$ та $e\Phi_p$ – квазірівні Фермі електронів і дірок, відповідно.

Підставимо густину зарядів (13.43) із врахуванням (13.44)–(13.46) до (13.40) і дістанемо рівняння Пуассона

$$\frac{d^2\psi(x)}{dx^2} = -\frac{4\pi en_i}{\epsilon_r} \left\{ e^{\beta(\Phi_p - \psi(x))} - e^{\beta(\Phi_p - \psi_b)} - e^{\beta(\psi(x) - \Phi_n)} + e^{\beta(\psi_b - \Phi_n)} \right\}. \quad (13.47)$$

У цьому рівнянні, крім змінної $\psi(x)$, квазірівні Фермі $e\Phi_n$ і $e\Phi_p$ також залежать від координати. Квазірівні Фермі змінюються зі зміною x від значень $e\Phi_n^s = e\Phi_n(x=0)$ та $e\Phi_p^s = e\Phi_p(x=0)$ на поверхні до $e\Phi_b$ в об'ємі. Причому, оскільки збурення напівпровідника від рівноважного стану відбувається на поверхні, то набуття квазірівнями Фермі рівноважного значення – на відстанях порядку дифузійної довжини нерівноважних носіїв. За незначних відхилень від термодинамічної рівноваги та невеликих значень товщини ОПЗ можна вважати ці величини сталими. Такі припущення справджуються для германію та кремнію, де дифузійні довжини

² Як обговорювалось у розд. 8, розташування рівня Фермі у власному напівпровіднику визначається відношенням ефективних мас електронів і дірок. Рівень Фермі зміщений від середини забороненої зони на величину $3/4kT \ln(m_p^*/m_e^*)$ що, наприклад для кремнію, за кімнатної температури становить 0.02 еВ, у той час як шириною забороненої зони є 1,1 еВ.

на кілька порядків більші за довжину екранування. Зважаючи на ці обставини, можна обчислити перший інтеграл рівняння (13.47): домножимо обидві його частини на $2(d\psi/d)x$ і проінтегруємо його від x до ∞ або від ψ до ψ_b . При цьому необхідно врахувати граничні умови

$$\left. \frac{d\psi}{dx} \right|_{x \rightarrow \infty} = 0.$$

У результаті дістанемо рівняння, що визначає потенціал поля у довільній точці ОПЗ

$$\left\{ \frac{d(\psi - \psi_b)}{dx} \right\}^2 = \frac{8\pi en_i}{\varepsilon} \left\{ \frac{1}{kT} \left[e^{\beta(\Phi_p - \psi)} - e^{\beta(\Phi_p - \psi_b)} \right] + (\psi - \psi_b) e^{\beta(\Phi_p - \psi_b)} + \right. \\ \left. + \frac{1}{kT} \left[e^{\beta(\Phi_p - \psi_b)} - e^{\beta(\Phi_p - \psi)} \right] + (\psi - \psi_b) e^{\beta(\Phi_p - \psi_b)} \right\}. \quad (13.48)$$

Для аналізу цього рівняння введемо стандартні позначення

$$Y = \beta(\psi - \psi_b), \quad (13.49)$$

$$\lambda = \frac{P_b}{n_i} = \sqrt{\frac{P_b}{n_b}} = e^{\beta(\Phi_p - \psi)}, \quad (13.50)$$

$$P = e^{\frac{\Phi_p - \Phi_b}{kT}}, \quad (13.51)$$

$$N = e^{\frac{\Phi_n - \Phi_b}{kT}}. \quad (13.52)$$

Параметр λ характеризує об'ємне легування так, що для напівпровідників n -типу $\lambda < 1$. Для напівпровідника p -типу $\lambda > 1$. У власному напівпровіднику $\lambda = 1$. Взавши до уваги, що

$$\left(\frac{\varepsilon_r kT}{8\pi e^2 n_i} \right)^{1/2} = l_D$$

– довжина екранування ($l_D = L_s/\sqrt{2}$, де L_s – довжина екранування Дебая), рівняння (13.48) у позначеннях (13.49)–(13.52) набуває вигляду:

$$\frac{dY}{dx} = l_D^{-1} F(Y, \lambda, P, N), \quad (13.53)$$

де

$$F(Y, \lambda, P, N) = \pm \left\{ \lambda \cdot e^P (e^{-Y} - 1) + \lambda^{-1} e^{-N} (e^Y - 1) + (\lambda - \lambda^{-1}) \cdot Y \right\}^{1/2}. \quad (13.54)$$

Перший і другий доданки у фігурних дужках визначаються внеском до об'ємного заряду вільних дірок та електронів. Останній доданок визначається об'ємним зарядом некомпенсованих домішок.

Оскільки $\psi - \psi_b$ визначає зміну потенціалу щодо його значення в об'ємі, то додатні значення Y відповідають вигину зон униз, а від'ємні – вигину зон уверх. Таким чином, при $Y > 0$ у приповерхневій області об'ємний заряд негативний, а при $Y < 0$ – позитивний, тому у першому випадку необхідно у (13.54) брати знак мінус, а у випадку $Y < 0$ – плюс.

Залежність електростатичного потенціалу ψ від координати в ОПЗ визначається другим інтегралом рівняння Пуассона

$$x = l_D \int_{Y_s}^Y dY [F(Y, \lambda, P, N)]^{-1}, \quad (13.55)$$

що являє собою форму потенціального бар'єра для вільних носіїв, який створюється у приповерхневій області внаслідок локалізації зарядів на поверхневих станах. Обчислення інтегралу в (13.55) є складною задачею і, взагалі кажучи, вимагає проведення чисельного інтегрування. Але в окремих випадках, що можуть бути легко реалізовані в експерименті, рівняння (13.55) можна розв'язати аналітично. Розглянемо окремі випадки.

Область збагачення. У цьому випадку у приповерхневому шарі концентрація основних носіїв більша за їхню концентрацію в об'ємі зразка. У матеріалі n -типу такий випадок реалізується за вигину зон униз, тобто при $Y_s > 0$, а в матеріалі p -типу – за вигину зон угору ($Y_s < 0$). Визначальний внесок до формування об'ємного заряду біля поверхні у випадку збагачення дають основні носії заряду. Це означає, що функцію $F(Y, \lambda)$ в основному визначає перший доданок для кристалу p -типу, або другий доданок – у кристалі n -типу. Розглянемо для визначеності напівпровідник n -типу. Зважаючи на те, що у цьому випадку $F(Y, \lambda) < 0$, маємо

$$F(Y, \lambda, P, N) \approx -\lambda^{-1/2} e^{Y/2}. \quad (13.56)$$

Підставимо цей вираз під інтеграл рівняння (13.55). Дістанемо

$$x = -l_D \lambda^{-1/2} \int_{Y_s}^Y e^{Y/2} dY = -l_{eff} (e^{Y/2} - e^{Y_s/2}), \quad (13.57)$$

де $l_{eff} = 2\lambda^{-1/2} l_D$ – довжина екранування, що визначає товщину ОПЗ. Дійсно, знайдемо із (13.57) залежність електростатичного потенціалу від координати x

$$Y = 2 \ln(e^{Y_s/2} - x/l_{eff}). \quad (13.58)$$

Із формули, зокрема видно, що при $x \rightarrow 0$ $Y \rightarrow Y_s$. При $x = l_{eff}(e^{Y_s/2} - 1)$ $Y = 0$, тобто на відстанях $x \approx l_{eff}$ вигин зон зникає. Таким чином, електростатичний потенціал швидко спадає зі збільшенням координати x , а товщина ОПЗ визначається довжиною екранування l_{eff} . За високого рівня легування ця довжина може бути істотно меншою за довжину Дебая L_s , тобто при збагаченні об'ємний заряд в основному зосереджується у

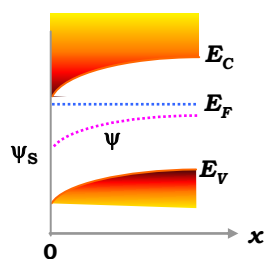


Рис. 13.5. Область збагачення поблизу поверхні напівпровідника *n*-типу

приповерхневому шарі. Для визначення товщини ОПЗ можна використати (13.57), звідки знайти величину x , що відповідає значенню $Y = 0$ (відсутність викривлення зон). У результаті дістаємо

$$L_{ОПЗ} = l_{eff} e^{Y_s/2}. \quad (13.59)$$

Таким чином, ми отримали, що довжина ОПЗ визначається ефективною довжиною екранування l_{eff} і значно залежить від вигину зон на поверхні Y_s . Хід потенціалу ψ у напівпровіднику *n*-типу у випадку збагачення подано на рис. 13.5.

Область збіднення. Якщо у приповерхневому шарі концентрація основних і неосновних носіїв заряду є значно меншою від концентрації іонізованих домішок, то це означає, що у приповерхневій області виникає область збіднення. Тут концентрація основних носіїв заряду може бути значно меншою за її значення в об'ємі. Така область може існувати у напівпровіднику *n*-типу, якщо викривлення зон відбувається вгору, тобто $Y_s < 0$. У напівпровіднику *p*-типу область збіднення може реалізуватись, якщо зони вигнуті згори ($Y_s > 0$).

При збідненні основний внесок до (13.54) надає останній доданок, тобто в цьому випадку для напівпровідника *n*-типу

$$F(Y, \lambda, P, N) = \{(\lambda - \lambda^{-1}) \cdot Y\}^{1/2}. \quad (13.60)$$

Підставляючи цю функцію під інтеграл до (13.55), дістанемо

$$x = l_D \int_{Y_s}^Y \{(\lambda^{-1} - \lambda)(-Y)\}^{-1/2} dY = -l_{eff} \int_{|Y_s|}^{|Y|} t^{-1/2} dt = l_{eff} (|Y_s|^{1/2} - (Y)^{1/2}), \quad (13.61)$$

де

$$l_{eff} = 2l_D(\lambda^{-1} - \lambda)^{-1/2}. \quad (13.62)$$

Звідси легко визначити, що потенціал **змінюється із** координатою за параболічним законом (Y – модуль потенціалу)

$$Y = \left\{ |Y_s|^{1/2} - \frac{x}{l_{eff}} \right\}^2. \quad (13.63)$$

Як видно із формули, при $x=0$ $Y = |Y_s|$, а при $x = l_{eff} \sqrt{|Y_s|}$ $Y = 0$. Оскільки ефективна довжина екранування

$$l_{eff} = \left(\frac{\epsilon_r kT}{2\pi e^2} \right)^{1/2} \frac{1}{(n_0 - p_0)^{1/2}}, \quad (13.64)$$

то для високоомних напівпровідників, де концентрація носіїв невелика, область збіднення затягується далеко до об'єму кристала. Довжина області ОПЗ у цьому випадку визначається із (13.60) умовою $Y(x = L_{\Gamma} \bar{i} \zeta) = 0$, звідки

$$L_{ОПЗ} = l_{eff} |Y_s|^{1/2} = \left(\frac{\epsilon_r kT}{2\pi e^2} \right)^{1/2} \left(\frac{|Y_s|}{n_0 - p_0} \right)^{1/2}. \quad (13.65)$$

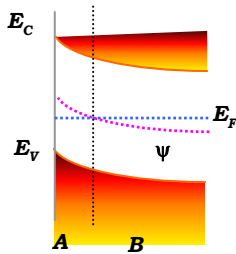


Рис. 13.6. Области збіднення (B) та інверсії (A) у напівпровіднику n-типу

За значного збіднення у приповерхневій області край зони неосновних носіїв розміщується ближче до рівня Фермі, ніж зона основних носіїв заряду (рис. 13.6). У таких умовах у приповерхневій області формується шар, в якому тип провідності відрізняється від об'ємного, – *інверсійний шар*. Умовою реалізації цього явища є $Y_s > 0$ у напівпровіднику p-типу та $Y_s < 0$ – у напівпровіднику n-типу. Із визначення інверсного шару випливає, що критерієм його утворення є перетин рівнем Фермі середини забороненої зони. Це, зокрема означає, що для існування області інверсії має виконуватись нерівність $Y_s > \beta(\Phi_b - \psi_b)$ або $Y_s > \ln \lambda$.

В області інверсії головний внесок до об'ємного заряду дають неосновні носії заряду. Тоді для напівпровідника n-типу у виразі (13.54) перший доданок буде набагато більшим за решту, тобто

$$F(Y, \lambda, P, N) = \lambda^{-1/2} \cdot e^{Y/2}. \quad (13.66)$$

Підставляючи цей вираз до (13.55), дістанемо

$$x = l_D \int_{Y_s}^Y \lambda^{-1/2} e^{Y/2} = l_{eff} \left[e^{Y/2} - e^{Y_s/2} \right]. \quad (13.67)$$

У цьому випадку

$$l_{eff} = \left(\frac{\varepsilon k T}{2\pi e^2 p_0} \right)^{1/2}, \quad (13.68)$$

і зміна потенціалу із координатою відбувається за логарифмічним законом

$$Y = 2 \ln \left[e^{Y_s/2} + x/l_{eff} \right]. \quad (13.69)$$

Іншими словами, потенціал досить повільно змінюється із координатою й за малих значень x майже не відрізняється від Y_s .

У випадку незначних вигинів зон, коли $|Y_s| \ll 1$, для оцінки форми потенціалу розкладемо експоненти у (13.54) у ряд Тейлора, обмежуючись доданками другого порядку. У результаті для напівпровідника n -типу і $Y_s > 0$ дістанемо

$$x = -\sqrt{2} l_D \int_{Y_s}^Y Y^{-1} (\lambda + \lambda^{-1})^{-1/2} dY = -l_{eff} \ln(Y/Y_s), \quad (13.70)$$

де

$$l_{eff} = \left(\frac{\varepsilon k T}{4\pi e^2} \right)^{1/2} \frac{1}{(n_0 - p_0)^{1/2}}. \quad (13.71)$$

Звідси отримуємо залежність потенціалу від координати

$$Y = Y_s e^{-x/l_{eff}}. \quad (13.72)$$

■ 13.4. Ефект поля

Зміна поверхневого потенціалу спричинятиме зміни концентрації носіїв заряду у приповерхневому шарі напівпровідника, отже – зміни його електропровідності. Зрозуміло, що змінювати поверхневий потенціал можна за допомогою створення поблизу поверхні напівпровідника електричного поля. Це зовнішнє електричне поле, таким чином, спричинятиме зміни електропровідності зразка. Явище зміни провідності напівпровідника під дією зовнішнього електричного поля, нормального до поверхні зразка,

називається *ефектом поля*. Перед детальним обговоренням ефектів поля розглянемо вплив поверхневого потенціалу на електропровідність.

13.4.1. Вплив поверхневого потенціалу на електропровідність

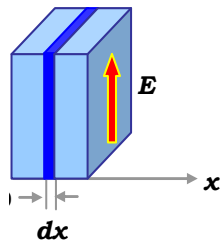


Рис. 13.7 До визначення латеральної провідності

Припустимо, що необхідно визначити латеральну провідність тонкої пластинки напівпровідника як відгук на електричне поле, паралельне поверхні. Нехай $n(x)$ і $p(x)$ – концентрації електронів і дірок у приповерхневому шарі пластинки, відповідно (рис. 13.7), що відрізняються від їхніх значень в об'ємі (або як кажуть у випадку невикривлених зон) n_0 і p_0 . Такі позначення для концентрації носіїв в об'ємі (замість n_b і p_b , які використовувалися раніше) ми застосовуємо задля підкреслення, що відмінність концентрації носіїв від n_0 і p_0 має сенс надлишкової концентрації носіїв, які можуть вважатись нерівноважними.

Вважатимемо, що рухливості електронів і дірок у приповерхневій області майже не відрізняються від їхніх значень в об'ємі – μ_n та μ_p . Струм через нескінченно тонкий шар пластинки тоді запишемо як

$$dj_d = e[\mu_n(n(x) - n_0) + \mu_p(p(x) - p_0)]Edx. \quad (13.73)$$

Тоді відмінністю струму від випадку невикривлених зон є

$$dj_d = e[\mu_n(n(x) - n_0) + \mu_p(p(x) - p_0)]Edx. \quad (13.74)$$

Інтегруючи це рівняння, знайдемо зміну електропровідності пластинки (віднесено на одиницю площі), спричинену впливом поверхні

$$G_S = e(\mu_n\Gamma_n + \mu_p\Gamma_p), \quad (13.75)$$

де надлишок поверхневої концентрації електронів і дірок

$$\Gamma_n = \int_0^{\infty} (n(x) - n_0) dx, \quad (13.76)$$

$$\Gamma_p = \int_0^{\infty} (p(x) - p_0) dx. \quad (13.77)$$

Таким чином, визначення впливу поверхні на електропровідність зводиться до обчислення розподілів концентрації електронів $n(x)$ і дірок $p(x)$ у шарі об'ємного заряду. У свою чергу цей розподіл повністю визначається концентраціями n_b , p_0 і ходом електричного потенціалу $\varphi(x)$. Такі обчислення є досить складними й зазвичай проводяться чисельно; тут ми лише коротко подамо їх загальну схему.

Вихідним пунктом схеми є рівняння Пуассона

$$\frac{d^2\varphi}{dx^2} = -\frac{4\pi\rho(x)}{\varepsilon}, \quad (13.78)$$

де ε – відносна діелектрична сприйнятливість напівпровідника, $\rho(x)$ – густина заряду. При проведенні розрахунків припускається, що домішки повністю іонізовані та рівномірно розподілені по об'єму напівпровідника. Це дає змогу записати

$$\rho(x) = -e(n - p + N_A - N_D), \quad (13.79)$$

або із врахуванням вимоги електронейтральності об'єму напівпровідника

$$n_0 - p_0 = N_D - N_A, \quad (13.80)$$

$$\rho(x) = -e[(n(x) - n_0) - (p(x) - p_0)]. \quad (13.81)$$

При розв'язку задачі зручно відраховувати енергію від рівня Фермі F_i власного напівпровідника³, де ефективні маси щільності станів $m_{dn}^* = m_{dp}^*$. Тоді

$$n(x) = n_0 e^{e\varphi/kT}, \quad p(x) = p_0 e^{-e\varphi/kT}. \quad (13.82)$$

Із врахуванням (13.82) густину зарядів (13.80) можна записати у вигляді

$$\rho(x) = -e[n_0(e^{e\varphi/kT} - 1) + p_0(e^{-e\varphi/kT} - 1)]. \quad (13.83)$$

У подальшому вводяться стандартні позначення: безрозмірний електростатичний потенціал

$$Y = \frac{e\varphi}{kT}, \quad Y(x=0) = Y_S = \frac{e\varphi_S}{kT}; \quad (13.84)$$

характеристика ступеня легування напівпровідника (є оберненою до параметра λ , визначеного (13.50), і також широко застосовується)

³ У зв'язку з таким вибором початку відліку енергії використовуватимемо інші позначення, ніж у попередньому пункті. Подальші обчислення подібні до проведених у п. 13.3.

$$\lambda = \frac{n_0}{n_i} = \frac{n_i}{p_0}. \quad (13.85)$$

Нерівність $\lambda > 1$ відповідає напівпровіднику n -типу, а $\lambda < 1$ – напівпровіднику p -типу. Введемо також параметр розмірності довжини, що менший у $\sqrt{2}$ рази від дебаєвського радіусу екранування власного напівпровідника

$$L_D = \frac{1}{\sqrt{2}} \sqrt{\frac{\epsilon k T}{4\pi e^2 n_i}}. \quad (13.86)$$

Підставляючи вираз (13.83) до (13.78) і використовуючи позначення (13.84–13.86), дістанемо рівняння

$$2 \frac{d^2 Y}{dx^2} = L_D^{-2} [\lambda(e^Y - 1) - \lambda^{-1}(e^{-Y} - 1)], \quad (13.87)$$

яке необхідно до визначити граничними умовами

$$Y(0) = Y_S, \quad Y(\infty) = 0, \quad \left. \frac{dY}{dx} \right|_{x=\infty} = 0. \quad (13.88)$$

Домножимо обидві частини (13.87) на dY/dx і візьмемо до уваги, що

$$2 \frac{d^2 Y}{dx^2} \frac{dY}{dx} = \frac{d}{dx} \left(\frac{dY}{dx} \right)^2. \quad \text{У результаті}$$

$$d \left(\frac{dY}{dx} \right)^2 = L_D^{-2} \{ \lambda(e^Y - 1) - \lambda^{-1}(e^{-Y} - 1) \} dY. \quad (13.89)$$

У подальшому необхідно врахувати, що за зміни координати від x до ∞ похідна dY/dx змінюється від величини dY/dx до 0, а Y – від Y до 0. Це означає, що при визначенні першого інтегралу (13.89) інтегрування необхідно проводити від dY/dx до 0 – у лівій, і від Y до 0 – у правій частині рівняння. У результаті дістаємо

$$(dY/dx)^2 = L_D^{-2} \{ \lambda(e^Y - 1) + \lambda^{-1}(e^{-Y} - 1) + Y(\lambda^{-1} - \lambda) \}.$$

Звідси, використовуючи позначення

$$F(\lambda, Y) = \{ \lambda(e^Y - 1) + \lambda^{-1}(e^{-Y} - 1) + Y(\lambda^{-1} - \lambda) \}^{1/2}, \quad (13.90)$$

знайдемо

$$(dY/dx) = \pm L_D^{-1} F(\lambda, Y). \quad (13.91)$$

Для від'ємних значень Y вибирається знак плюс, а для позитивних – мінус. Повертаючись тепер до (13.76–13.77), можна записати, що

$$\Gamma_n = n_0 \int_0^\infty (e^Y - 1) dx = n_0 \int_{Y_s}^0 (e^Y - 1) (dY/dx)^{-1} dY = n_0 L_D \int_0^{Y_s} (e^Y - 1) F^{-1}(\lambda, Y) dY, \quad (13.92)$$

$$\Gamma_p = p_0 \int_0^\infty (e^{-Y} - 1) dx = p_0 \int_{Y_s}^0 (e^{-Y} - 1) (dY/dx)^{-1} dY = p_0 L_D \int_0^{Y_s} (e^{-Y} - 1) F^{-1}(\lambda, Y) dY \quad (13.93)$$

Таким чином, розрахунок впливу поверхневого потенціалу на електропровідність зводиться до обчислення розподілу $n(x)$ і $p(x)$ у шарі об'ємного заряду згідно із (13.92-13.93). Але загальні риси поведінки G_S так само, як й у випадку ОПЗ, можна визначити із якісних міркувань. Розглянемо для

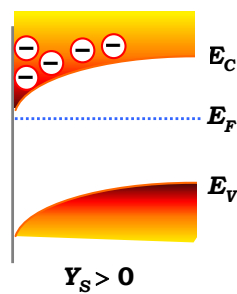


Рис. 13.8. Збагачений приповерхневий шар у напівпровіднику n-типу

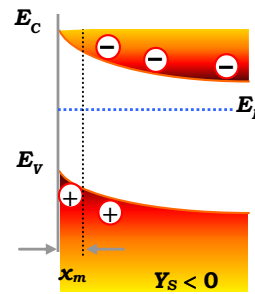


Рис. 13.9. Збіднений приповерхневий шар у напівпровіднику n-типу

визначеності напівпровідник n-типу. Якщо потенціал поверхні Y_s є позитивним, зони викривляються догори, і край зони провідності наближається до рівня Фермі (рис. 13.8). Завдяки цьому у приповерхневій області з'являється шар, збагачений основними носіями заряду, і як впливає із (13.75) та (13.92), поверхнева провідність G_S збільшуватиметься за зростання Y_s . У випадку, коли поверхневий потенціал є від'ємним, зони викривляються вгору. Приповерхневий шар збіднюватиметься електронами. У цьому випадку, доки величина $E_F - E_V$ буде більшою за $E_C - E_F$, концентрація дірок залишатиметься набагато меншою за концентрацію електронів, і G_S зменшуватиметься за зростання абсолютної величини Y_s (рис. 13.9). Але коли викривлення зон стає значним – таким, що рівень Фермі у приповерхневій області стає ближчим до валентної зони, ніж до зони провідності, різко зростає концентрація дірок. Таким чином, у приповерхневій області виникає інверсний шар. Основну роль тепер відіграють дірки, і подальше зростання $|Y_s|$, згідно із (13.75) та (13.93), спричинить зростання поверх-

невої провідності. При цьому G_S пройде через мінімум. Аналогічно поводитиметься й дірковий напівпровідник. Оскільки поблизу мінімуму у випадку $F - E_V > E_C - E_F$ електронний внесок до провідності зменшується, а дірковий – збільшується при $F - E_V < E_C - E_F$, у точці мінімуму G_S електрони та дірки дають однаковий внесок. Іншими словами, у деякому перерізі $x = x_m$, де, наприклад потенціал $Y_S = (1/2) \cdot (Y_S)_{\min}$, виконується співвідношення

$$\mu_n n(x_m) = \mu_p p(x_m). \quad (13.94)$$

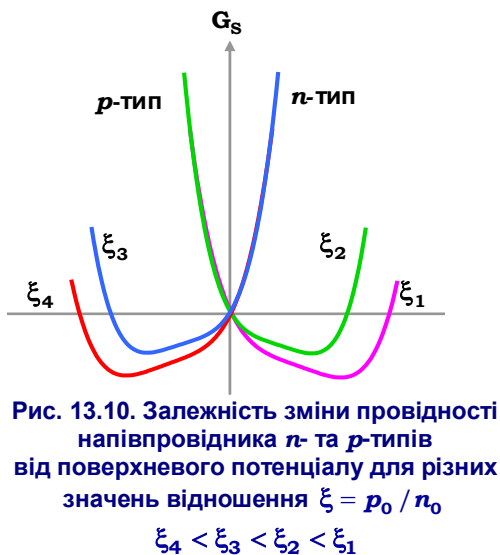
Вважаючи напівпровідник невідродженим, тобто $n = n_0 e^Y$, $p = p_0 e^{-Y}$, отримаємо із (13.94)

$$\mu_n n_0 \exp[(Y_S)_{\min} / 2] = \mu_p p_0 \exp[-(Y_S)_{\min} / 2], \quad (13.95)$$

звідси дістанемо оцінку

$$(Y_S)_{\min} = \ln \left(\frac{p_0 \mu_n}{n_0 \mu_p} \right). \quad (13.96)$$

Бачимо, що, оскільки зазвичай $(\mu_n / \mu_p) < 1$ для напівпровідника n -типу, де $(p_0 / n_0) < 1$, величина під логарифмом є меншою за одиницю, тобто значення поверхневого потенціалу, що відповідає мінімальній поверхневій провідності, буде негативним. При цьому його абсолютна величина буде тим більшою, чим меншим буде вираз $(p_0 \mu_n) / (n_0 \mu_p)$. Для напівпровідника



p -типу $(p_0 / n_0) > 1$, і значення поверхневого потенціалу, що відповідає мінімальній поверхневій провідності, буде позитивним. Зауважимо, що у власних напівпровідниках, як впливає із (13.96), мінімум поверхневої провідності спостерігатиметься за малих (за мірою відмінності відношення μ_n / μ_p від одиниці) негативних значень поверхневого потенціалу. На рис. 13.10 схематично подано хід залежності величини G_S від поверхневого потенціалу для різних значень відношення p_0 / n_0 . Цей графік наводить на думку про можливість експериментального визначення поверхневого потенціалу Y_S .

Дійсно, якщо відомо відношення концентрацій неосновних до основних носіїв p_0/n_0 за відомого співвідношення рухливостей μ_n/μ_p , то за допомогою (13.96) визначають величину $(Y_S)_{\min}$. Далі змінюють величину Y_S (напр., змінюючи газову атмосферу, підбираючи різноманітні гази та варіюючи їхній тиск). При цьому намагаються підібрати оточуюче середовище так, щоб електропровідність проходила через мінімум G_{\min} . Тоді

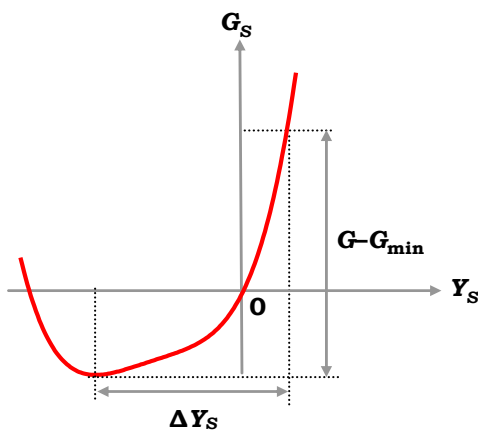


Рис. 13.11. Визначення поверхневого потенціалу Y_S

за будь-якого іншого стану поверхні, якому відповідає електропровідність G , поверхневий потенціал визначається за формулою

$$Y_S = (Y_S)_{\min} + \Delta Y_S, \quad (13.97)$$

де ΔY_S – різниця між $(Y_S)_{\min}$ і величиною Y , що відповідає поверхневій електропровідності $G_S = G - G_{\min}$ (рис. 13.11).

13.4.2. Вплив зовнішнього нормального поля на електропровідність

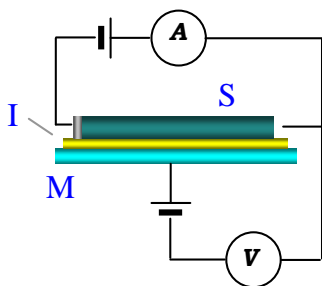


Рис. 13.12. Схема дослідження стаціонарного ефекту поля

Величину поверхневого потенціал можна, як зазначалось раніше, варіювати не тільки змінюючи навколишнє середовище, а прикладаючи зовнішнє поле до напівпровідникового зразка, тобто використовуючи ефект поля. Існує багато експериментальних методів вивчення ефекту поля. Розглянемо, наприклад метод вивчення стаціонарного ефекту поля (рис. 13.12). Схема базується на вивченні латеральної провідності напівпровідникової півки S , на торцях якої розташовані низькоомні контакти. Півка через тонкий шар діелектрика I утворює із тон-

кою металічною пластинкою M конденсатор. До конденсатора прикладають сталу напругу, величина і знак якої контролюються. За допомогою низькоомних контактів на торцях напівпровідникового зразка напівпровідник включається до схеми для точного вимірювання малих змін провідності. Набагато зручніше використовувати схему нестационарного ефекту поля. Для цього до обкладинок конденсатора прикладають змінну напругу низької частоти. Тоді, подаючи на один із входів осцилографа сигнал, пропорційний прикладеній напрузі, а на інший – сигнал, пропорційний зміні провідності, можна отримати залежність G_S від V . За допомогою ефекту поля можна отримати цінну інформацію про поверхневі стани напівпровідника. Дійсно, електрони та дірки у напівпровіднику створюють заряд, що віднесений на одиницю площі півки та можна записати як

$$Q = Q_S + e(\Gamma_p - \Gamma_n), \quad (13.98)$$

де Q_S – заряд, зв'язаний на поверхневих рівнях, а величина $e(\Gamma_p - \Gamma_n)$ визначає рухомий заряд у приповерхневому шарі, спричинений рухом надлишкових дірок та електронів, повна кількість яких визначається величинами Γ_p та Γ_n із (13.76) і (13.77). Крім цих зарядів у системі існує ще один, що створюється іонами домішок в об'ємі напівпровідника. За відсутністю поля сумарний заряд іонізованих домішок компенсується зарядом Q . У більшості випадків можна вважати, що в ефекті поля заряд іонів не змінюється. Тоді індукований заряд за прикладення зовнішнього поля змінюватиметься на величину

$$\delta Q = \delta Q_S + e(\delta\Gamma_p - \delta\Gamma_n). \quad (13.99)$$

Оскільки величини $\delta\Gamma_p$ і $\delta\Gamma_n$ можна визначити із (13.92) та (13.93) за відомої величини поверхневого потенціалу Y_S , який, у свою чергу, також легко визначається із вимірювань ефекту поля, то величина δQ_S однозначно визначається зміною заряду δQ . Але індукований заряд δQ у простій схемі плоского конденсатора визначається його ємністю C , що віднесена на одиницю площини, і різницею потенціалу V , що є експериментальним параметром $\delta Q = CV$. Таким чином

$$\delta Q_S = CV - e(\delta\Gamma_p - \delta\Gamma_n). \quad (13.100)$$

Іншими словами, можна експериментально знаходити залежність Q_S від Y_S . У свою чергу такі залежності дозволяють визначати енергетичну структуру поверхневих рівнів E_S та їхню концентрацію. Дійсно, за зміни Y_S поверхневі рівні так само, як і краї зон у приповерхневому шарі зміщуються щодо

рівня Фермі. При проходженні будь-якого рівня E_S через F зарядовий стан рівня зміниться. За наявності рівнів одного типу така зміна зарядового стану спричинить появу східців на кривих залежності Q_S від Y_S . Нехай E_S відраховується за невикривлених зон від рівня Фермі власного напівпровідника F_i . Тоді, визначаючи $\varepsilon_S = E_S - F_i$, дістанемо

$$E_S - E_F = \varepsilon_S - (E_F - F_i) - kTY_S. \quad (13.101)$$

Оскільки в об'ємі напівпровідника

$$E_F - F_i = \frac{1}{2} kT \ln \frac{n_0}{p_0}, \quad (13.102)$$

знайдемо ймовірність заповнення рівня E_S за заданої величини Y_S . Для цього використаємо функцію розподілу заповнення локалізованих рівнів

$$f(E) = \left(1 + \frac{g_1}{g_0} \exp\left(\frac{E_F - E}{kT}\right) \right)^{-1}, \quad (13.103)$$

де знехтуємо можливість виродження цих рівнів (тобто приймемо, що $g_1/g_0 = 1$). Маємо

$$f(\varepsilon_S, Y_S) = \left(1 + \exp\left(\frac{\varepsilon_S}{kT} - \frac{1}{2} \ln \frac{n_0}{p_0} - Y_S\right) \right)^{-1}. \quad (13.104)$$

Можна визначити заряд, обумовлений іонізацією цих рівнів:

- ♦ **для акцепторних рівнів із поверхневою концентрацією S_A**

$$Q_S = -eS_A f(\varepsilon_S, Y_S); \quad (13.105)$$

- ♦ **для донорних рівнів із поверхневою концентрацією S_D**

$$Q_S = eS_D [1 - f(\varepsilon_S, Y_S)]. \quad (13.106)$$

Проведений аналіз показав, що для обох випадків на кривих залежності Q_S від Y_S з'являтиметься східцеподібна особливість і точка перегину, що визначається нулем у показнику експоненти. Ця точка відповідає потенціалу

$$Y_S^{(i)} = \frac{\varepsilon_S}{kT} - \frac{1}{2} \ln \frac{n_0}{p_0}. \quad (13.107)$$

Таким чином, визначаючи із даних вимірювань ефекту поля точку $Y_S^{(i)}$, можна знайти ε_S . Концентрацію поверхневих рівнів можна знайти, вимірюючи величину східцеподібного стрибка, який дорівнює ev .

Експериментальні дослідження ефекту поля намагаються проводити так, щоб потенціал V на керуючому електроді (затворі) набагато перевищував поверхневий потенціал у напівпровіднику. Це дає змогу визначати

індукований заряд як $\frac{d^2\varphi}{dx^2} = -\frac{4\pi e}{\varepsilon}(n_0 - n(x))$ і будувати експериментальні

залежності $G_S(Q)$. Таким самим чином можна побудувати й теоретичну залежність: відома ємність конденсатора дає змогу визначати індукований заряд, що у свою чергу, – визначити поверхневий потенціал. А визначений поверхневий потенціал дозволяє знайти поверхневу провідність. Варто зауважити, що реально залежність поверхневої провідності від індукованого заряду, яка спостерігається в експерименті, суттєво відрізняється від залежності, отриманої теоретично. На це є кілька причин. По-перше між напівпровідником і металом існує контактна різниця потенціалів, що не враховується у простих моделях ефекту поля. По-друге, межа розподілу **напівпровідник-діелектрик** і приповерхнева область діелектрика зазвичай містять стани, що несуть некомпенсований заряд. При цьому такі стани є за своєю природою глибокими рівнями, які не змінюють заряд за експериментально можливих значень поверхневого електростатичного потенціалу. Крім того, існують стани, що можуть бути заповнені електронами за тих значень Y_S , які можуть бути реалізовані в експерименті. Таким чином, у реальній ситуації виникають кілька неконтрольованих моментів, що не дозволяють будувати адекватні моделі. Наприклад, перші дві із зазначених причин приводять до наявності у напівпровідника поверхневої провідності у нульовому полі (при $V = 0$). Наявність такої провідності викликає паралельний зсув експериментальної кривої щодо теоретичної $G_S(Q)$ уздовж вісі зарядів на деяку величину Q_{S0} . Електронні стани у приповерхневій області діелектрика спричиняють не тільки зсув, але й деформацію форми кривої $G_S(Q)$. Отже, якісні та кількісні відмінності між розрахованими у рамках простої теорії та експериментально отриманими кривими залежності поверхневої провідності від індукованого на поверхні заряду дають змогу визначати концентрацію та енергетичні спектри поверхневих станів у структурі **напівпровідник-діелектрик**. Дійсно, за наявності поверхневих станів лише частина індукованого у напівпровіднику заряду Q_{Sp} бере участь у створенні поверхневої провідності. Інша частина заряду Q_{Ss} захоплюється на поверхневі стани. Тому

$$CV = \mp Q_{Sp} \mp Q_{Ss}. \quad (13.108)$$

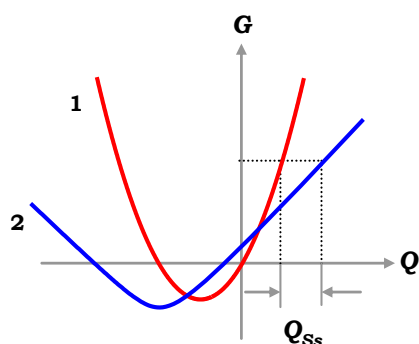


Рис. 13.13. Поверхнева провідність як функція індукованого заряду:
 1 – теоретичний розрахунок;
 2 – експеримент $Q_{Ss} = Q_{ind} - Q_{Sp}$

На рис. 13.13 подано теоретичну та експериментальну криві залежності поверхневої провідності від індукованого заряду. Вплив нормального електричного поля на величину поверхневого електростатичного потенціалу Y_S та поверхневої провідності G_S визначається співвідношенням між змінами заряду Q_{Sp} у приповерхневій області та захопленого заряду Q_{Ss} за заданої величини V . Зрозуміло, що чим більшою буде величина зміни заряду Q_{Ss} за зміни потенціалу V , тим меншими будуть значення зміни поверхневого електростатичного по-

тенціалу ΔY_S і поверхневої провідності ΔG_S . Оскільки за заданого знаку похідної dG_S/dQ_{ind} поверхнева провідність визначається тільки величиною Y_S , то очевидно, що за різних значень захопленого заряду Q_{Ss} необхідно індукувати у напівпровідник різний заряд для досягнення одного й того самого значення поверхневої провідності. Різниця між індукованими зарядами для досягнення заданої величини поверхневої провідності в експериментальній і теоретичній кривих і дає величину захопленого заряду за визначеного значення Y_S . Зауважимо, що обробка даних щодо ефекту поля можлива лише, якщо експериментальна крива $G_S(Q)$ має мінімум, який свідчить, що щільність станів, які здатні змінювати заряд у процесі вимірювань, є відносно малою, а напруженість поля, нормального до поверхні зразку, – досить великою. Значення G_S , яке відповідає мінімуму, визначається рівнянням

$$\frac{\mu_n}{\mu_p} e^{2e\phi_0/kT} = e^{-Y_S}, \quad (13.109)$$

що легко отримати з умови $dG_S/dQ = 0$ (див. додаток N). Цю величину можна прийняти за точку відліку абсолютних значень Y_S . Величину поверхневої провідності, яку можна знайти експериментально, також зручно відраховувати від її значення у мінімумі. Для цього мінімум експериментальної кривої сполучають із мінімумом теоретичної (тобто будують криву, паралельно зсунуту вздовж осі G_S щодо отриманої експериментально (рис. 13.13)).

Варто зауважити, що дані досліджень ефекту поля на неякісних зразках показали, що вимірювання величини сталого струму після прикладення

напруги до польового електроду поверхнева провідність релаксує до нуля за всіх можливих значень напруги V незалежно від її знаку. Цей факт свідчить про наявність у такому зразку великої кількості акцепторних і донорних станів на межі розподілу **напівпровідник-діелектрик**, які захоплюють індукований заряд та екранують об'єм напівпровідника від проникнення електричного поля. Детальне вивчення кінетики ефекту поля показало, що існують дві області спаду на кривій поверхневої провідності, які характеризуються різними часами релаксації. Цей факт дозволяє припустити, що у приповерхневій області існують два типи локалізованих станів із різними часами захоплення – так звані швидкі та повільні стани. До швидких станів, що характеризуються часом захоплення $\sim 10^{-3} \div 10^{-8}$ с, належать такі, які розташовуються безпосередньо на межі розподілу. До повільних станів, що характеризуються часами захоплення від кількох мілісекунд до кількох годин, належать стани, які розташовуються на зовнішньому боці діелектрика (окисла). Застосування в експериментах змінної напруги відповідної частоти на затворі дозволяє виключити вплив повільних поверхневих станів на результати вимірювань ефекту поля.

■ 13.4. Поверхнева рекомбінація

У розд. 11 обговорювались генераційно-рекомбінаційні процеси у напівпровідниках. При цьому одним із найбільш ефективних таких механізмів була рекомбінація через локальні пастки, роль яких відігравали різноманітні домішки в об'ємі напівпровідника. Поверхня характеризується великою кількістю локальних станів, що зумовлені як наявністю самої поверхні, так і домішками, що розташовані у приповерхневій області. Такі

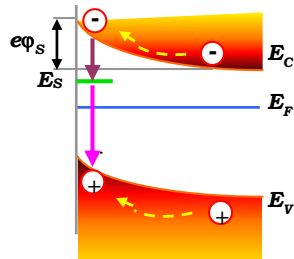


Рис. 13.14. Рекомбінація через поверхневий рівень

поверхневі локальні рівні можуть виступати у ролі ефективних рекомбінаційних пасток (рис. 13.14). Вплив поверхні на кінетику електронних процесів характеризують швидкістю поверхневої рекомбінації. Із такою ідеєю вперше виступив Шоклі, розглядаючи щільності потоку носіїв заряду до поверхні

$$\mathbf{j}_n / e = \mathbf{s}_n \delta n, \quad \mathbf{j}_p / e = \mathbf{s}_p \delta p, \quad (13.108)$$

де \mathbf{j}_n та \mathbf{j}_p – щільності струму електронів і дірок, що течуть до поверхні, δn і δp – надлишкові кон-

центрації електронів і дірок, відповідно. Вважається, що такі направлені потоки зумовлені наявністю на поверхні напівпровідника рекомбінаційних центрів, які відіграють роль місць стоку носіїв. Величини \mathbf{s}_n та \mathbf{s}_p мають розмірність см/с (в СІ – м/с) і називаються швидкостями поверхневої рекомбінації електронів і дірок.

Фізичну інтерпретацію s як швидкості стоку можна зрозуміти, розглядаючи, наприклад щільність потоку дірок $j_p^{(v-s)}$ з об'єму на поверхню, який можна записати у вигляді

$$j_p^{(v-s)} = \frac{1}{4} e v_{Ts} p_s, \quad (13.109)$$

де v_{Ts} – середня теплова швидкість, p_s – концентрація дірок у приповерхневій області. Щільність потоку дірок від поверхні у глибину зразка складається із двох потоків – потоку частинок, що відбиваються від поверхні, і потоку дірок $\mu_n n_0 \exp[(Y_S)_{\min}/2] = \mu_p p_0 \exp[-(Y_S)_{\min}/2]$, згенерованих на поверхні

$$j_p^{(s-v)} = \frac{1}{4} e v_{Ts} p_s R_p + j_p^{(gs)}, \quad (13.110)$$

де R_p – середня ймовірність відбиття дірки поверхнею. Якщо ввести аналогічно формулі (13.108) швидкість поверхневої генерації дірок $(\mu_n / \mu_p) < 1$, то у стані рівноваги, коли потоки частинок до поверхні та від неї компенсують один одного, швидкість рівноважної генерації дірок визначатиметься формулою

$$G_{S0} = \frac{1}{4} (1 - R_p) v_{Ts} p_{s0}. \quad (13.111)$$

Зрозуміло, що у слабонерівноважному стані, але в умовах стаціонарних потоків (оскільки має виконуватись закон збереження заряду), можна записати аналогічно (13.111)

$$G_S = \frac{1}{4} (1 - R_p) v_{Ts} p_s. \quad (13.112)$$

Таким чином, швидкість генерації нерівноважних дірок дається різницею $G_S - G_{S0}$

$$\delta G_S = \frac{1}{4} (1 - R_p) e v_{Ts} \delta p_s. \quad (13.113)$$

Оскільки на поверхні рекомбінують частинки, згенеровані у приповерхневій області, із (13.108) та (13.113) знайдемо

$$s_p = \frac{1}{4}(1 - R_p)v_{Ts}. \quad (13.114)$$

Зрозуміло, що аналогічний вираз можна отримати й для електронів. Таким чином, можна стверджувати, що для поверхонь із низькою здатністю відбиття частинок ($R \approx 0$) швидкість поверхневої рекомбінації становить за порядком величини середню теплову швидкість носіїв, і для більшості використовуваних напівпровідників (Si, Ge, GaAs) може досягати 10^5 см/с. Обмеження швидкості поверхневої рекомбінації має прозорий фізичний сенс: зростання рекомбінації через поверхню обмежено швидкістю дифузійної доставки нерівноважних носіїв заряду з об'єму на поверхню. Оскільки швидкість поверхневої рекомбінації значно залежить від здатності поверхні до відбиття частинок, то вона може змінюватись у широких межах залежно від стану поверхні, тобто від її механічної обробки, вигину зон тощо.

Таким чином, введення швидкості поверхневої рекомбінації за Шоклі означає, що результуючий потік частинок до поверхні напівпровідника визначається різницею між потоком частинок, що захоплюються на поверхневій рівні E_S , і потоком частинок, термічно викинутих із цих рівнів у зону. Іншими словами, ця швидкість залежить тільки від пропускної здатності поверхневого рекомбінаційного рівня та є повним аналогом рекомбінації за механізмом Шоклі-Ріда в об'ємі. Тоді для швидкості зміни концентрації дірок у результаті рекомбінації через поверхневі рівні

$$R_p^{(S)} = -\frac{d\delta p_S}{dt} \quad (13.115)$$

можна використати формулу (див. п. 10.2)

$$R_p^{(S)} = \gamma_{ns}\gamma_{ps}N_S \frac{n_S p_S - n_1 p_1}{\gamma_{ns}(n_S + n_1) + \gamma_{ps}(p_S + p_1)}, \quad (13.116)$$

де γ_{ns} та γ_{ps} – ймовірності захоплення поверхневим рівнем електронів і дірок за одиницю часу (величини мають розмірність см³/с); n_S і p_S – нерівноважні поверхневі концентрації електронів і дірок, n_1 і p_1 – рівноважні концентрації електронів і дірок, якщо рівень Фермі збігається із рівнем рекомбінаційної пастки E_S (тобто виконується закон діючих мас, $n_1 p_1 = n_0 p_0 = n_i^2$). Оскільки у стаціонарних умовах швидкості зміни концентрацій електронів і дірок мають дорівнювати одна одній, звідси випливає, що

$$R_p^{(S)} = R_n^{(S)} = R_S, \quad (13.117)$$

тобто у приповерхневій області потоки дірок та електронів до поверхні мають бути рівними. Інакше на поверхні напівпровідника може відбуватися накопичування зарядів одного знаку, що спричинить виникнення додаткового електричного поля, яке викличе зміни потоків носіїв заряду. Тим самим система не перебуватиме у стаціонарному стані. Результуючий потік електронів і дірок до поверхні з'являється завдяки градієнту надлишкових носіїв заряду, що виникає у приповерхневій області через різницю у темпах рекомбінації на поверхні та в об'ємі напівпровідника. І якщо швидкість поверхневої рекомбінації невелика, то число електрон-діркових пар, що рекомбінують на поверхні напівпровідника, є невеликим порівняно із потоками частинок до поверхні та від неї. Ці потоки пов'язані із дифузійними та дрейфовими струмами в області просторового заряду. Тому за невеликої швидкості поверхневої рекомбінації можна вважати, що дифузійні та дрейфові потоки практично врівноважують один одного, тобто концентрація носіїв у цих умовах визначатиметься розподілом Больцмана. Якщо вигин зон на поверхні є $e\varphi_s$, а n та p – концентрації нерівноважних електронів і дірок в об'ємі напівпровідника, то на поверхні

$$n_s = ne^{Y_s} \quad \text{і} \quad p_s = pe^{-Y_s}. \quad (13.118)$$

В умовах термодинамічної рівноваги

$$n_{s0} = n_0 e^{Y_s} \quad \text{і} \quad p_{s0} = p_0 e^{-Y_s}. \quad (13.119)$$

Із врахуванням (13.118) і закону діючих мас швидкість зміни концентрації дірок завдяки поверхневій рекомбінації запишемо у вигляді

$$R_p^{(S)} = \gamma_{ns}\gamma_{ps}N_S \frac{np - n_0p_0}{\gamma_{ns}(n_s + n_1) + \gamma_{ps}(p_s + p_1)}. \quad (13.120)$$

Вважатимемо, що надлишкова концентрація носіїв в об'ємі напівпровідника є невисокою, тобто $\delta p \ll p_0$, $\delta n \ll n_0$ і $\delta Q_s = CV - e(\delta\Gamma_p - \delta\Gamma_n)$. Оскільки $p = p_0 + \delta p$, $n = n_0 + \delta n$, то можна записати

$$np - n_0p_0 \cong (n_0 + p_0)\delta p, \quad n_s \cong n_{s0}, \quad p_s = p_{s0}. \quad (13.121)$$

Підставляючи ці вирази до (13.120) і використовуючи (13.119), отримаємо

$$R^{(S)} = \gamma_{ns}\gamma_{ps}N_S \frac{n_0 + p_0}{\gamma_{ns}(n_0 e^{Y_s} + n_1) + \gamma_{ps}(p_0 e^{-Y_s} + p_1)} \delta p. \quad (13.122)$$

Оскільки $R^{(S)} = s\delta p$, запишемо співвідношення, що пов'язує швидкість поверхневої рекомбінації s із поверхневим потенціалом Y_s

$$s = \gamma_{ns}\gamma_{ps}N_S \frac{n_0 + p_0}{\gamma_{ns}(n_0 e^{Y_S} + n_1) + \gamma_{ps}(p_0 e^{-Y_S} + p_1)}. \quad (13.123)$$

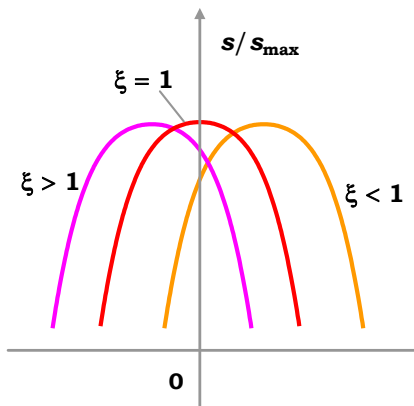


Рис. 13.15. Залежність швидкості поверхневої рекомбінації від поверхневого потенціалу

Ця формула у багатьох випадках задовільно описує характер поведінки швидкості поверхневої рекомбінації за зміни поверхневого потенціалу, зокрема вона показує, що крива залежності $s = s(Y_S)$ має максимум за деякого значення $Y_S = Y_S^{\max}$ (рис. 13.15). Для оцінки величини Y_S^{\max} можна дорівняти до нуля похідну від знаменника у (13.123)

$$Y_S^{\max} = \frac{1}{2} \ln \left(\frac{p_0 \gamma_{sp}}{n_0 \gamma_{sn}} \right). \quad (13.124)$$

Визначаючи з експериментальних даних величину поверхневого потенціалу Y_S за відомих співвідношень $\xi = p_0/n_0$, можна

знайти відношення коефіцієнтів захоплення γ_{sp}/γ_{sn} для даного поверхневого рівня. Якщо величина $\gamma_{sp}/\gamma_{sn} \gg 1$, то (оскільки захоплення за наявності кулонівського притягання зазвичай характеризуються великими значеннями коефіцієнта захоплення) скоріше за все зарядовий стан рівня від'ємний або нейтральний, тобто рівень – акцепторний. Якщо ж $\gamma_{sp}/\gamma_{sn} \ll 1$, то із великою ймовірністю рівень є донорним. Зазвичай із всієї сукупності поверхневих рівнів тільки частина із них бере участь у процесах рекомбінації. Інші ж рівні можуть обмінюватись носіями заряду тільки з однією із зон. Тобто вони є рівнями "прилипання".

■ 6.7. Задачі

1. Визначити хід потенціалу у збагаченому приповерхневому шарі напівпровідника n -типу. Вважаючи в об'ємі напівпровідника донори повністю іонізованими, оцінити товщину шару збагачення.

Розв'язок. Оскільки всі донори у напівпровіднику іонізовані, рівняння Пуассона, що визначає хід потенціалу поблизу поверхні, запишемо у вигляді

$$\frac{d^2\varphi}{dx^2} = -\frac{4\pi e}{\varepsilon} (n_0 - n(x)). \quad (13.125)$$

Як граничні умови використовуватимемо, те що в об'ємі напівпровідника потенціал і його похідна зникають, тобто

$$\text{при } x \rightarrow \infty, \quad \varphi \rightarrow 0 \quad \text{і} \quad \frac{d\varphi}{dx} \rightarrow 0. \quad (13.126)$$

На поверхні напівпровідника потенціал має бути сталим і від'ємним:

$$\text{при } x = 0, \quad \varphi = \varphi_s < 0. \quad (13.127)$$

Оскільки згідно із (13.82) $n(x) = n_0 e^{e\varphi/kT}$, рівняння (13.125) перепишемо у вигляді

$$\frac{d^2\varphi}{dx^2} = -\frac{4\pi en_0}{\varepsilon} (1 - e^{e\varphi/kT}). \quad (13.128)$$

Діючи так само, як при виведенні (13.91), отримаємо

$$\frac{e}{kT} \frac{d\varphi}{dx} = L_D^{-1} [e^{e\varphi/kT} - e\varphi/kT - 1]^{1/2}, \quad (13.129)$$

де
$$L_D = L_S / \sqrt{2}, \quad L_S = \sqrt{\varepsilon kT / (4\pi e^2 n_0)} \quad (13.130)$$

– дебаєвський радіус екранування. Звідси маємо, що залежність потенціалу від координати визначається таким інтегралом

$$\int_{e\varphi/kT}^{e\varphi_s/kT} [e^u - u - 1]^{-1/2} du = x / L_D. \quad (13.131)$$

Цю формулу можна аналізувати чисельно. Але для визначення товщини шару збагачення L_0 , що можна визначити з умови $e\varphi(L_0)/kT = 1$, цей інтеграл можна оцінити таким чином.

При $e\varphi_s/kT \gg 1$ можна наближено записати

$$\int_1^{\infty} [e^u - u - 1]^{-1/2} du \approx 1,2 = L_0 / L_D. \quad (13.132)$$

Звідси маємо, що $L_0 \approx 1,2 \cdot L_D$, тобто навіть за значного збагачення ширина шару збагачення не значно відрізняється від довжини дебаєвського екранування.

2. Нормально до плоскої поверхні власного напівпровідника прикладено стале слабке електричне поле. Як воно впливає на вигин зон?

Розв'язок. Нехай прикладене поле настільки слабке, що можна вважати, що в об'ємі напівпровідника $e\varphi/kT \ll 1$. Запишемо рівняння, якому задовольняє потенціал поля

$$\frac{d^2\varphi}{dx^2} = -\frac{4\pi e}{\varepsilon} (p(x) - n(x)). \quad (13.133)$$

Концентрації носіїв заряду визначаються формулами

$$n = n_i e^{e\varphi(x)/kT}, \quad p = n_i e^{-e\varphi(x)/kT}, \quad (13.134)$$

де n_i – концентрація носіїв в об'ємі напівпровідника за даної температури.

Рівняння (13.133) необхідно доповнити граничними умовами

$$\varphi = 0 \quad \text{при} \quad x \rightarrow \infty, \quad (13.135)$$

$$E = -\varepsilon d\varphi/dx \quad \text{при} \quad x = 0, \quad (13.136)$$

Підставимо вирази (13.134) до рівняння (13.133)

$$\frac{d^2\varphi}{dx^2} = -\frac{4\pi e}{\varepsilon} n_i \left[e^{-e\varphi/kT} - e^{e\varphi/kT} \right]. \quad (13.137)$$

Розкладаючи праву частину за малим параметром $e\varphi/kT$, отримаємо

$$\frac{d^2\varphi}{dx^2} = \frac{8\pi e^2 n_i}{\varepsilon kT} \varphi, \quad (13.138)$$

або, оскільки $\sqrt{\frac{\varepsilon kT}{8\pi e^2 n_i}} = L_D$,

$$\frac{d^2\varphi}{dx^2} - \frac{\varphi}{L_D^2} = 0. \quad (13.139)$$

Розв'язок рівняння (13.139) шукаємо у вигляді

$$\varphi(x) = C_1 e^{-x/L_D} + C_2 e^{x/L_D}. \quad (13.140)$$

Коефіцієнти C_1 та C_2 знайдемо із граничних умов:

$$x = \infty, \quad \varphi(\infty) = C_1 \cdot 0 + C_2 e^\infty \rightarrow C_2 = 0$$

$$\frac{d\varphi}{dx} = -C_1 \frac{1}{L_D} e^{-x/L_D} \Big|_{x=0} = -\frac{E}{\varepsilon} \rightarrow C_1 = \varepsilon L_D / E$$

Таким чином, хід потенціалу у приповерхневій області напівпровідника

$$\varphi(x) = \frac{\varepsilon L_D}{E} e^{-x/L_D}. \quad (13.141)$$

Оскільки електростатичний потенціал викликає вигин зон, то в умовах, коли зовнішнє поле направлено вглиб напівпровідника (тобто в його приповерхневій області накопичуватимуться електрони), зони вигинатимуться вниз

$$E_C(x) = E_{C0} - (eEL_D / \varepsilon) \cdot e^{-x/L_D}, \quad (13.142)$$

$$E_V(x) = E_{V0} - (eEL_D / \varepsilon) \cdot e^{-x/L_D}. \quad (13.143)$$

3. Визначити, як відрізняється потенціал у точці інверсії у кремнії і германії за кімнатної температури, якщо обидва зразка леговані акцепторними домішками так, що концентрації дірок у них однакова і становить $p_0 = 10^{15} \text{ см}^{-3}$. Вважати, що концентрація носіїв заряду у

власних кремнії і германії Si: $n_i = 1,05 \cdot 10^{10} \text{ см}^{-3}$; Ge: $n_i = 2,2 \cdot 10^{13} \text{ см}^{-3}$. Відношення рухливості електронів і дірок: Si – $b_{\text{Si}} = 2,9$; Ge – $b_{\text{Ge}} = 2,1$.

Список літератури

1. *Смит Р.* Полупроводники. – М.: Мир, 1982.
3. *Зи С.М.* Физика полупроводниковых приборов. – М.: Мир, 1973.
4. *Бонч-Бруевич В.Л., Калашников С.Г.* Физика полупроводников. – М.: Наука, 1977.
5. *Пека Г.П., Стріха В.І.* Поверхневі та контактні явища у напівпровідниках. – К.: Либідь, 1992.
6. *Kasap S.O.* Principles of Electronic Materials and Devices, McGraw-Hill, 2005.

Розділ 14

КОНТАКТНІ ЯВИЩА У НАПІВПРОВІДНИКАХ

У попередньому розділі було розглянуто ефекти, пов'язані з існуванням меж розподілу в напівпровідниках і показано, що наявність меж розподілу спричиняє появу цілої низки цікавих фізичних явищ. Іншою важливою обставиною, що спонукає детальніше розглядати ефекти, обумовлені наявністю меж розподілу, є той факт, що, по-перше, на практиці ми завжди маємо справу із просторово обмеженими системами, по-друге, сучасні технології напівпровідникових приладів великою мірою базуються на використанні так званих планарних систем, тобто структур, що складаються із набору плоских шарів напівпровідників, металів і діелектриків.

Оскільки вся сучасна електроніка базується на планарних структурах, важливою характеристикою яких є взаємодія різного типу плоских контактів, розглянемо окремо, що відбувається на межах розподілу **метал-метал**, **метал-напівпровідник** та **напівпровідник-напівпровідник**. Для такого розгляду виникає необхідність визначення деякої енергетичної характеристики, яка була б універсальною величиною для різного типу матеріалів. Такою є термоелектронна робота виходу, що має сенс енергії, яку необхідно додати до системи для вилучення електрона із твердого тіла і переведення його у вакуум.

■ 14.1. Термоелектронна робота виходу

Для переведення електрона із твердого тіла у вакуум необхідно витратити деяку енергію, тобто можна стверджувати, що на межі розподілу **тверде тіло-вакуум** існує енергетичний бар'єр, що заважає виходу електронів із кристалу. Зрозуміло, що для того, щоб електрон міг покинути зразок, він повинен мати енергію, достатню для подолання цього бар'єру. Чим вищою буде температура, тим більша частина електронів буде здатною подолати бар'єр. Вихід електронів із речовини внаслідок теплового збудження називається термоелектронною емісією. Для визначення густини струму термоелектронної емісії підсумуємо число електронів, що спроможні подолати потенціальний бар'єр і вийти із напівпровідникового зразка. Нехай E_{vac} – енергія електрона, що вийшов із напівпровідника у вакуум і перебуває у стані спокою

щодо зразка. Тоді енергія, необхідна для переведення електрона із дна зони провідності у вакуумний стан, який характеризується нульовою кінетичною енергією, можна визначити як

$$\chi = E_{vac} - E_C. \quad (14.1)$$

Ця енергія має спеціальну назву – енергія електронної спорідненості. Ця характеристика зазвичай використовується для визначення взаємного положення дна зони провідності та стелі валентної зони за контакту двох напівпровідників (див. далі).

При вивченні ефектів на контактах важливу роль відіграє потенціальний бар'єр, що необхідно подолати електрону для переходу з одного середовища до іншого. Для подолання потенціального бар'єру кінетична енергія електрона має бути більшою за висоту бар'єру. Відомості про висоту цього бар'єру – дуже важливі з практичного погляду. Уявимо, що досліджуваний зразок поміщується у деяку адіабатичну оболонку, що підтримується при температурі T . Електронний газ над зразком знаходитиметься у стані термодинамічної рівноваги зі зразком:

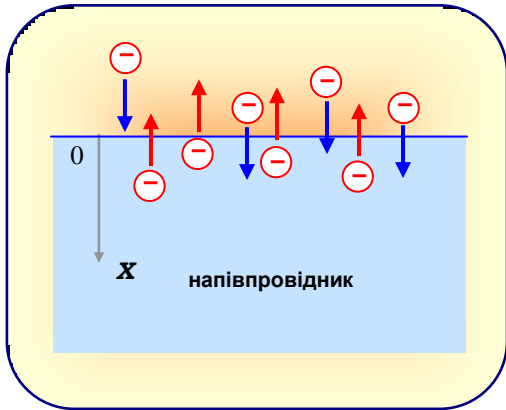


Рис. 14.1. До визначення термоелектронної енергії виходу

кількість електронів, що випаровуються із напівпровідника (або метала) дорівнює кількості електронів, що потрапляють із вакууму в напівпровідник. Іншими словами, якщо розглядати рух електрона вздовж вісі OX (рис. 14.1), то для того, щоб електрон потрапив у зразок, його кінетична енергія має бути більшою за потенціальний бар'єр V_{vac}

$$\frac{m_0 v_x^2}{2} \geq V_{vac}. \quad (14.2)$$

Кількість станів одиничного об'єму в інтервалі швидкостей $(\mathbf{v}, \mathbf{v} + d\mathbf{v})$ визначається формулою

$$dN = 2 \left(\frac{2\pi m_0}{\hbar} \right)^3 dv_x dv_y dv_z. \quad (14.3)$$

Кількість електронів, що завдяки тепловому руху характеризуються швидкостями з інтервалу $(\mathbf{v}, \mathbf{v} + d\mathbf{v})$

$$dn = f_0 dN = 2 \left(\frac{2\pi m_0}{\hbar} \right)^3 e^{-(E-E_F)/kT} dv_x dv_y dv_z. \quad (14.4)$$

Вважатимемо, що всі електрони, які рухаючись вздовж вісі ОХ, подолали потенціальний бар'єр, не повертаються до зовнішнього середовища. Таким чином, до поверхні напівпровідника рухається потік електронів зі швидкістю v_x , що утворює струм, який протікає із вакууму до зразка

$$I = e \int dnv_x. \quad (14.5)$$

Підставляючи сюди (14.4), отримаємо

$$I = 2e \left(\frac{2\pi m^*}{\hbar} \right)^3 e^{E_F/kT} \int_0^\infty dv_x v_x \int_{-\infty}^\infty dv_y \int_{-\infty}^\infty dv_z e^{-E/kT}. \quad (14.6)$$

Нижня границя інтегрування у першому інтегралі (14.6) вибрана із таких міркувань: для того, щоб електрон міг потрапити до зразка, його швидкість має бути направлена у бік зразка (вздовж вісі ОХ). Оскільки енергія електрона у вакуумі $E = E_{vac} + m_0(v_x^2 + v_y^2 + v_z^2)/2$, струм (14.6) визначатиметься формулою

$$I = 2e \frac{(4\pi^2 m_0)^2}{\hbar^3} kT \cdot e^{-(E_{vac} - E_F)/kT} \int_0^\infty dv_x v_x e^{-(m_0/2kT)v_x^2}. \quad (14.7)$$

При цьому використано відоме співвідношення $\int_{-\infty}^\infty e^{-at^2} dt = \sqrt{\pi/a}$. Обчислюючи інтеграл у (14.7), отримуємо

$$\int_0^\infty v_x dv_x e^{-(m_0/2kT)v_x^2} = \frac{kT}{m_0}. \quad (14.8)$$

Оскільки в умовах термодинамічної рівноваги такий самий струм протікатиме із напівпровідника до вакууму, використовуючи формулу (14.8), із виразу (14.7) отримуємо для густини струму термоелектронної емісії

$$I = AT^2 e^{-\Phi/kT}, \quad (14.9)$$

де $A = 2(2\pi)^4 e m_0 k^2 \hbar^{-3}$ – універсальна константа, а термоелектронна робота виходу

$$\Phi = E_{vac} - E_F. \quad (14.10)$$

Виключаючи із цього виразу величину E_{vac} , за допомогою (14.1) отримуємо

$$\Phi = \chi + E_C - E_F. \quad (14.11)$$

Ця енергія для напівпровідників становить величину порядку кількох електрон-вольт. Оскільки термоелектронна робота виходу визначається положенням рівня Фермі, то вона залежить від температури,

концентрації домішок і природи напівпровідника. Наприклад, для власного напівпровідника, де $E_F = (E_C + E_V)/2 + (3/4)kT \ln(m_{pd}^* / m_{nd}^*)$, робота виходу

$$\Phi_i = \chi + E_g / 2 + (3kT / 4) \ln(m_{nd}^* / m_{pd}^*) \quad (14.12)$$

визначається шириною забороненої зони і відношенням ефективних мас електронів і дірок. Для напівпровідника n -типу у випадку слабкої іонізації домішок, де $E_F = (E_C + E_D) \cdot 2 - kT \ln(2N_C / N_D)$, термоелектронна робота виходу

$$\Phi_n = \chi + (E_C - E_D) / 2 + (kT / 2) \ln(2N_C / N_D). \quad (14.13)$$

У випадку сильної іонізації домішок, коли $E_F = E_C + kT \ln(N_C / N_D)$, маємо

$$\Phi_n = \chi + kT \ln(N_C / N_D). \quad (14.14)$$

Аналогічно для напівпровідника p -типу у випадку слабкої іонізації домішок

$$\Phi_p = \chi + E_g + (E_V - E_A) / 2 + (kT / 2) \ln(2N_V / N_A). \quad (14.15)$$

У випадку сильної іонізації домішок термоелектронна енергія виходу акцепторного напівпровідника є

$$\Phi_p = \chi + E_g - kT \ln(N_V / N_A). \quad (14.16)$$

■ 13.2. Контакти метал-метал та метал-напівпровідник

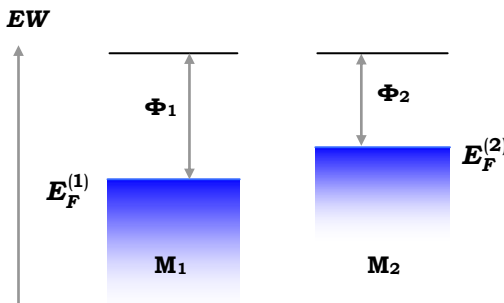


Рис. 14.2. Енергетичні діаграми двох ізольованих металів

Розглянемо два металічних зразки, що мають плоскі поверхні та характеризуються роботою виходу Φ_1 та Φ_2 , і $\Phi_1 > \Phi_2$, тобто енергія Фермі $E_F^{(1)}$ першого металу є меншою за енергію Фермі $E_F^{(2)}$ другого (рис. 14.2). Утворимо контакт цих зразків. **Метали почнуть обмінюватися електронами, й оскільки $E_F^{(2)} > E_F^{(1)}$, в основному електрони із другого металу переходитимуть до**

першого: на першому зразку з'явиться негативний заряд, а на другому – позитивний. У результаті перпендикулярно до площини контакту виникне електричне поле, яке буде локалізоване у приповерхневій

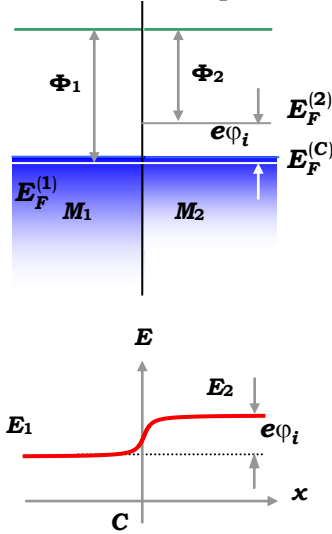


Рис. 14.3. Контакт двох металів.
Виникнення контактної різниці потенціалів

області товщиною порядку довжини екранування (у металах ця величина становить $\sim 10^{-8}$ см). Завдяки потоку електронів із другого металу до першого почнуть вирівнюватись рівні Фермі. Коли рівні Фермі вирівнюються й в об'єднаному зразку виникне один рівень Фермі $E_F^{(C)}$ (рис. 14.3), струм електронів припиниться. При цьому на контакті у рівноважному стані встановиться деяка різниця потенціалів ϕ_i , що називають контактною різницею потенціалів. Для визначення величини контактної різниці потенціалів обчислимо потік електронів, що переходять із металу 1 до металу 2, і навпаки. Так само як при визначення термоелектронної роботи виходу знайдемо, що із металу 2 до металу 1 перейдуть

$$dN_2 = 2 \frac{(2\pi)^3}{m_2^* \hbar^3} p_x dp_x \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z \{e^{(E_2 - E_F^{(2)})/kT} + 1\}^{-1} \quad (14.17)$$

електронів. Аналогічно у зворотному напрямку із зразка 1 до зразка 2 перейдуть

$$dN_1 = 2 \frac{(2\pi)^3}{m_1^* \hbar^3} p_x dp_x \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z \{e^{(E_1 - E_F^{(1)})/kT} + 1\}^{-1} \quad (14.18)$$

електронів. Оскільки зразок 2 на межі розподілу заряджений позитивно, то його електрони набудуть додаткової енергії

$$E_2^{(e)} = -e\phi_i \quad (14.19)$$

так, що повна енергія електронів у металі 2 визначатиметься формулою

$$E_2^{(tot)} = E_2 - e\phi_i. \quad (14.20)$$

Оскільки термодинамічна рівновага означає, що число прямих переходів електронів (із зразка 2 до зразка 1) дорівнює числу зворотних переходів, то має виконуватись умова

$$dN_1 = dN_2. \quad (14.21)$$

З іншого боку, за термодинамічної рівноваги повна енергія електронів в обох частинах системи має бути однаковою, тобто

$$E_1 = E_2 - e\phi_i . \quad (14.22)$$

Крім того, умова термодинамічної рівноваги вимагає також рівності диференціалів енергії електронів у різних частинах системи, а саме $dE_2 = dE_1$, звідки

$$\frac{\partial E_1}{\partial p_x^{(1)}} dp_x^{(1)} = \frac{\partial E_2}{\partial p_x^{(2)}} dp_x^{(2)} \text{ або } (m_1^*)^{-1} p_x^{(1)} dp_x^{(1)} = (m_2^*)^{-1} p_x^{(2)} dp_x^{(2)} .$$

Маючи на увазі ці вирази та (14.17) і (14.18), запишемо рівняння (14.21) в явному вигляді

$$\int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z \{e^{(E_1 - E_F^{(1)})/kT} + 1\}^{-1} = \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z \{e^{(E_2 - E_F^{(2)})/kT} + 1\}^{-1} . \quad (14.23)$$

Ця рівність виконується тотожно лише за умови рівності підінтегральних виразів, що означає рівність показників ступеню в експонентах

$$E_1 - E_F^{(1)} = E_2 - E_F^{(2)} . \quad (14.24)$$

Звідси, зважаючи на рівняння (14.22) отримуємо

$$\phi_2 = e^{-1}(E_F^{(2)} - E_F^{(1)}) . \quad (14.25)$$

Таким чином, контактна різниця потенціалів на контакті двох металів визначається різницею енергій Фермі ізольованих зразків цих металів. У разі утворення контактної різниці потенціалів виникає електричне поле, величину якого можна оцінити як $E_{cont} \sim \phi_2 / d$, де d – відстань між плоскими поверхнями металів, що утворили спільну систему. Відповідно до теореми Гауса електричне поле пов'язане із зарядом, поверхнева густина якого $\sigma_S \sim \epsilon\phi_2 / d$. За щільного контакту двох металів можна вважати, що відстань d приблизно дорівнює сталій ґратки. Тоді величина поверхневого заряду, що забезпечує існування контактної різниці потенціалів становить $\sim 10^{13} \text{ см}^{-2}$, що на один-два порядки менше за поверхневу концентрацію електронів вільної металічної поверхні, що становить $\sim 10^{15} \text{ см}^{-2}$. Це означає, що електричне поле контактної різниці потенціалів не заглиблюється у метал на відстань, більшу за величину сталої ґратки (10^{-7} – 10^{-8} см), оскільки поверхневим зарядом σ_S , що зумовлений контактною різницею потенціалів, можна знехтувати порівняно із власним поверхневим зарядом металу.

Розглянемо тепер контакт між металом та невиродженим напівпровідником. Нехай для визначеності робота виходу металу Φ_M більше роботи виходу напівпровідника Φ_S . Тоді при дотиканні поверхонь металу та напівпровідника потік електронів із напівпровідника перевищуватиме потік електронів із металу. У результаті метал в області

контакту заряджатиметься негативно, а напівпровідник набуде позитивного заряду. Електричне поле E_{st} , що виникне при цьому, тепер заважатиме проникненню електронів із напівпровідника, і надлишковий струм припиниться. При цьому вирівнюються рівні Фермі металу та напівпровідника та встановиться рівність струмів термоелектронної емісії. Вираз для струму термоелектронної емісії задається (14.9). Тоді, оскільки наявність контактної різниці потенціалів приведе до збільшення термоелектронної роботи виходу напівпровідника на величину $e\phi_C$, із рівності $I_M = I_S$ отримаємо

$$e\phi_C = \Phi_M - \Phi_S. \quad (14.29)$$

Напруженість електричного поля в шарі об'ємного заряду $\rho(x)$

$$E_C = \frac{4\pi}{\varepsilon} \int \rho(x) dx. \quad (14.30)$$

Таким чином, у силу рівності об'ємних зарядів у контактуючих зразках, різниця потенціалів буде пропорційна товщині об'ємного заряду.

$$U = -\int_0^l E_C dx = -E_C \cdot l. \quad (14.31)$$

Як було з'ясовано вище, товщина шару об'ємного заряду в металі $\sim 10^{-7}$ – 10^{-8} см, а в напівпровіднику перевищує цю величину як мінімум на чотири порядки (тобто складає величину $\sim 10^{-4}$ см – див. далі). Таким чином можна стверджувати, що контактна різниця потенціалів ϕ_C практично повністю падає у приконтактній області напівпровідника. Напруженість електричного поля, викликана контактною різницею потенціалів, не перевищує $\sim 10^6$ В/см. З іншого боку, напруженість внутрішньокристалічного поля $\sim 10^8$ В/см. Таким чином, незначне додавання контактної різниці потенціалів не може спричинити перебудову енергетичної структури напівпровідника, тобто змінити ширину забороненої зони, але може призвести до викривлення зон у приконтактній області щодо рівня Фермі. Таке викривлення, у свою чергу, приведе до зміни концентрації носіїв напівпровідника в його приповерхневій області. Таким чином, якщо робота виходу електронів із напівпровідника менша за роботу виходу із металу, то на контакті **метал–напівпровідник** поверхня металу заряджається негативно, а приповерхневий шар напівпровідника – позитивно. Тепер, щоб потрапити у метал, електронам у напівпровіднику необхідно подолати бар'єр $e\phi_C = \Phi_M - \Phi_S$ проти сил електричного поля E_C . Це означає, що в напівпровіднику енергетичні зони викривляються вгору так, що дно зони провідності віддаляється від рівня Фермі F_n , а стеля валентної зони наближається до цього рівня. При цьому поблизу контакту кон-

центрація електронів у зоні провідності зменшується, а концентрація дірок у валентній зоні збільшується порівняно з її об'ємним значенням. Таким чином, поблизу контакту виникне шар із дещо зниженою питомою провідністю. Такий шар, збіднений основними носіями заряду, називають запірним шаром (рис. 14.4 а).

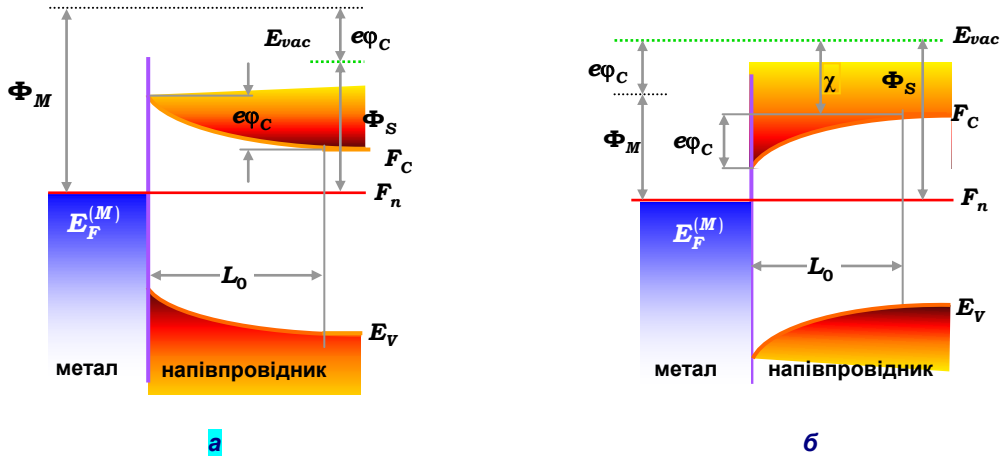


Рис. 14.4. Контакт метал–напівпровідник *n*-типу
 а – $\Phi_M > \Phi_S$ (запірний шар); б – $\Phi_M < \Phi_S$ (антизапірний шар)

Якщо робота виходу із напівпровідника є більшою за роботу виходу із металу, тобто виконується нерівність $\Phi_M < \Phi_S$, то напівпровідник заряджається негативно, і його зони у приконтактній області вигинаються **вниз** – у бік валентної зони (рис. 14.4 б): у приконтактній області число електронів у зоні провідності збільшується, а число дірок у валентній зоні – зменшується. У такому випадку говорять, що в електронному напівпровіднику виникає антизапірний шар. У випадку контакту металу із напівпровідником *p*-типу якісна картина вигину зон збережеться такою, як і для електронного напівпровідника (рис. 14.5), оскільки вигін зон визначається тільки різницею робіт виходу. Головна відмінність для напівпровідника *p*-типу, що контактує з металом, полягає в тому, що у випадку $\Phi_M > \Phi_S$ концентрація дірок у приконтактній області буде більшою за їхню концентрацію в об'ємі, й у приконтактній області виникне збагачений основними носіями (дірками) антизапірний шар. Коли для контактуючих металу та напівпровідника *p*-типу виконується нерівність $\Phi_M < \Phi_S$, то відбувається вигин зон **униз**, число дірок у приконтактній області, порівняно з об'ємом зразку, зменшується та виникає запірний шар. За значного збагачення приконтактної області неосновними носіями заряду може виникнути інверсний шар що спричинить виникнення фізичного *p-n* переходу.

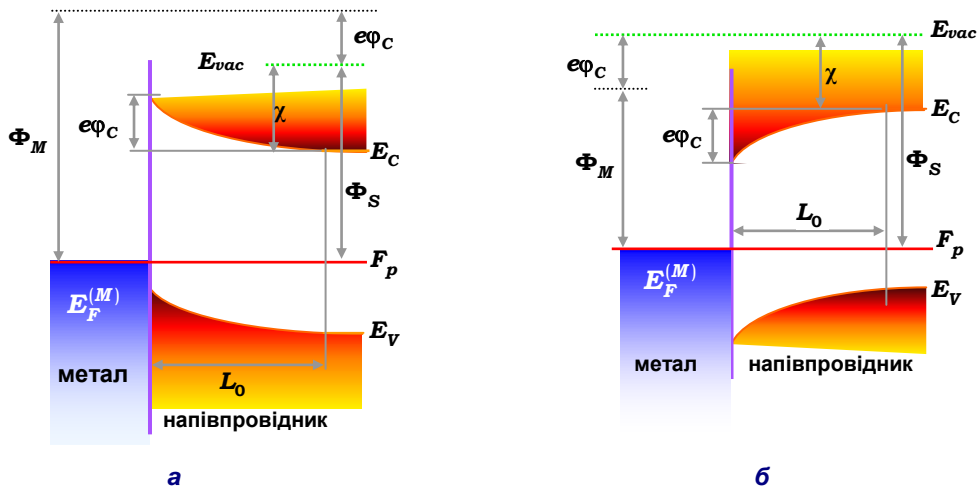


Рис. 14.5. Контакт метал–напівпровідник *p*-типу:
 а – $\Phi_M > \Phi_S$ (антизапірний шар); б – $\Phi_M < \Phi_S$ (запірний шар)

■ 14.3. Розподіл концентрації електронів на контакті метал-напівпровідник

Визначимо тепер структуру приконтактної області (довжину шару у напрямку, нормальному до площини контакту), на якій відбувається вигин зон, або глибину області напівпровідника, куди проникає контактне поле.

Розглянемо контакт між металом та напівпровідником *n*-типу. Нехай напівпровідник містить донори із концентрацією N_D та акцептори з концентрацією N_A . При цьому $N_D > N_A$. Тоді щільність об'ємного заряду **запишемо** у вигляді

$$\rho = e(N_D^+ - N_A^- - n). \quad (14.32)$$

Зрозуміло, що у стаціонарному стані концентрації іонізованих домішок (N_D^+ та N_A^-) є функціями тільки координат. В такому разі із рівняння неперервності

$$\frac{\partial \rho}{\partial t} = -\frac{\partial j}{\partial x} \quad (14.33)$$

отримуємо

$$j = \text{const}. \quad (14.34)$$

Користуючись співвідношенням Ейнштейна, запишемо густину струму, що протікає через контакт

$$j = en\mu E + \mu kT \frac{dn}{dx}. \quad (14.35)$$

Якщо домішки напівпровідника є мілкими, то можна вважати, що всі донори та акцептори іонізовані. Тоді для визначення параметрів області об'ємного заряду до (14.35) необхідно додати рівняння Пуассона

$$\frac{d^2\varphi}{dx^2} - \frac{4\pi e}{\varepsilon}(n - n_0) = 0, \quad (14.36)$$

де n_0 означає сталу концентрацію електронів в об'ємі напівпровідника. Вважатимемо, що площина контакту розташована на перерізі $x = 0$. Позначимо потенціал на контакті через φ_C . Тоді граничні умови набудуть вигляду

$$\varphi = \begin{cases} \varphi_C, & x = 0 \\ 0, & x = \infty \end{cases}, \quad n = \begin{cases} n_C, & x = 0 \\ n_0, & x = \infty \end{cases}. \quad (14.37)$$

Із рівняння (14.35) знайдемо концентрацію носіїв на контакті за відсутності струму через контакт

$$\frac{1}{kT} \frac{d\varphi}{dx} = \frac{1}{n} \frac{dn}{dx}, \quad (14.38)$$

яке можна проінтегрувати в квадратурах

$$\frac{1}{kT} \int_{\varphi_C}^0 e d\varphi = \int_{n_C}^{n_0} \frac{dn}{n},$$

звідки

$$n_C = n_0 e^{e\varphi_C/kT}. \quad (14.39)$$

Із цієї формули, зокрема впливає: якщо у напівпровіднику n -типу зони вигнуті вниз, тобто потенціал φ_C є від'ємним і збільшується до 0 при переході в об'єм зразка, концентрація електронів на межі розподілу більша за подібну в об'ємі та експоненційно зменшується за мірою віддалення від межі розподілу. Якщо ж зони вигнуті вгору, то концентрація електронів на межі розподілу менша за її об'ємне значення та експоненційно збільшується за віддалення від границі.

У випадку, коли через контакт протікає струм, концентрація носіїв на контакті – $n(x=0)$ – взагалі не має дорівнювати рівноважній концентрації n_C . Її необхідно знаходити окремо. Але за невеликих значень струму величина $n(x=0)$ не дуже відрізняється від n_C і її можна вважати незалежною від струму. Для оцінювання величини струмів, які допускають таке наближення, необхідно визначити сумарний струм, що визначається потоком частинок із напівпровідника до ме-

талу, і навпаки – із металу до напівпровідника. Електрони, що рухаються із напівпровідника до металу та дійшли до контакту (ті, що подолали бар'єр контактної різниці потенціалів), вже не зустрічають завад і безперешкодно проникають у метал. Без зовнішньої різниці потенціалів кількість електронів, які проходять через одиничний переріз за одиницю часу, пропорційна добутку $n_C v$. З іншого боку, із металу до напівпровідника проходить кількість електронів, пропорційна $n(0)v$. Тоді сумарний струм через контакт є пропорційним різниці між граничними концентраціями електронів на контакті та їхній середній швидкості у напрямку, нормальному до площини контакту

$$j = \text{const} \cdot ev(n(0) - n_C). \quad (14.40)$$

Зрозуміло, що чим меншою є різниця $(n(0) - n_C)$, тим меншим має бути струм. Оскільки для концентрацій $n_C \sim 10^{13} - 10^{14} \text{ см}^{-3}$ і швидкостей руху електронів $\sim 10^7 \text{ см/с}$ другий доданок в (14.40) дає значення $\sim 10 \text{ А/см}^2$, то для струмів через бар'єр $\sim 0.1 \text{ А/см}^2$ відносна різниця між $n(x=0)$ та n_C не перевищує 0.01. Таким чином, із точністю до одного відсотка наближення $n(0) \approx n_C$ буде припустимим.

За відсутності струму залежність концентрації носіїв у приконтактній області від величини різниці потенціалів, як впливає із (14.35), можна визначити, виконуючи інтегрування рівняння (14.38)

$$\frac{1}{kT} \int_{\varphi(x)}^0 e d\varphi = \int_{n(x)}^{n_0} \frac{dn}{n}.$$

Звідси отримуємо співвідношення

$$n(x) = n_0 e^{e\varphi(x)/kT}. \quad (14.41)$$

Підставляючи цей вираз до рівняння Пуассона (14.36), отримуємо

$$\frac{d^2\varphi}{dx^2} = \frac{4\pi e}{\varepsilon} n_0 (e^{e\varphi(x)/kT} - 1). \quad (14.42)$$

Розв'язки цього рівняння за різних припущень є головною метою цього розгляду.

Розглянемо спочатку випадок слабкого викривлення зон, що відповідає нерівності $e\varphi \ll kT$. Розкладення експоненти у (14.42) у ряд дає

$$\frac{d^2\varphi}{dx^2} = \frac{4\pi e^2}{\varepsilon kT} n_0 \varphi. \quad (14.43)$$

Розв'язок цього рівняння із врахуванням граничних умов (14.37) має вигляд

$$\varphi(x) = \varphi_C e^{-x/L_S}, \quad (14.44)$$

де $L_S = \sqrt{(\varepsilon kT)/(4\pi e^2 n_0)}$ – довжина екранування. Бачимо, що потенціал експоненційно зменшується за мірою віддалення від контакту та за великих x , тобто в об'ємі зразку досягає нуля. Дебаєвська довжина екранування L_S у напівпровідниках має порядок $10^{-5} - 10^{-4}$ см. За зменшення концентрації носіїв довжина екранування збільшується. Довжину екранування не можна зробити як завгодно великою, зменшуючи концентрацію носіїв, тому концентрація неконтрольованих домішок не може бути меншою за $10^{10} - 10^{11}$ см $^{-3}$, що зумовлено технологічними особливостями вирощування напівпровідникових зразків. Це означає, що довжина екранування у напівпровідниках не може перевищувати величину $\sim (1 \div 10) \cdot 10^{-4}$ см.

14.3.1. Збагачений (антисепічний) контакт за відсутності струму

Розглянемо розподіл електричного потенціалу та щільності електронів для збагаченого контакту, тобто для випадку, коли у приконтактному шарі $e\varphi_C < 0$ і виконується нерівність $|e\varphi_C| \gg kT$. Тоді в області контакту, що безпосередньо примикає до межі розподілу, можна записати

$$1 \ll e^{e\varphi/kT}, \quad (14.45)$$

тим самим спростивши (14.42)

$$\frac{d^2\varphi}{dx^2} = \frac{4\pi e}{\varepsilon} n_0 e^{e\varphi/kT}. \quad (14.46)$$

Для розв'язку цього рівняння домножимо обидві його частини на $d\varphi/dx = \zeta$ і візьмемо до уваги, що

$$\frac{d^2\varphi}{dx^2} \frac{d\varphi}{dx} = \zeta \frac{d\zeta}{dx} = \frac{1}{2} \frac{d\zeta^2}{dx}.$$

Тоді рівняння (14.46) можна привести до вигляду

$$d\zeta^2 = \frac{8\pi e}{\varepsilon} n_0 e^{e\varphi/kT} d\varphi.$$

Пряме інтегрування дає

$$\zeta^2 = \frac{8\pi kT}{\varepsilon} n_0 e^{e\varphi/kT} + C. \quad (14.47)$$

Для визначення сталої інтегрування зазначимо, що далеко від контакту в об'ємі напівпровідника $\varphi = 0$ і $\zeta = 0$. Тоді

$$C = -\frac{8\pi kT}{\varepsilon} n_0. \quad (14.48)$$

Однак, у силу нерівності (14.45), сталою C у розв'язку (14.47) можна знехтувати порівняно із першим доданком. Таким чином, отримуємо

$$\frac{d\varphi}{dx} = \pm \sqrt{\frac{8\pi kT n_0}{\varepsilon}} \cdot e^{e\varphi/2kT}. \quad (14.49)$$

Для визначення знаку згадаємо, що в електронному напівпровіднику потенціал зменшується за переходу до об'єму зразку, тобто у розв'язку (14.49) потрібно вибрати знак мінус. Інтегруючи рівняння за x у межах від 0 до x , знаходимо що

$$\exp\left(-\frac{e\varphi}{2kT}\right) - \exp\left(-\frac{\varphi_C}{2kT}\right) = \sqrt{\frac{4\pi e^2 n_0}{2\varepsilon kT}} x. \quad (14.50)$$

Це рівняння зручно переписати у вигляді

$$\exp\left(-\frac{e\varphi}{2kT}\right) = \exp\left(-\frac{\varphi_C}{2kT}\right) \left\{ 1 + \sqrt{\frac{4\pi e^2 n_0 \exp(\varphi_C/kT)}{2\varepsilon kT}} x \right\}. \quad (14.51)$$

Звідси, беручи до уваги, що $n_0 \exp(\varphi_C/kT) = n_C$, отримуємо: у безпосередньому околі контакту потенціал розподілено як

$$\varphi(x) = \varphi_C - \frac{2kT}{e} \ln\left(1 + \frac{x}{\sqrt{2} \cdot l}\right), \quad (14.52)$$

де $l = \sqrt{(\varepsilon kT)/(4\pi e^2 n_C)}$. Ця характеристична довжина є ні чим іншим, як дебаєвською довжиною екранування, в якій концентрація електронів у глибині зразка замінена на її значення на контакті. Таким чином, ми отримали, що в приконтактній області потенціал змінюється за логарифмічним законом. Підставляючи до (14.41) величину потенціалу (14.52) і беручи до уваги, що

$$\exp\left\{\frac{e\varphi}{kT}\right\} = \exp\left\{\frac{e\varphi_C}{kT}\right\} \exp\left\{\ln\left(\frac{\sqrt{2}l}{\sqrt{2}l+x}\right)\right\},$$

отримаємо розподіл концентрації електронів

$$n(x) = n_C \left(\frac{\sqrt{2} \cdot l}{x + \sqrt{2} \cdot l}\right)^2. \quad (14.53)$$

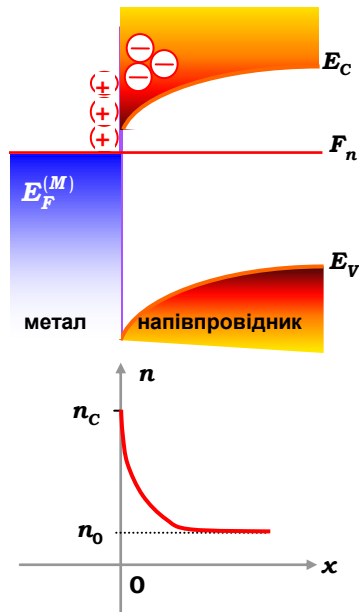


Рис. 14.6. Залежність концентрації електронів у напівпровіднику *n*-типу, що контактує з металом

Далеко від контакту маємо, що потенціал $\varphi = \varphi_N$, а концентрація електронів

$$n = n_C \exp(e\varphi_C / kT) = n_0.$$

Таким чином, в області напівпровідника довжиною $\sim \sqrt{2} \cdot l$, що безпосередньо межує з металом заливаються основні носії заряду, концентрація яких, згідно із (14.66), не залежить від їхньої концентрації у зразку (рис. 14.6). Іншими словами, навіть для ізоляторів концентрація n може бути помітною. Можна передбачати, що електропровідність тонкоплівкових структур **метал-тонкий шар діелектрика-метал** може бути досить великою.

14.3.2. Запірний (вичерпаний) контакт за відсутності струму

Розглянемо тепер випадок, коли приконтактний шар у напівпровіднику збіднений основними носіями, тобто так званий випадок запірного або блокуючого контакту. Обмежимося спеціальним випадком дуже сильного збіднення, який не тільки досить легко аналітично обчислити, але й часто зустрічається у практиці. Нехай в електронному напівпровіднику донорна домішка повністю іонізована, тобто $n_0 = N_D^+$. Оскільки у приконтактній області енергія електронів на дні зони провідності визначається виразом $E_C + e\varphi(x)$, об'ємний заряд у ній

$$\rho = en_0(1 - e^{-e\varphi(x)/kT}). \tag{14.54}$$

Як зазначалося вище, контактна різниця потенціалів повністю приходить на приконтактну область у напівпровіднику. Тоді можна вважати, що $|e\varphi| \gg kT$, тобто об'ємний заряд визначатиметься виразом

$$\rho = en_0. \tag{14.55}$$

Це означає, що із приконтактної області напівпровідника товщиною L_0 вільні електрони витіснені полем, і в цій області зосереджено позитивний заряд іонізованих донорів. Із врахуванням цього в області об'ємного заряду рівняння Пуассона запишемо у вигляді

$$\frac{d^2\varphi}{dx^2} = \frac{4\pi e}{\varepsilon} n_0. \quad (14.56)$$

Розв'язком цього рівняння може бути функція

$$\varphi(x) = \frac{2\pi e n_0}{\varepsilon} (L_0 - x)^2 + C_1(L_0 - x) + C_2. \quad (14.57)$$

Оскільки ми вважаємо, що поле проникає всередину напівпровідника на відстань L_0 , то до (14.57) треба додати граничні умови

$$\varphi(L_0) = 0 \text{ та } \left. \frac{d\varphi}{dx} \right|_{x=L_0} = 0, \quad (14.58)$$

використання яких дозволяє визначити коефіцієнти $C_1 = 0$ і $C_2 = 0$. Таким чином, у приконтактній запірній області електронного напівпровідника електричний потенціал залежить від координати **як**

$$\varphi(x) = \frac{2\pi e n_0}{\varepsilon} (L_0 - x)^2. \quad (14.59)$$

Оскільки контактна різниця потенціалів на контакті визначається різницею термоелектронних робіт виходу (див. (14.29)), можна отримати глибину проникнення контактного поля (т. зв. шар Шотткі)

$$L_0 = \sqrt{\frac{\varepsilon}{4\pi e n_0}} \varphi_C. \quad (14.60)$$

Беручи до уваги, що $\varphi_C = e(\Phi_M - \Phi_S)$, із (14.60) отримуємо

$$\frac{L_0}{L_S} = \sqrt{\frac{2(\Phi_M - \Phi_S)}{kT}}. \quad (14.61)$$

Звідси бачимо, що, оскільки різниця робіт виходу метала та напівпровідника приблизно становить величину 1 еВ, то за кімнатної температури (де **$kT \approx 0,026$ еВ**) глибина проникнення поля може перевищувати довжину екранування майже у десять разів.

Запірний приконтактний шар збіднений основними носіями заряду, завдяки чому має більший опір, ніж об'єм напівпровідника. **Це означає, що контакт метал-напівпровідник** за сутністю становить специфічну структуру, що включає шар великого опору в **обкладках**, що є хорошими провідниками – металу та напівпровідника із великою провідністю. Таку структуру можна розглядати як конденсатор. Тому запі-

рний шар можна характеризувати бар'єрною ємністю, яку можна визначити згідно із формулою $C = \varepsilon / (4\pi L_0)$, або із врахуванням (14.60)

$$C = \sqrt{\frac{\varepsilon \varepsilon n_0}{8\pi \varphi_C}}. \quad (14.62)$$

Аналогічна формула визначає бар'єрну ємність й у випадку контакту **метал-напівпровідник** p -типу.

■ 14.4. Випрямлення струму у контакті метал-напівпровідник

Розглянемо тепер, яким чином може формуватись специфічна вольт-амперна характеристика запірних (блокувальних) контактів. Вплив потенціального бар'єру на проходження електричного струму через контакт **метал-напівпровідник** істотно залежить від співвідношення між шириною бар'єру та де-бройлівською довжиною хвилі електронів. Дійсно, якщо товщина бар'єру є порівняною із довжиною хвилі де-Бройля електронів, то завдяки квантовому тунелюванню електрони проходять бар'єр. Якщо ж довжина хвилі де-Бройля істотно менша за товщину бар'єра, то ефектами тунелювання можна знехтувати, і можливість проходження бар'єрної області визначатиметься класичною умовою перевищення енергією електронів висоти бар'єра. Оцінки показують, що за кімнатної температури та концентрації електронів $\sim 10^{15} \text{ см}^{-3}$ товщина бар'єру становить $\sim 10^{-5} \text{ см}$. Характерна величина довжини хвилі де-Бройля зонного електрона за кімнатної температури $\lambda_B \sim 10^{-6} \text{ см}$. Отже виконуються умови класичного проходження бар'єрної області. Необхідно, однак, мати на увазі, що збільшення концентрації електронів спричинить зменшення товщини потенціального бар'єру $\sim (n_0)^{-1/2}$ (див (14.60)) і за концентрацій електронів $\sim 10^{18} \text{ см}^{-3}$ виконуватиметься нерівність $L_0 < \lambda_B$, і рух електронів через бар'єр визначатиметься квантовим тунелюванням.

Із квантової теорії відомо, що прозорість потенціального бар'єру D (відношення числа електронів, що долають бар'єр до повного числа електронів, що падають на бар'єр) визначається товщиною бар'єра δ та різницею між його висотою Φ_0 і кінетичною енергією електрона, що набігає на бар'єр $E_{\perp} = m^* v_{\perp}^2 / 2$, де v_{\perp} – проекція швидкості електрона перпендикулярна до поверхні: $D = \exp\left[-(2\delta/\hbar)\sqrt{2m^*(\Phi_0 - E_{\perp})}\right]$. Для

електронів, що рухаються із металу до напівпровідника висотою бар'єру буде Φ_M , а для електронів, що рухаються із напівпровідника до металу – Φ_S . Характерні величини $\Phi_0 \sim 1 \text{ eV}$ і $\delta \sim 10^{-8} \text{ см}$, і для $E_{\perp} \ll \Phi_0$ прозорість бар'єру становить $\sim 1/3$. Можна стверджувати, що тонкі контакти є практично прозорими. Але навіть у випадку тонких контактів не всі електрони, що рухаються із металу, проникатимуть у напівпровідник. Наприклад, електрони, що характеризуються енергією, яка попадає у заборонену зону, відіб'ються від напівпровідника. Електрони, що мають енергію, яка попадає у валентну зону, також не можуть попадати у стани, що вже заповнені, якщо тільки там немає незаповнених станів, тобто дірок. Тільки електрони, енергії яких відповідають зоні провідності, можуть проникнути у напівпровідник. Для того, щоб потрапити у зону провідності, електрону метала необхідно подолати потенціальний бар'єр висотою $\Phi_M - \Phi_S$, який визначається тільки роботами виходу q не залежить від прикладеної напруги.

Розглянемо проходження електронів через бар'єрну область контакту **метал–електронний напівпровідник** у випадку не дуже сильного легування. За відсутності зовнішнього електричного поля енергія електронів напівпровідника на дні зони провідності менша за таку на межі розподілу з металом на величину потенціального бар'єру $e\phi_C$. Іншими словами, для електрона, що рухається із глибини напівпровідника до металу існує потенціальний бар'єр $e\phi_C$. У рівноважному стані потік електронів із напівпровідника до металу врівноважується таким самим потоком із металу до напівпровідника, що спричиняє вирівнювання рівнів Фермі по обидва боки контакту (рис. 14.7 а). Якщо у системі утворено зовнішню різницю потенціалів так, що енергію електронів у напівпровіднику збільшено на величину $eV > 0$, для чого напівпровідник n -типу має бути з'єднаний із негативним полюсом джерела живлення (для напівпровідника p -типу знак напруги має бути плюсовим). Тоді потенціальний бар'єр для електронів напівпровідника зменшується на величину eV , і потік електронів із напівпровідника до металу (тобто струм з металу в напівпровідник j_{M-S}) збільшується (рис. 14.7 б). У металі ж ні бар'єр, ні потік електронів у бік напівпровідника не змінюються. Таким чином, у системі виникає ненульовий струм із металу до напівпровідника $j = j_{M-S} - j_{S-M}$, який швидко збільшується за збільшення різниці потенціалів V .

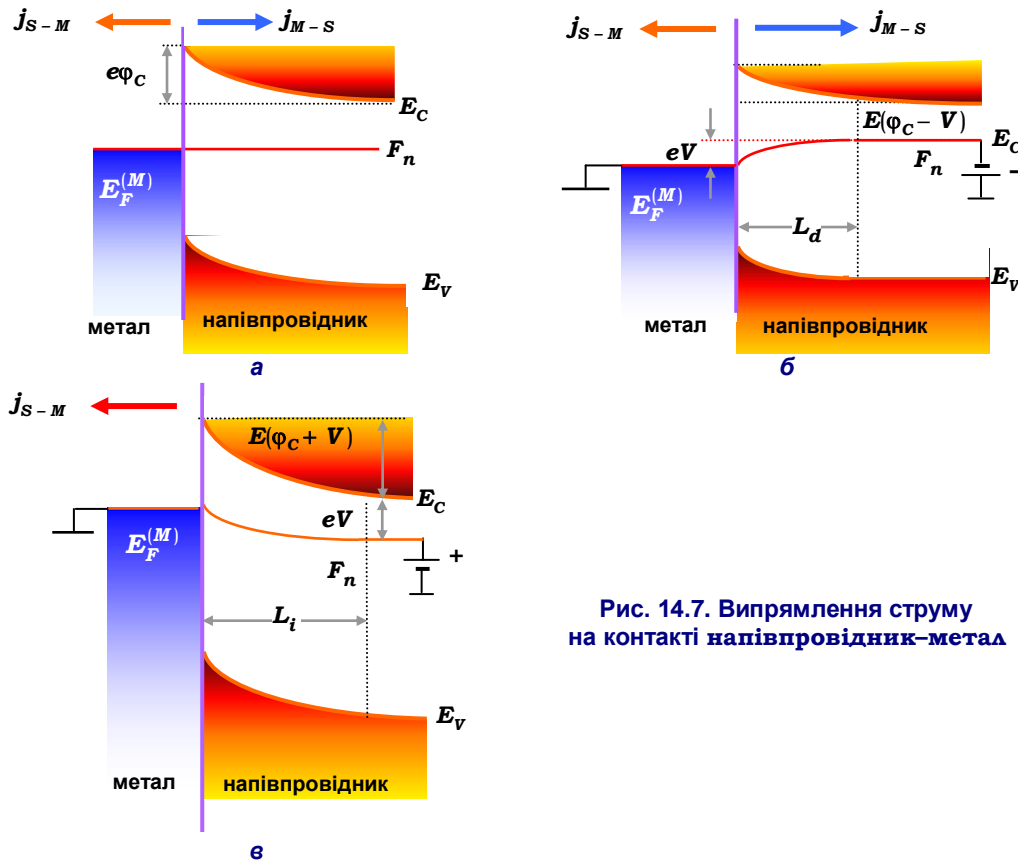


Рис. 14.7. Випрямлення струму на контакті напівпровідник-метал

На вольт-амперній характеристиці переходу спостерігатиметься швидке збільшення струму (рис. 14.8). Якщо до контакту прикласти різницю потенціалів іншого знаку (тобто плюс для напівпровідника *n*-типу), потенціальний бар'єр у напівпровіднику збільшиться, і потік електронів із напівпровідника до металу зменшиться: виникне струм із напівпровідника до металу (рис. 14.7 в). За умови $eV \sim kT$ електрони із напівпровідника практично не зможуть подолати потенціальний бар'єр, і струм із напівпровідника до металу визначатиметься тунелюванням – він досягне деякого значення насичення, яке не залежить від зовнішньої напруги.

Розрахунки вольт-амперних характеристик контакту **метал-напівпровідник** необхідно проводити із врахуванням важливості ієрархії довжин – ширини запірного шару L_0 (області просторового заряду) та довжиною вільного пробігу електронів l . Дійсно, якщо $l \gg L_0$, то внесок процесів розсіяння електронів усередині ОПЗ до

формування струмів через контакт є малим, і ним можна знехтувати. Тому струми j_{M-S} та j_{S-M} можна розглядати як струми термоелектронної емісії. При цьому треба брати до уваги, що емісія відбувається із напівпровідника до металу, і навпаки. Оскільки висота потенціального бар'єра у цьому випадку зазвичай є меншою, ніж за термоелектронної емісії у вакуум, то струми термоелектронної емісії можуть бути достатньо помітними при кімнатних температурах. Якщо виконується зворотна нерівність $l \ll L_0$ за час проходження через товщину бар'єра, електрони встигають багато разів розсіятись. Тоді струм через контакт необхідно визначати із врахуванням і дифузійного та дрейфового руху.

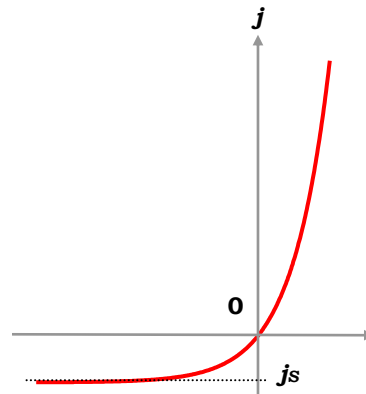


Рис. 14.8. Вольт-амперна характеристика контакту метал-напівпровідник

14.4.1. Діодна теорія випрямлення струму

Як зазначалось раніше, при виконанні нерівності $l \gg L_0$ можна вважати, що електрони пролітають запірну область контакту майже без зіткнень. У цьому випадку запірний шар об'ємного заряду напівпровідника схожий на вакуумний проміжок між електродами вакуумної лампи. У зв'язку з цим теорія, що використовується для обчислення вольт-амперної характеристики контакту **метал-напівпровідник**, у випадку, що розглядається, називається діодною теорією. Оскільки товщина запірного шару L_0 є тим меншою, чим більшою є рівноважна концентрація електронів в об'ємі напівпровідника n_0 , а довжина вільного пробігу електронів є тим більшою, чим більшою є рухливість цих електронів, то можна стверджувати, що діодна теорія випрямлення струму контакту **метал-напівпровідник** може застосовуватись до напівпровідників, що характеризуються значною концентрацією та високою рухливістю носіїв. Як відомо із квантової теорії, прозорість потенціального бар'єра висотою Φ_0 і товщиною d визначається формулою

$$D = \exp \left\{ -\frac{d}{\pi \hbar} \sqrt{2m^* (\Phi_0 - E')} \right\}, \quad (14.63)$$

де $E' = m^* v_x^2 / 2$ – кінетична енергія руху електронів у напрямку бар'єра (за вісь ОХ системи координат береться цей напрямок). Якщо прийняти, що $\Phi_0 \sim 1$ еВ, а $d \sim 10^{-8}$ см, то для не дуже швидких електронів, для

яких $E' \ll \Phi_0$, прозорість бар'єру становить 0.3: бар'єр є майже прозорим. Тому вважатимемо, що всі електрони з енергією, більшою за Φ_0 , здатні пройти через бар'єр. Для переходу із напівпровідника до металу електрони повинні мати швидкість v_x , достатню для проходження бар'єру висотою $e(\varphi_C - V)$. Електричний струм, зумовлений рухом таких електронів, можна обчислити методом, аналогічним використаним нами при обчисленні струму термоелектронної емісії (див. (14.9)). Для отримання виразу для струму електронів із металу в напівпровідник треба замість енергії електронної спорідненості записати енергію бар'єру $e(\varphi_C - V)$. У результаті отримаємо

$$j_{M-S} = \frac{em^*(kT)^2}{2\pi^2\hbar^3} \exp\left\{\frac{F - E_C}{kT}\right\} \exp\left\{-\frac{e(\varphi_C - V)}{kT}\right\}. \quad (14.64)$$

Оскільки рівноважна концентрація електронів в об'ємі напівпровідника

$$n_0 = 2 \left(\frac{m^* kT}{2\pi\hbar} \right)^{3/2} e^{(F - E_C)/kT} \quad (14.65)$$

та середня теплова швидкість цих електронів

$$v_0 = \sqrt{\frac{8kT}{\pi m^*}}, \quad (14.66)$$

формула (14.64) може бути представлена у вигляді

$$j_{M-S} = (en_S v_0 / 4) \cdot e^{eV/kT}. \quad (14.67)$$

Тут для концентрації електронів на поверхні напівпровідника за нульової різниці потенціалів використано позначення

$$n_S = n_0 e^{-e\varphi_C/kT}. \quad (14.68)$$

Оскільки з боку металу висота потенціального бар'єру така сама, як і з боку напівпровідника, то потік електронів із металу до напівпровідника не залежить від прикладеної різниці потенціалів і дорівнює потоку електронів із напівпровідника до металу в умовах термодинамічної рівноваги. Іншими словами, враховуючи, що напрямок електричного струму є протилежним до напрямку руху електронів, $j_{S-M} = j_{M-S}(V=0)$. Таким чином, використовуючи (14.67), знайдемо

$$j_{S-M} = \frac{en_S v_0}{4}. \quad (14.69)$$

Сумарний струм через контакт дорівнюватиме різниці

$$j = j_{M-S} - j_{S-M}. \quad (14.70)$$

Підставляючи сюди вирази (14.67) та (14.69), отримаємо, що струм через контакт пов'язаний із різницею потенціалів, що прикладений до контакту, таким співвідношенням

$$j = j_S (e^{eV/kT} - 1), \quad (14.71)$$

де струм насичення
$$j_S = \frac{1}{4} en_S v_0. \quad (14.72)$$

Формула (14.71) описує залежність струму через контакт від різниці потенціалів на контактах. Наприклад, за прямого зміщення ($V > 0$) струм через контакт зростає за експоненційним законом за зростання V . За зворотного зміщення ($V < 0$) струм зменшується за зменшення V і прямує до струму насичення j_S (рис. 14.8). Тобто, контакт **метал-напівпровідник** характеризується випрямляючими властивостями та може бути використаний як випрямляч змінного струму.

14.4.2. Дифузійна теорія випрямлення струму

У випадку, коли процеси розсіювання електронів при проходженні контакту відіграють помітну роль, необхідно враховувати як дрейфову складову, так і дифузійну складові руху електронів. Для цього запишемо струм через контакт у вигляді

$$j = en\mu_n E + eD_n \frac{dn}{dx}. \quad (14.73)$$

Помножимо цей вираз на $\exp(eV/kT)$ і використаємо співвідношення Ейнштейна. Дістанемо

$$je^{eV/kT} = n\mu_n e^{eV/kT} \frac{deV}{dx} + \mu_n kT e^{eV/kT} \frac{dn}{dx}. \quad (14.74)$$

Звідси знаходимо, що

$$je^{eV/kT} = \mu_n kT \frac{d}{dx} \{ne^{eV/kT}\}. \quad (14.75)$$

Проінтегруємо це рівняння за координатою від 0 до L (товщина бар'єрної області) і визначимо струм, що не залежить від координати

$$j = \frac{\mu_n kT}{S} \{n(L)e^{eV(L)/kT} - n(0)e^{eV(0)/kT}\}, \quad (14.76)$$

де

$$S = \int_0^L dx e^{eV(x)/kT}. \quad (14.77)$$

Для спрощення ситуації вважатимемо, що електричне поле зосереджено тільки на ОПЗ, падінням напруги на квазінейтральному шарі знехтуємо, тоді можна записати граничні умови

$$V(L) = 0, \quad n(L) = n_0. \quad (14.78)$$

На межі розподілу для випадку, коли зовнішня різниця потенціалів не дуже велика, але у приконтатній області існує запірний шар (тобто $e(\varphi_N - V) \gg kT$), маємо

$$V(0) = \varphi_C - V, \quad n(0) = n_S. \quad (14.79)$$

Тоді (14.76) матиме вигляд

$$j = \frac{n_0 \mu_n kT}{S} \{1 - e^{-eV/kT}\}. \quad (14.80)$$

Далі необхідно оцінити величину S у знаменнику (14.91). Перейдемо від інтегрування за координатою до інтегрування за потенціалом

$$S = \int_{\varphi_C - V}^0 e^{eV/kT} \left(\frac{dx}{dV} \right) dV. \quad (14.81)$$

Цей інтеграл зручно оцінювати методом перевалу. Для цього необхідно пам'ятати, що максимальний внесок до інтегралу надаватиме та область, де величина $V(x)$ є максимальною, – область поблизу межі розподілу ($x = 0$). Зі збільшенням x підінтегральна функція експоненційно спадатиме та за достатньо великих x стане настільки малою, що її внеском можна буде знехтувати. Зрозуміло, що основний внесок до зміни підінтегральної функції даватиме швидкозмінна експонента, а функція $(dV/dx)^{-1}$ змінюватиметься набагато повільніше. Тому можна покласти, що

$$\left(\frac{dV}{dx} \right)^{-1} = -E_S^{-1} \quad (14.82)$$

і винести її за знак інтеграла. Як результат дістанемо

$$S = -E_S^{-1} \int_{\varphi_C - V}^0 dV e^{eV/kT} = -E_S^{-1} \cdot \frac{kT}{e} \cdot \{1 - e^{e(\varphi_C - V)/kT}\}. \quad (14.83)$$

Оскільки ми вважаємо, що виконується нерівність $e(\varphi_N - V) \gg kT$, другий доданок у фігурних дужках формули набагато перевищує одиницю, якою, таким чином, можна знехтувати і записати, що

$$S = E_S^{-1} (kT/e) e^{e(\varphi_C - V)/kT}. \quad (14.84)$$

Підставляючи цю формулу до (14.80), маємо

$$j = en_0 \mu_n E_S e^{-e\varphi_C/kT} (e^{eV/kT} - 1). \quad (14.85)$$

Оскільки $n_S = n_0 e^{-e\varphi_C/kT}$, введемо позначення

$$j_S = en_S \mu_n E_S \quad (14.86)$$

для струму за великих зворотних зміщень та отримаємо величину струму через контакт **метал-напівпровідник** як функцію прикладеної зовнішньої різниці потенціалів, обчислену у наближенні дифузійної теорії

$$j = j_S(e^{eV/kT} - 1). \quad (14.87)$$

Таким чином, дві представлені теорії випрямлення струму на контакті **метал-напівпровідник** відповідають двом можливим варіантам поведінки носіїв заряду в області просторового заряду. Перший варіант, коли довжина вільного пробігу носіїв більше ширини шару Шотткі (діодна теорія), що фізично відображає ситуацію, за якої в області просторового заряду практично відсутні зіткнення носіїв заряду. Другий – відповідає дифузійній теорії, що означає зміну енергії носія заряду при доланні потенціального бар'єру області просторового заряду за рахунок розсіювання електронів.

Наостанок проаналізуємо наслідки реалізації різних типів контакту **метал-напівпровідник** для транспортних властивостей напівпровідникових структур. Дійсно, збагачення приконтактного шару напівпровідника неосновними носіями заряду може спричинити суттєві зміни властивостей контактів. Наприклад, при прикладенні зовнішнього потенціалу виникає струм через контакт. При цьому формування струму відбувається як електронною, так і дірковою компонентами. Оскільки концентрація носіїв є різною у різних точках напівпровідника, склад струму змінюється за незмінної його сумарної величини. Розглянемо для визначеності контакт металу із напівпровідником n -типу. Нехай $\gamma = j_p(x)/j$ – частка струму, що переноситься неосновними носіями заряду, тобто дірками. У загальному випадку

$$\gamma(x) = j_p(x) [j_p(x) + j_n(x)]^{-1} = p(x) \left[p(x) + \frac{|\mu_n|}{\mu_p} n(x) \right]^{-1}. \quad (14.88)$$

$\gamma(x)$ визначає відношення струму, що переноситься неосновними носіями заряду до повного струму у нерівноважному випадку. У випадку рівноважної концентрації носіїв заряду n_0 та p_0 введемо величину

$$\gamma_0 = p_0 \left[p_0 + \frac{|\mu_n|}{\mu_p} n_0 \right]^{-1}. \quad (14.89)$$

Тоді очевидно: якщо у заданій точці x концентрація неосновних носіїв заряду зменшується, то $\gamma(x)/\gamma_0 < 1$. Якщо у заданій точці концентрація неосновних носіїв заряду при проходженні струму збільшується, то $\gamma(x)/\gamma_0 > 1$. Таким чином, при проходженні струму через контакт **метал-напівпровідник** може відбуватись зміна концентрації основних і неосновних носіїв заряду в об'ємі напівпровідника. Зрозуміло, що це

спричинить зміни властивостей напівпровідника. Отже можна стверджувати, що у випадку $\gamma(x)/\gamma_0 > 1$ об'єм напівпровідника збагачується неосновними нерівноважними носіями заряду. Це, до речі, може відбуватись як для запірного, так і для антизапірного шарів. У випадку запірного шару відбувається інжекція неосновних носіїв заряду через контакт **метал-напівпровідник**. У випадку антизапірного шару відбувається акумуляція неосновних носіїв. У випадку $\gamma(x)/\gamma_0 < 1$ об'єм збіднюється неосновними носіями заряду.

Якщо при ввімкненні зовнішнього потенціалу контакт не змінює концентрацію неосновних носіїв заряду $\gamma = \gamma_0$, то він називається омичним. Створення омичних контактів являє собою важливу технологічну задачу, яка не завжди має розв'язки. Тому для виключення впливу контактних явищ при вимірюванні властивостей напівпровідників застосовують компенсаційні методи, завдяки яким струм через контакт не протікає і, як результат, контактні явища, наприклад інжекція неосновних носіїв, не впливають на результати вимірювань.

■ 14.5. Розподіл концентрації носіїв на контакті напівпровідник-напівпровідник

Основна частина напівпровідникових приладів основана на використанні властивостей переходів між напівпровідників легованих різними типами домішок, зокрема *p-n*-переходів. Зрозуміло, що будь-які контакти між по-різному легованими напівпровідниками можна розглядати як сильно неоднорідну напівпровідникову систему. Нехай напівпровідник характеризується деяким довільним розподілом акцепторної і донорної домішок – $N_A(\mathbf{r})$ та $N_D(\mathbf{r})$, відповідно (рис. 14.10). Це

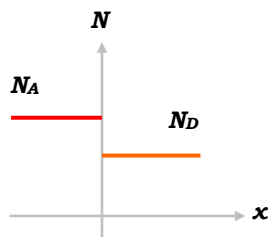


Рис. 14.10. Розподіл концентрації домішок на *p-n*-переході

означає, що концентрація носіїв є залежною від координат. Оскільки положення рівня Фермі щодо енергетичних зон визначається концентрацією носіїв, то у такій системі дістанемо залежність рівня Фермі від координати, $F = F(\mathbf{r})$, тобто стан системи є нерівноважним. Це означає виникнення потоку носіїв заряду, який намагається вирівняти концентрацію електронів і дірок у зразку. Виникнення дифузійного струму \mathbf{j}_D спричинить розділення зарядів, завдяки чому виникне об'ємний заряд, що індукуватиме електричне поле – це призведе до викривлення енергетичних зон. Дифузійний струм компенсувати-

меться дрейфовим струмом, тому у підсумку сумарний струм у напівпровіднику дорівнюватиме нулю.

$$\mathbf{j} = \mathbf{j}_D + \mathbf{j}_E = 0. \quad (14.90)$$

Іншими словами, напівпровідник перейде у стан термодинамічної рівноваги, за якого рівень Фермі не залежить від координати.

Розглянемо для визначеності явища, що виникають у контакті між напівпровідниками n - та p -типів. Такий контакт, зокрема може бути утворений, якщо до однієї частини напівпровідника вводяться донорні, а до іншої – акцепторні домішки, тобто у деякій області напівпровідника відбуватиметься зміна електропровідності від електронної на діркову. Такий перехід між матеріалами, що характеризуються p - та n -типами провідності, називається p - n -переходом. Припустимо, що такий перехід є нескінченно вузьким, та акцепторна область напівпровідника легована сильніше за донорну область, тобто $N_A > N_D$ (див. рис. 14.10). Позначимо концентрацію електронів (неосновні носії заряду) і дірок (основні носії заряду) в області $x < 0$ через n_p і p_p , відповідно. Так само в n -області концентрації основних і неосновних носіїв заряду позначимо як n_n і p_n , відповідно. Припустимо також, що домішкові рівні є мілкими настільки, що можна вважати всі домішки іонізованими. В умовах термодинамічної рівноваги в обох напівпровідниках виконується співвідношення

$$n_p p_p = n_n p_n = n_i^2. \quad (14.91)$$

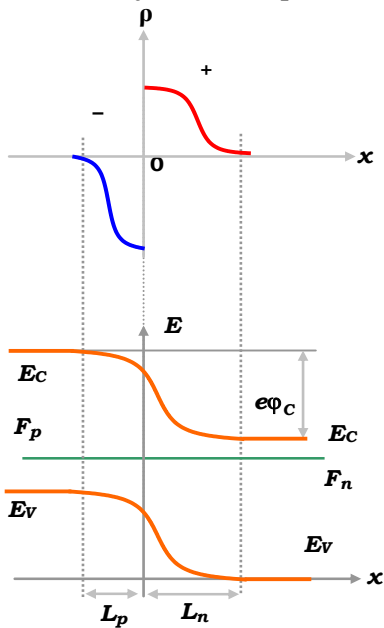


Рис. 14.11. Розподіл об'ємного заряду та зона структура p - n -переходу

Припустимо тепер, що обидві частини напівпровідника приходять в контакт. У перший момент стикання поблизу межі розподілу існуватиме великий градієнт концентрації носіїв заряду, що спричинить дифузію дірок з p - до n -області та електронів з n - до p -області. Ці дифузійні потоки заряджених частинок приведуть, як зазначалось вище, до розділення зарядів, унаслідок якого в приконтактному шарі p -області виникне негативний заряд, а в приконтактному шарі n -області – позитивний. Ці об'ємні заряди створюють сильне електричне поле, направлене з n - до p -області, яке заважатиме руху електронів і дірок. У результаті встановиться рівноважний стан зі сталим зна-

ченням рівня Фермі для всього зразка. В області переходу під дією поля виникне викривлення зон (рис. 14.11).

Разом із тим, в області контакту виникне перерозподіл концентрації електронів і дірок. Вигін зон і перерозподіл концентрації носіїв спричинить виникнення потенціального бар'єру, що заважатиме руху носіїв

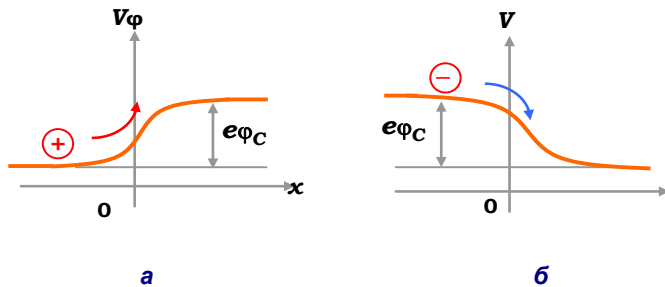


Рис. 14.12. Перехід основних (а) і неосновних (б) носіїв заряду із сильно легованої *p*-області до слабо легованої *n*-області *p-n* переходу

із однієї області до іншої. Дійсно, основні носії заряду при переході через контакт мають долати потенціальний бар'єр заввишки $e\phi_c$. Перехід неосновних носіїв відбуватиметься завдяки дії електричного поля *p-n*-переходу (рис. 14.12 та 14.13).

На рис. 14.12 подано перехід основних (що долають бар'єр) та неосновних носіїв заряду із *p*-області контакту до *n*-області. На рис. 14.13 вказано, як відбувається перехід основних (електрони) і неосновних (дірки) носіїв заряду із *n*-області контакту до *p*-області. У стані термодинамічної рівноваги щільність дифузійного струму основних носіїв заряду $J_p^{(M)}$ та $J_n^{(M)}$ урівноважується щільністю дрейфового струму неосновних носіїв заряду $J_p^{(m)}$ та $J_n^{(m)}$ так, що їхня сума (сумарний струм через *p-n*-перехід) дорівнюватиме нулю. Нехай Φ_p – термоелектронна робота виходу із напівпровідника *p*-типу, а Φ_n – термоелектронна робота виходу з електронного напівпровідника. Тоді можна записати, що потенціальний бар'єр на контакті в умовах термодинамічної рівноваги

$$e\phi_c = \Phi_p - \Phi_n = (\chi - E_C + F_p) - (\chi - E_C + F_n) = F_p - F_n. \quad (14.92)$$

Зважаючи на припущення про повну іонізацію домішок у напівпровіднику та використовуючи (14.13) та (14.15), запишемо рівняння (14.92) у вигляді

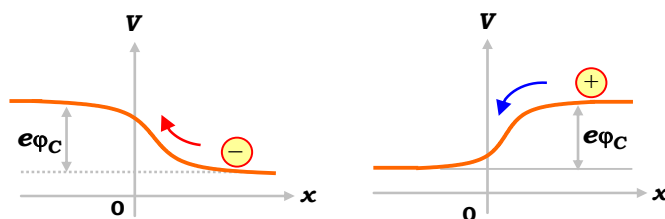


Рис. 14.13. Перехід основних (а) і неосновних (б) носіїв заряду із слабо легованої *n*-області до сильно легованої *p*-області *p-n*-переходу

$$e\varphi_C = E_g - kT \ln(N_V / N_A) - kT \ln(N_C / N_D). \quad (14.93)$$

Далі згадаємо, що $n_i^2 = N_C N_V \exp(-E_g / kT)$. Тоді отримуємо

$$e\varphi_C = kT \ln \left(\frac{n_n p_p}{n_i^2} \right), \quad (14.94)$$

Таким чином, контактна різниця потенціалів на p - n -переході тим більша, чим більше леговані n - та p -області напівпровідника. Із (14.93) видно, що максимальне значення контактної різниці потенціалів для невідродженого напівпровідника

$$\varphi_N^{(\max)} = E_g / e. \quad (14.95)$$

Потенціюючи (14.94), із врахуванням (14.91), отримаємо співвідношення

$$\frac{p_n}{p_p} = \frac{n_p}{n_n} = \exp \left(-\frac{e\varphi_C}{kT} \right). \quad (14.96)$$

Звідси легко дістаємо співвідношення між концентраціями основних і неосновних носіїв на p - n переході

$$p_n = p_p e^{-e\varphi_C / kT} \quad (14.97)$$

та

$$n_p = n_n e^{-e\varphi_C / kT}. \quad (14.98)$$

Як зазначалось раніше, чим сильніше легований напівпровідник, тим вужчий шар об'ємного заряду виникає в ньому. Щодо розглядуваної моделі це означає, що товщина шару об'ємного заряду p -області – L_p менша за товщину шару об'ємного заряду n -області – L_n . Уся область об'ємного заряду є сумою цих товщин

$$L = L_n + L_p. \quad (14.99)$$

У геометрії зразка, де n -область розташована в інтервалі $-L_p \leq x < 0$, об'ємний заряд від'ємний і визначається концентрацією іонів акцепторної домішки. Поведінка електричного потенціалу у цій області визначається рівнянням Пуассона

$$\frac{d^2\varphi}{dx^2} = 4\pi \frac{ep_p}{\varepsilon} \quad (14.100)$$

із граничними умовами

$$\varphi(-L_p) = 0, \quad \left. \frac{d\varphi}{dx} \right|_{x=-L_p} = 0. \quad (14.101)$$

В області $0 < x \leq L_n$ об'ємний заряд позитивний і визначається концентрацією донорної домішки. При цьому поведінка електричного потенціалу в цій області визначається рівнянням Пуассона

$$\frac{d^2\varphi}{dx^2} = -4\pi \frac{en_n}{\varepsilon} \quad (14.102)$$

із граничними умовами

$$\varphi(L_n) = \varphi_c, \quad \left. \frac{d\varphi}{dx} \right|_{x=L_n} = 0. \quad (14.103)$$

Розв'язуючи (14.100) та (14.102) із граничними умовами (14.101) та (14.103), отримуємо, що потенціал в p -області переходу ($-L_p \leq x < 0$)

$$\varphi_p = 2\pi \frac{ep_p}{\varepsilon} (x + L_p)^2. \quad (14.104)$$

В n -області ($0 < x \leq L_n$) потенціал

$$\varphi_n = \varphi_c - 2\pi \frac{en_n}{\varepsilon} (L_n - x)^2. \quad (14.105)$$

Візьмемо до уваги, що на границі областей при $x = 0$ потенціал і поле мають бути неперервними, тобто

$$\varphi_p(0) = \varphi_n(0) \quad \text{та} \quad \left. \frac{d\varphi_p}{dx} \right|_{x=0} = \left. \frac{d\varphi_n}{dx} \right|_{x=0}. \quad (14.106)$$

Звідси, використовуючи (14.104) та (14.105), отримуємо співвідношення

$$n_n L_n = p_p L_p, \quad (14.107)$$

що віддзеркалює принцип електронейтральності: в обох областях напівпровідника, які утворюють p - n -перехід, об'ємні заряди дорівнюють один одному. Підставляючи до першого рівняння в (14.106) вирази для φ_p та φ_n із (14.104) і (14.105), дістанемо

$$\varphi_c = 2\pi \frac{e}{\varepsilon} (p_p L_p^2 + n_n L_n^2). \quad (14.108)$$

Додаючи до обох частин (14.107) величини $p_p L_n$ та $n_n L_p$ і використовуючи вираз (14.99), отримаємо, що

$$L_p^2 = \frac{n_n^2}{(n_n + p_p)^2} L^2 \quad \text{та} \quad L_n^2 = \frac{p_p^2}{(n_n + p_p)^2} L^2. \quad (14.109)$$

Підставляючи ці співвідношення до (14.108), отримаємо зв'язок між контактною різницею потенціалів і товщиною області об'ємного заряду

$$\varphi_c = 2\pi \frac{e}{\varepsilon} \frac{n_n p_p}{(n_n + p_p)} L^2. \quad (14.110)$$

Звідси отримаємо вираз для товщини об'ємного заряду p - n -переходу

$$L = \sqrt{\frac{\varepsilon}{2\pi e} \cdot \frac{n_n + p_p}{n_n p_p} \cdot \varphi_c}. \quad (14.111)$$

Таким чином, ми отримали важливий результат: чим більшим є ступінь легування областей напівпровідника, що утворюють p - n -перехід, тим меншою є товщина шару об'ємного заряду. Крім того, як випливає із (14.108), якщо одна із областей переходу легована значно сильніше за іншу, то більша частина падіння електростатичного потенціалу приходить на високоомну область. Ще одним важливим висновком із (14.111) є те, що вона показує, що товщина p - n -переходу залежить від прикладеної напруги. Дійсно за ненульової зовнішньої напруги V_e у (14.111) необхідно здійснити заміну $\varphi_C \rightarrow \varphi_C + V_e$ (цей факт буде детально обговорено у наступному підрозділі). В області p - n -переходу значно зменшується концентрація носіїв заряду, що спричиняє суттєве підвищення електричного опору, тобто електронно-дірковий перехід являє собою шар низької питомої провідності, розташований між областями високої провідності. Тому таку структуру можна розглядати як плоский конденсатор, що характеризується деякою ємністю. Така ємність називається бар'єрною ємністю, її можна обчислити за формулою $C = \varepsilon / 2\pi L$. Із врахуванням (14.111) отримуємо, що бар'єрна ємність

$$C = \sqrt{\frac{e\varepsilon}{4\pi\varphi_C} \cdot \frac{n_n p_p}{n_n + p_p}}. \quad (14.112)$$

■ 14.6. Випрямлення струму на p - n -переході

Розглянемо тепер p - n -перехід, до контактів якого прикладено різницю потенціалів V . При цьому вважатимемо, що різниця потенціалів є такою, що p -область переходу заряджається позитивно. У такому випадку кажуть, що p - n -перехід відбувається в умовах прямого зміщення. Оскільки область об'ємного заряду характеризується високим електричним опором, падіння напруги буде в основному відбуватись в цій області. Таким чином, за прямого зміщення висота потенціального бар'єру знижується на величину eV порівняно із рівноважним станом, за якого $V = 0$. Зрозуміло, що відповідно зміниться й товщина запірного шару: тепер замість φ_C у (14.111) необхідно записати $\varphi_C - V$, тобто

$$L_d = \sqrt{\frac{\varepsilon}{2\pi e} \cdot \frac{n_n + p_p}{n_n p_p} \cdot (\varphi_C - V)}. \quad (14.113)$$

Нехай спочатку до зразка не прикладено зовнішньої електричної напруги. Тобто, p - n -перехід перебуває у рівноважному стані, в якому енергетична діаграма p - n -переходу має вигляд, що подано на рис. 14.14.

При цьому ширина переходу виражається формулою (14.111). Прикладення різниці потенціалів V , що утворює пряме зміщення (рис. 14.15), спричинить зниження потенціального бар'єру та збільшення потоку

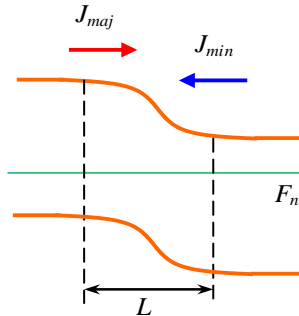


Рис.14.14

Енергетична діаграма p - n -переходу в умовах термодинамічної рівноваги

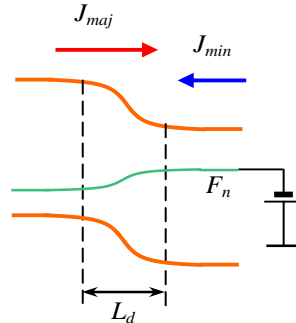


Рис. 14.15

Енергетична діаграма p - n -переходу в умовах прикладення прямого зміщення

носіїв заряду, оскільки більша кількість електронів із n -області переходить до p -області, а із p -області – переходить більша кількість дірок до n -області. Іншими словами, потік основних носіїв різко збільшиться, а потік неосновних носіїв через перехід залишиться практично сталим. У результаті у зовнішньому колі протікатиме електричний струм, який називають прямим. Він визначається різницею струмів основних і неосновних носіїв зарядів і в даному випадку буде направлений від p - до n -області. Таким чином, можна стверджувати, що в n -області під дією зміщення з'являться надлишкові неосновні носії (дірки) із концентрацією Δp . Вони створять поблизу контакту позитивний об'ємний заряд, який дуже швидко (за час максвелловської релаксації) буде скомпенсований об'ємним зарядом основних носіїв заряду, які під дією електричного поля, індукованого дірками, підтягнуться із об'єму зразка. До n -області зразка, у свою чергу, електрони надійдуть від зовнішнього джерела. Таким чином у всіх областях напівпровідника зберігатиметься електронейтральність. Але у приконтактній області p - n -переходу концентрація носіїв буде підвищена: основних носіїв – на величину Δn , неосновних – на величину Δp так, що спостерігатиметься рівність $\Delta n = \Delta p$. Таке впорскування у напівпровідник із p - n -переходом носіїв заряду за подачі на нього прямого зміщення називається інжекцією. Підкреслимо, що за подачі на p - n -перехід прямого зміщення відбуватиметься інжекція неосновних носіїв заряду. Система перейде у нерівноважний стан, що характеризуватиметься концентрацією носіїв поблизу контакту $p = p_n + \Delta p$, яку у стаціонарному випадку можна визначити на перерізі $x = L_n$.

$$p(L_n) = p_p e^{-e(\phi_c - V)/kT}. \quad (14.114)$$

Звідси, беручи до уваги (14.97), отримаємо

$$p(L_n) = p_n e^{eV/kT}. \quad (14.115)$$

Концентрація надлишкових дірок визначається як $\Delta p = p - p_n$ або

$$\Delta p(L_n) = p_n (e^{eV/kT} - 1). \quad (14.116)$$

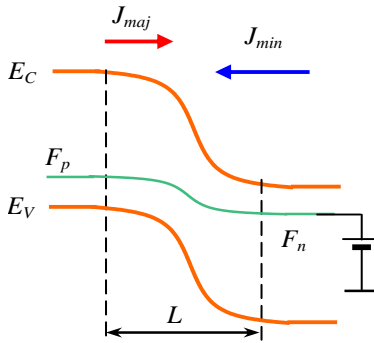


Рис. 14.16
Енергетична діаграма *p-n*-переходу
в умовах зворотного зміщення

Таким чином, бачимо, що за збільшення прямого зміщення на *p-n*-переході концентрація інжектованих неосновних носіїв заряду експоненційно зростає, що спричиняє різке зростання струму через контакт.

Розглянемо тепер випадок, коли до *p-n*-переходу прикладене зворотне зміщення. У цьому випадку *p*-область контакту заряджатиметься негативно, і потенціальний бар'єр підвищиться на eV (рис. 14.16). Разом із цим збільшиться товщина запірного шару

$$L = \sqrt{\frac{\varepsilon}{2\pi e} \cdot \frac{n_n + p_p}{n_n p_p} \cdot (\phi_c + V)}. \quad (14.117)$$

Чим більше зворотне зміщення, тим вищим буде потенціальний бар'єр, який зможе подолати все менша кількість основних носіїв заряду. Відповідно до цього кількість неосновних носіїв заряду в приконтактній області зменшуватиметься порівняно з їх рівноважною концентрацією. Згідно з умовою електронейтральності зменшуватиметься й концентрація основних носіїв. Концентрація неосновних носіїв заряду в *p*-області на перерізі $x = -L_p$ визначатиметься формулою, аналогічною до (14.116)

$$\Delta n(-L_p) = n_p (e^{-eV/kT} - 1). \quad (14.118)$$

Іншими словами, за зворотного зміщенні концентрація основних носіїв заряду зменшуватиметься, а неосновних – лишатиметься майже незмінною. Тому за зворотного зміщення струм через *p-n*-перехід буде направлений від *n*- до *p*-області. Збільшення зміщення спричинить незначне збільшення струму і подальше насичення. Подальше збільшення зворотного зміщення викличе різке збільшення струму завдяки електричному пробою на контакті. Оскільки струм через *p-n*-перехід

визначається концентрацією носіїв, то згідно із (14.117) та (14.118), можна записати

$$j = j_s(e^{eV/kT} - 1), \quad (14.119)$$

де j_s – струм насичення.

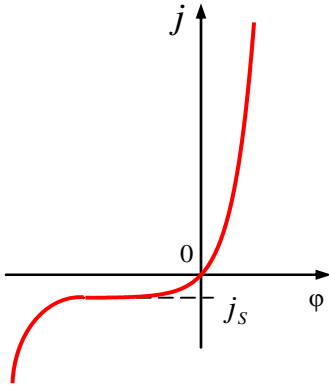


Рис.14.17
Вольт-амперна характеристика **p-n-переходу**

Якісний вигляд вольт-амперної характеристики **p-n-переходу** подано на рис.14.17. Для детальнішого аналізу цієї характеристики припустимо що утворюється контакт між двома напівпровідниками різного типу провідності **p**- та **n**-. Для спрощення припустимо, що межа розподілу між ними є різкою, тобто при $x > 0$ – розташована область напівпровідника **n**-типу, а при $x < 0$ – напівпровідника **p**-типу. У системі встановився рівень Фермі, однаковий для обох її частин. Хід потенціалу

в області межі розподілу та його залежність від координати в наближенні сталої концентрації носіїв заряду в ОПЗ знайдено раніше ((14.104) і (14.105)). Залежність довжини L від зовнішньої різниці потенціалів можна знайти, оскільки зовнішня різниця потенціалів, як і у випадку бар'єру Шоттки або зменшує, або збільшує висоту бар'єру на межі поділу. Використовуючи (14.117) для випадку $p_p \gg n_n$, знайдемо

$$L = \sqrt{\frac{\varepsilon}{2\pi e n_n}} (\varphi_c \pm V). \quad (14.120)$$

Ясно, що змінюючи прикладену зовнішню напругу, ми змінюємо величину області просторового заряду, яка утворює плоский конденсатор. Таким чином, зміна V спричиняє зміну заряду цього конденсатора Q . Величина $C = dQ/dV$ – ємністю, зумовлена просторовим (фіксованим) зарядом на іонізованих домішках. Вона називається бар'єрною, і за аналогією з ємністю плоского конденсатора її можна записати як

$$C_b = S \sqrt{\frac{\varepsilon e n_n}{8\pi(\varphi_c + V)}}. \quad (14.121)$$

На ефекті зміни ємності за зміни зовнішньої різниці потенціалів будуються прилади, що називаються варикапами або варакторами.

Розглянемо динаміку носіїв заряду у **p-n-переході**. Для цього припустимо, що легування обох частин напівпровідника відбувалось так,

що $p_p \gg n_n$. Тоді, як випливає із (14.111) та (14.109), $L_n \approx L$. Концентрація дірок (неосновних носіїв) в n -області на межі розподілу $x = 0$ можна записати як

$$p_n(0) = N_V e^{(F_p - \phi_c)/kT} = N_V e^{F_p/kT} e^{-\phi_c/kT} = p_p e^{-\phi_c/kT}. \quad (14.122)$$

Прикладемо до p - n -переходу зовнішню напругу величиною V , що відповідає напрямку прямого зміщення. Згідно із (14.122), концентрацію неосновних носіїв заряду можна записати у вигляді

$$p_n^{(V)}(0) = p_p e^{-e(\phi_c - V)/kT} = p_n(0) e^{eV/kT}. \quad (14.123)$$

Таким чином, на межі розподілу p - n -переходу концентрація неосновних носіїв заряду збільшується у $e^{V/kT}$ разів. Таке явище, як нам вже відомо, називається інжекцією неосновних носіїв. При прикладенні напруги у зворотному напрямку концентрацію неосновних носіїв в n -області переходу можна записати як

$$p_n^{(V)}(0) = p_p e^{-e(\phi_c + V)/kT} = p_n(0) e^{-eV/kT}. \quad (14.124)$$

Таким чином, у випадку прикладення до p - n -переходу напруги у зворотному напрямку на межі розподілу відбувається зменшення концентрації неосновних носіїв у $e^{V/kT}$ разів. Це явище має назву екстракції носіїв.

Розглянемо тепер, як відбуватиметься випрямлення струму на p - n -переході. Для простоти уявимо тонкий p - n -перехід, в якому товщина квазінейтральних n - і p -областей мала, і вся прикладена зовнішня напруга падає на p - n -переході. Нехай до переходу прикладене пряме зміщення. Тоді відповідно до (14.123) на границі розподілу спостерігатиметься збільшення концентрації неосновних носіїв і, якщо падіння напруги на квазінейтральних областях немає, то струм буде обумовлений дифузиею неосновних носіїв – електронів до p -області та дірок до n -області, тобто

$$j_n = -eD_n \frac{dn_p}{dx}, \quad j_p = -eD_p \frac{dp_n}{dx}. \quad (14.125)$$

Оскільки концентрація неосновних носіїв заряду поблизу межі розподілу спадає за експонентою залежно від відстані із характерними довжинами L_p та L_n , то можна записати

$$p_n(x) = p_{n0} + [p_n(0) - p_{n0}] \cdot e^{-x/L_p}, \quad (14.126)$$

$$n_p(x) = n_{p0} + [n_p(0) - n_{p0}] \cdot e^{-x/L_n}. \quad (14.127)$$

Підставляючи ці вирази до (14.125), отримаємо для відповідних струмів

$$j_n = \frac{eD_n}{L_n} [n_p(0) - n_{p0}] = \frac{eD_n n_{p0}}{L_n} (e^{eV/kT} - 1), \quad (14.128)$$

$$j_p = \frac{eD_p}{L_p} [p_n(0) - p_{n0}] = \frac{eD_p p_{n0}}{L_p} (e^{eV/kT} - 1). \quad (14.129)$$

Тоді повний струм визначатиметься формулою

$$j = j_s (e^{eV/kT} - 1), \quad (14.130)$$

де

$$j_s = \frac{eD_n n_{p0}}{L_n} + \frac{eD_p p_{n0}}{L_p}. \quad (14.131)$$

За зворотної напруги матимемо

$$j = j_s (e^{-eV/kT} - 1). \quad (14.132)$$

Таким чином, вольт-амперна характеристика p - n -переходу матиме вигляд, що поданий на рис. 14.17. Необхідно особливо підкреслити, що струм через p - n -перехід переноситься неосновними носіями заряду. У цьому полягає принципова відмінність p - n -переходу від приладів на основі бар'єру Шоттки. Наслідком цього факту є те, що **швидкодія p - n -переходу визначається часом життя неосновних носіїв заряду, а швидкодія приладів на основі бар'єру Шоттки визначається, як зазначалось вище, часом максвеллівської релаксації.**

■ 14.7. Гетеропереходи

Раніше ми розглядали p - n -переходи, утворені напівпровідниками з однаковою величиною забороненої зони. Це були зразки напівпровідника, по-різному леговані в один та в інший бік від деякої межі, що мала назву p - n -переходу. Такі переходи називають гомопереходами, що підкреслює, що по обидва боки переходу ми маємо один і той самий матеріал. При контакті напівпровідників із різною шириною забороненої зони виникають гетеропереходи. Найчастіше їх отримують методом молекулярно променевої епітаксії. Застосовують також методи епітаксiального нарощування монокристалічних шарів із газової фази, напилення у вакуумі та ін. Однією із проблем при створенні гетероструктур є незбіг сталих ґратки компонентів гетероструктури. Дійсно, незбіг значень сталих ґратки різко погіршує умови епітаксії, зокрема він спричиняє виникнення механічних напружень на межі розподілу. Ці напруження викликають утворення дислокацій невідповідності та інших дефектів упаковки. Зрозуміло, що така дефектність інтерфейсу

може призводити до багатьох неприємностей – від різкого зменшення рухливості носіїв до відшарування плівок від підкладок – тобто до неможливості утворення гетероструктури як такої.

Згідно із табл. 14.1 систему AlGaAs/GaAs можна розглядати як приклад, де сталі ґратки збігаються. Дійсно, ця система має незначний незбіг сталих ґратки ($\sim 0,1\%$) у широкому діапазоні співвідношень концентрацій від Al до Ga. Такі структури виростають вільними від механічних напружень і надзвичайно досконалими – вони є основою для виготовлення промислово значимих гетероструктур. Іншим прикладом широко використовуваної системи для конструювання гетероструктур є GaAs/Ga_{1-x}In_xAs і Ga_{1-x}Al_xAs/Ga_{1-x}In_xAs. Такі структури використовуються, наприклад для виготовлення польових транзисторів (field-effect transistor).

Таблиця 14.1

Напівпровідник	E_g (eВ) при $T = 300\text{ К}$	Стала ґратки (нм)	Коефіцієнт лінійного розширення при $T = 300\text{ К}$ (10^{-4} К^{-1})	Енергія електронної спорідненості (eВ)
Ge	0,67	0,5658	5.7	4.13
Si	1,11	0.5431	2.33	4.01
GaAs	1,43	0.5653	5.8	4.07
AlAs	2,15	0.5661	5.2	
GaP	2,25	0.5451	5.3	4.3
GaSb	0,68	0.6095	6.9	5.06
AlSb	1,6	0.6136	3.7	3.65

Тепер, коли ми розібрались з деякими технологічними проблемами, що виникають при утворенні гетерограниць, спробуємо з'ясувати, яким чином формуватимуться енергетичні бар'єри у таких структурах. Для цього уявимо, що ми створюємо гетероперехід із двох напівпровідників, які характеризуються забороненими зонами E_{g1} та E_{g2} . Нехай для визначеності $E_{g2} > E_{g1}$. Яку енергетичну діаграму матиме утворений гетероперехід (рис. 14.18) а), б) чи в)?

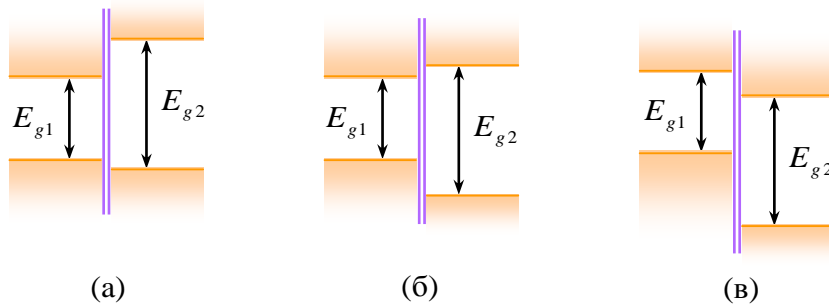


Рис. 14.18
Можливі типи гетеропереходів

Зрозуміло, що три можливі гетеропереходи мають принципово різний вигляд енергетичних діаграм, отже й різні електронні властивості. Який все ж таки тип енергетичної структури можна отримати, з'єднавши два напівпровідника із визначеними значеннями зонної щільності? Або що ми отримаємо у результаті такого "зонного конструювання" (до речі, в англomовній літературі цей процес так і називається band engineering). Відповідь на питання про формування гетеробар'єру надає так зване правило Андерсона, що базується на фундаментальному енергетичному параметрі – електронній спорідненості матеріалів χ – енергії, яка необхідна для видалення електрона із дна зони провідності E_C на вакуумний рівень (тобто видалення електрона із кристала). Ідея Андерсона базується на очевидному припущенні, що абсолютне значення вакуумного рівня є однаковим для всіх матеріалів. Звідси маємо, що різниця між енергіями дна зони провідності в гетеропереході

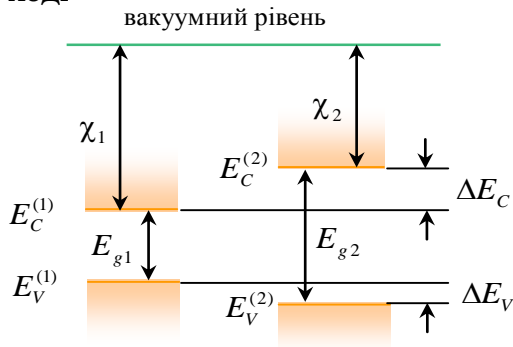


Рис. 14.19
Побудова енергетичної діаграми гетеро переходу за правилом Андерсона

$$\Delta E_C = E_C^{(2)} - E_C^{(1)} = \chi_1 - \chi_2. \quad (14.133)$$

Тоді, якщо відоме значення різниці енергій забороненої зони ΔE_g , то легко визначається різниця між рівнями стель валентної зони (рис. 14.19)

$$\Delta E_V = \Delta E_g - \Delta E_C. \quad (14.134)$$

Наприклад, розглянемо гетеросистему GaAs-Al_{0.3}Ga_{0.7}As. Відомо, що $\chi_{\text{GaAs}} = 4,07$, а $\chi_{\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}} = 3,74$ еВ.

Звідси $\Delta E_C = 0,33$ еВ. Оскільки

різниця енергії зонних щілин $\Delta E_g = 0,37 \text{ eV}$, то отримуємо $\Delta E_v = 0,04 \text{ eV}$.

Таким чином утворюється гетеропереход, де зонна щілина матеріалу, який має менше значення ширини забороненої зони, розташований усередині енергетичного зазору зонної щілини матеріалу, що має більшу величину забороненої зони. Такий гетероперехід називається гетеропереходом I типу (рис. 14.20 а).

Іншим прикладом гетероперехода є так званий гетеропереход II типу (рис. 14.20 б). Існує й третя можливість поєднання двох напівпро-

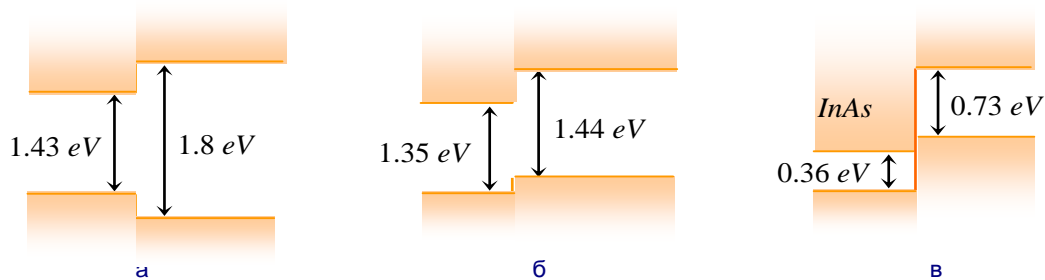


Рис. 14.20
Гетеропереходи I (а), II (б), та III (в) типів

відникових матеріалів у гетероструктуру – це так звана гетероструктура III типу. У цьому випадку на межі розподілу відбувається перекриття забороненої зони одного із компонентів гетеросистеми. Англійською мовою цей тип гетеросистеми так і називається – *broken-gap alignment*. Класичним прикладом гетеросистеми III типу є гетеропереход InAs-GaSb. Цей приклад як енергетичну діаграму подано на рис. 14.20 в. Завдяки тому, що тут перекриваються зона провідності одного компонента із валентною зоною іншого, електрони та дірки спонтанно переходитимуть з одного боку на інший, доки не будуть зупинені наведеним електричним полем, аналогічно тому, як формується область збіднення у *p-n* діоді. Залежно від наявності домішок обидва напівпровідники, що утворюють гетероперехід, можуть мати:

- ♦ однаковий тип провідності – тоді іноді їх називають ізотипними гетеропереходами,
- ♦ різний тип провідності – про такі переходи говорять як про анізотипні.

Зрозуміло, що енергетичні діаграми гетеропереходів, що містять вигини зон і виникаючі при цьому потенціальні бар'єри, відрізняються від діаграм для гомопереходів. Для демонстрації цього розглянемо побудову зонної діаграми ідеального різкого *n-p*-гетеропереходу,

вважаючи, що більшою величиною забороненої зони характеризується напівпровідник p -типу. Припустимо, що напівпровідник n -типу розташований в області $x < 0$, а напівпровідник p -типу – в області $x > 0$. Гетерограниця збігається із площиною $x = 0$. При приведенні напівпровідників (рис. 14.19) у контакт у системі встановиться термодинамічна рівновага, що характеризується єдиним Фермі-рівнем для обох напівпровідників і наявністю контактної різниці потенціалів $\varphi_C = (\Phi_1 - \Phi_2)/e$, де Φ_1 і Φ_2 – роботи виходу напівпровідників n - та p -типів, відповідно. В ідеальному різкому гетеропереході контактний потенціал $V(x)$ та енергія електрона поблизу поверхні зразка $e\varphi(x)$ мають бути неперервними функціями координати x , нормальної до гетерограниці. При цьому виконується рівність $e\varphi(x) = V(x)$. Це приводить до того, що нормальна компонента вектора електричної індукції також має бути неперервною: $D_1 = D_2$. Якщо напівпровідники характеризуються діелектричної сприйнятливостями ε_1 та ε_2 , то маємо

$$\varepsilon_1 E_1 = \varepsilon_2 E_2, \quad (14.135)$$

де E_1 та E_2 – нормальні компоненти електричного поля в n - та p -напівпровідниках поблизу межі розподілу. Тому на границі гетеропереходу нормальна компонента електричного поля має розрив. А оскільки

$$E(x) = -\frac{dV(x)}{dx}, \quad (14.136)$$

то величини $V(x)$ і $\varphi(x)$ мають злам. Якщо припустити, що електронна спорідненість та енергія забороненої зони напівпровідників, що утворюють гетероперехід, є сталими вздовж зразків аж до межі розподілу та не дорівнюють одне одному, то на границі гетеропереходу відбуваються стрибки згідно із правилом Андерсена (див. (14.133) і (14.134)) $\Delta E_C = \chi_1 - \chi_2$, $\Delta E_V = E_{g2} - E_{g1} - \Delta E_C$. Знайдемо хід країв зон $E_C(x)$ та $E_V(x)$ по обидва боки гетеропереходу із умови неперервності нормальної компоненти електричної індукції на межі розподілу (14.135), яку із врахуванням (14.136) можна записати у вигляді

$$\varepsilon_1 \left. \frac{d\varphi_1}{dx} \right|_{x=0} = \varepsilon_2 \left. \frac{d\varphi_2}{dx} \right|_{x=+0}. \quad (14.137)$$

При цьому вважатимемо, що повна контактні різниця потенціалів φ_C розподіляється між двома напівпровідниками на частини φ_{C_1} та φ_{C_2} так, що

$$\varphi_C = \varphi_{C_1} + \varphi_{C_2}. \quad (14.138)$$

Припустимо, що напівпровідники леговані мілкими домішками так, що донори та акцептори є повністю іонізованими, а концентрацією електронів і дірок у глибині зразків n - та p -типів є n_0 та p_0 , відповідно. Тоді рівняння Пуассона по обидва боки переходу можна записати у вигляді

$$\frac{d^2\varphi}{dx^2} = -\frac{4\pi en_0}{\varepsilon_1}, \quad x < 0, \quad (14.139)$$

$$\frac{d^2\varphi}{dx^2} = \frac{4\pi ep_0}{\varepsilon_2}, \quad x > 0. \quad (14.140)$$

Інтегруючи ці рівняння із граничними умовами

$$\varphi|_{x=-d_1} = \varphi_1, \quad \left. \frac{d\varphi}{dx} \right|_{x=-d_1} = 0; \quad \varphi|_{x=d_2} = \varphi_2, \quad \left. \frac{d\varphi}{dx} \right|_{x=d_2} = 0, \quad (14.141)$$

отримуємо хід потенціалу

$$\text{в } n\text{-області} \quad \varphi_1(x) = \varphi_1 - \frac{2\pi en_0}{\varepsilon_1} (x - d_1)^2, \quad (14.142)$$

$$p\text{-області} \quad \varphi_2(x) = \varphi_2 + \frac{2\pi ep_0}{\varepsilon_2} (x + d_2)^2. \quad (14.143)$$

Підставляючи ці розв'язки до (14.137), отримаємо

$$\frac{d_1}{d_2} = \frac{p_0}{n_0}. \quad (14.144)$$

Покладемо у формулах (14.143) та (14.144) $x = 0$. Зауважуючи, що $\varphi_1(0) - \varphi_1 = \varphi_{C_1}$ і $\varphi_2 - \varphi_2(0) = \varphi_{C_2}$, знайдемо

$$\varphi_{C_1} = \frac{2\pi en_0}{\varepsilon_1} d_1^2, \quad (14.145)$$

$$\varphi_{C_2} = \frac{2\pi ep_0}{\varepsilon_2} d_2^2. \quad (14.146)$$

Повна контактна різниця потенціалів звідси

$$\varphi_C = 2\pi e \left(\frac{n_0 d_1^2}{\varepsilon_1} + \frac{p_0 d_2^2}{\varepsilon_2} \right). \quad (14.147)$$

Використовуючи співвідношення (14.144), із (14.145-146) отримуємо

$$\frac{\varphi_{C_1}}{\varphi_{C_2}} = \frac{p_0}{n_0} \frac{\varepsilon_2}{\varepsilon_1}. \quad (14.148)$$

Виключаючи у (14.147) за допомогою співвідношення (14.144) послідовно d_2 і d_1 , отримуємо товщини шарів об'ємного заряду з n - та p -боквів гетеропереходу, відповідно

$$d_1 = \sqrt{\frac{\Phi_C \varepsilon_1 \varepsilon_2 p_0}{2\pi e n_0 (\varepsilon_2 p_0 + \varepsilon_1 n_0)}}, \quad (14.149)$$

$$d_2 = \sqrt{\frac{\Phi_C \varepsilon_1 \varepsilon_2 n_0}{2\pi e p_0 (\varepsilon_2 p_0 + \varepsilon_1 n_0)}}. \quad (14.150)$$

Отримані співвідношення повною мірою визначають енергетичну діаграму гетеропереходу (рис. 14.21). Зауважимо, що проведений розгляд енергетичних діаграм є ідеалізованим, оскільки не бралась до уваги наявність локальних поверхневих енергетичних рівнів. Ми згадували, що такі рівні завжди виникають у випадках порушення періодичності ґратки, наприклад поблизу поверхні, а також на межі розподілу якою є гетероперехід. Якщо концентрація поверхневих рівнів є великою, на

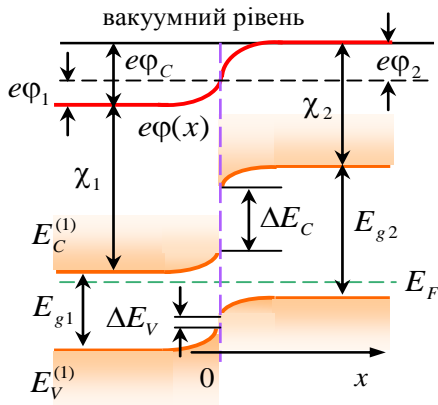


Рис.14.21
Енергетична діаграма ідеального різкого p - n -гетеропереходу

межі розподілу може виникнути досить значний електричний заряд, і вигин зон на границі гетероперехода може суттєво спотворитись. Гетеропереходи широко використовуються у сучасній мікроелектроніці, наприклад як ефективні інжектори дірок у матеріал n -типу. Це звільняє від необхідності мати сильно леговану область p -типу. Гетеропереходи є головним конструкційним елементом при створенні квантових ям і надґраток, що широко використовуються при створенні світловипромінюючих елементів і лазерів.

■ 14.8. Біполярний транзистор

Розглядаючи процес випрямлення струму p - n -переходом, ми, зокрема визначили, що при проходженні струму у пропускному режимі (за прямого зміщення) концентрація неосновних носіїв в області переходу значно збільшується. Ці надлишкові носії, наприклад дірки, що перейшли із p -області переходу до n -області, дифундують у глибину

n -області. Як результат опір частини n -переходу, що безпосередньо примикає до p - n -переходу, знижується. Якщо концентрація неосновних носіїв при цьому залишається меншою від концентрації електронів, то загальна зміна опору n -області залишається невеликою, незважаючи на те, що концентрація дірок може збільшуватись на кілька порядків порівняно із рівноважною. Якщо в n -області поблизу p - n -переходу утворити ще один p - n -перехід, що включений у запірному напрямку, то струм через нього, який визначається концентрацією неосновних носіїв (дірок) різко зросте. Оскільки внутрішній опір p - n -переходу, що включений у зворотному напрямку, дуже великий, тобто струм через такий перехід є майже незмінним навіть за значної зміни напруги, то до електричного кола можна включити великий зовнішній опір R_n (рис. 14.22). Тоді за незначної зміни напруги, прикладеної до p - n -переходу, що включений у прямому напрямку, струм через другий p - n -перехід сильно змінюватиметься, та із включеного до кола зовнішнього опору можна знімати великі напруги. Це означає, що такий прилад (*транзистор*) можна використовувати як підсилювач електричних сигналів. Такий прилад, що складається із двох p -областей, які розділені n -областю, називають p - n - p -транзистором. Існують прилади, що мають іншу структуру – n - p - n -транзистори, а також велика кількість методів виготовлення транзисторів, зокрема вплавлення, дифузія та епітаксіальне нашарування. Наприклад, виготовлення транзистором методом дифузії проводиться шляхом одночасної дифузії до p -матеріалу дононої та акцепторної домішки, наприклад In і Sb.

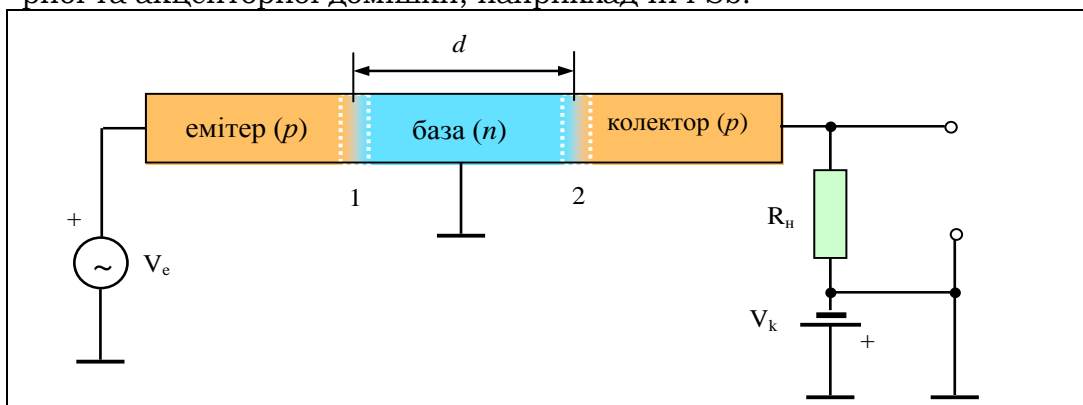
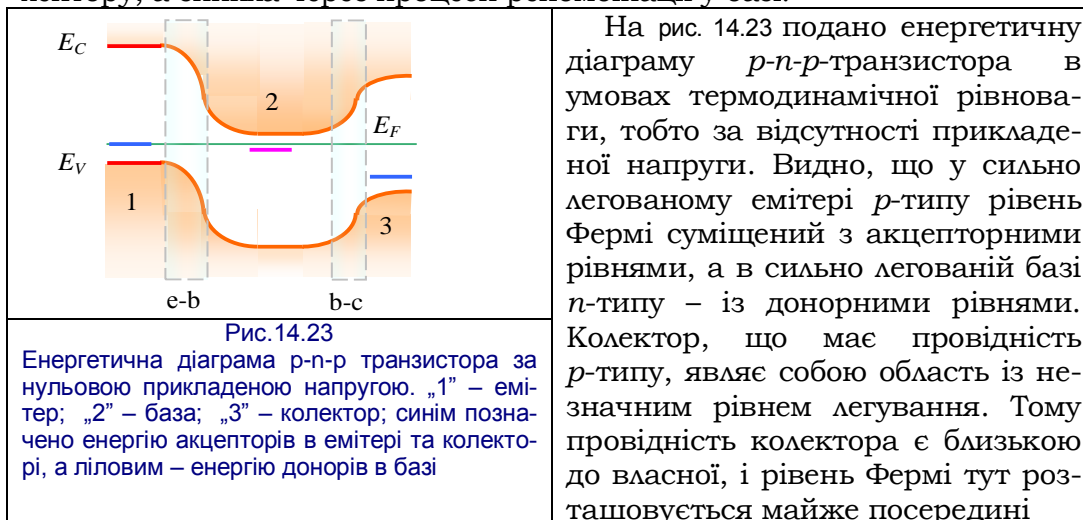


Рис. 14.22

Схема p - n - p -транзистора. „1” – емітерний p - n -перехід; „2” – колекторний p - n -перехід

У процесі виготовлення сурма дифундує швидше за індій та утворює n -область. Концентрація індію береться істотно більшою за концентрацію сурми. Із цієї причини та область, до якої продифундував індій,

знову має p -провідність. Ця область є емітером, тобто областю, звідки дірки вводяться у базу. Область зразку, де переважає сурмик, що має провідність n -типу, утворює базу, а основний матеріал, до якого домішки не продифундували, служить колектором, тобто областю, до якої дірки збираються із бази. У приладів, що виготовлені за такою технологією, колекторна область є слабо легованою, тобто в ній концентрація дірок є меншою за концентрацію електронів у базі. Емітерна ж область завжди є сильно легованою. Ще однією технологічною рисою транзисторів є те, що базова область завжди виготовляється тонкою – її товщина d має бути меншою за дифузійну довжину. У зворотному випадку більша частина інжекттованих із емітера дірок взагалі не досягла б колектору, а зникла через процеси рекомбінації у базі.



На рис. 14.23 подано енергетичну діаграму p - n - p транзистора в умовах термодинамічної рівноваги, тобто за відсутності прикладеної напруги. Видно, що у сильно легованому емітері p -типу рівень Фермі суміщений з акцепторними рівнями, а в сильно легованій базі n -типу – із донорними рівнями. Колектор, що має провідність p -типу, являє собою область із незначним рівнем легування. Тому провідність колектора є близькою до власної, і рівень Фермі тут розташовується майже посередині

забороненої зони. Потенціальна яма в області бази є бар'єром для дифузії дірок із емітера до колектора. Потік дірок через бар'єр керується напругою V_E між емітером і базою.

Для визначення залежності струмів у транзисторі від параметрів приладу та прикладених різниць потенціалів розглянемо спочатку область бази. Вважатимемо, що домішка у цій області приладу розподілена рівномірно, тобто концентрація електронів n_n є сталою вздовж довжини бази. Нехай на прилад подано сталі напруги. Між емітером і колектором прикладено напругу у пропускну режимі V_e , а між колектором та базою – напругу в запірному режимі, що має абсолютну величину V_c . За невеликих рівнів інжекції, коли концентрація дірок у базовій області p не перевищує рівноважну концентрацію електронів n_n , напруги V_e та V_c , в основному падають на переходах **емітер-база** та **база-колектор**. При цьому граничні концентрації неосновних носіїв

заряду визначаються характеристиками тонкого p - n -переходу, тобто на межі емітерного переходу (при $x = 0$)

$$p(0) = p_n e^{eV_E/kT}, \quad (14.151)$$

а на межі колекторного переходу (при $x = d$)

$$p(d) = p_n e^{-eV_C/kT}. \quad (14.152)$$

У цих формулах через p_n позначено рівноважну концентрацію неосновних носіїв в області бази. Малі рівні інжекції свідчать також, що дрейфовий струм неосновних носіїв є набагато меншим дифузійного. Це означає, що розподіл неосновних носіїв у базі визначається рівнянням дифузійного руху

$$\frac{d^2 p}{dx^2} = \frac{p - p_n}{L_b}, \quad (14.153)$$

Де $\Delta p = p - p_n$ - надлишкова концентрація, а L_b дифузійна довжина неосновних носіїв заряду. Розв'язок (14.153) зручно шукати, якщо представити Δp у вигляді суми $\Delta p_1 + \Delta p_2$ так, щоб Δp_1 і Δp_2 задовольняли це рівняння та граничні умови

$$\Delta p_1(0) = p_n (e^{eV_E/kT} - 1), \quad \Delta p_2(0) = 0, \quad (14.154)$$

$$\Delta p_1(d) = 0, \quad \Delta p_2(d) = p_n (e^{-eV_C/kT} - 1). \quad (14.155)$$

Розв'язок (14.153) шукатимемо у вигляді

$$\Delta p_1 = A_1 \operatorname{sh}\left(\frac{d-x}{L_b}\right) + B_1 \operatorname{ch}\left(\frac{d-x}{L_b}\right), \quad (14.156)$$

$$\Delta p_2 = A_2 \operatorname{sh}\left(\frac{x}{L_b}\right) + B_2 \operatorname{ch}\left(\frac{x}{L_b}\right). \quad (14.157)$$

Для виконання граничних умов (14.154–14.155) необхідно покласти

$$A_1 = \frac{p_n}{\operatorname{sh}(d/L_b)} [e^{eV_E/kT} - 1], \quad A_2 = \frac{p_n}{\operatorname{sh}(d/L_b)} [e^{-eV_C/kT} - 1], \quad B_1 = B_2 = 0. \quad (14.158)$$

Підставивши ці константи до (14.156–14.157), отримаємо

$$\Delta p = \frac{p_n}{\operatorname{sh}(d/L_b)} \left[(e^{eV_E/kT} - 1) \cdot \operatorname{sh}\left(\frac{d-x}{L_b}\right) - (1 - e^{-eV_C/kT}) \cdot \operatorname{sh}\left(\frac{x}{L_b}\right) \right]. \quad (14.159)$$

Використаємо отриманий вираз для визначення діркового струму емітера та колектора, маючи на увазі, що цей струм має дифузійну природу.

$$j_{p,e} = -eD_p \left. \frac{dp}{dx} \right|_{x=0} = \frac{eD_p p_n}{L_b} \left[\left(e^{eV_E/kT} - 1 \right) \cdot \operatorname{cth} \left(\frac{d}{L_b} \right) + \left(1 - e^{-eV_C/kT} \right) \cdot \operatorname{csch} \left(\frac{d}{L_b} \right) \right], \quad (14.160)$$

$$j_{p,c} = -eD_p \left. \frac{dp}{dx} \right|_{x=d} = \frac{eD_p p_n}{L_b} \left[\left(e^{eV_E/kT} - 1 \right) \cdot \operatorname{csch} \left(\frac{d}{L_b} \right) + \left(1 - e^{-eV_C/kT} \right) \cdot \operatorname{cth} \left(\frac{d}{L_b} \right) \right]. \quad (14.161)$$

Різниця цих струмів визначає рекомбінаційний струм бази

$$j_r = j_{p,e} - j_{p,c}. \quad (14.162)$$

Підставляючи сюди вирази для емітерного та колекторного струмів, знайдемо

$$j_r = \frac{eD_p p_n}{L_b} \left[\left(e^{eV_E/kT} + e^{-eV_C/kT} - 2 \right) \left(\operatorname{cth} \left(\frac{d}{L_b} \right) - \operatorname{csch} \left(\frac{d}{L_b} \right) \right) \right]. \quad (14.163)$$

Беручи до уваги, що $\operatorname{cth}x - \operatorname{csch}x = (\operatorname{ch}x - 1)/\operatorname{sh}x = \operatorname{th}(x/2)$, запишемо

$$j_r = \frac{eD_p p_n}{L_b} \left[\left(e^{eV_E/kT} + e^{-eV_C/kT} - 2 \right) \operatorname{th} \left(\frac{d}{2L_b} \right) \right]. \quad (14.164)$$

Оскільки у робочому режимі напруга на колекторі істотно перевищує величину kT/e , то доданком $\exp(-eV_C/kT)$ можна знехтувати. Необхідно також скористатись тим фактом, що для зменшення рекомбінаційних втрат база створюється тонкою, тобто гіперболічний тангенс можна розкласти у ряд за малим параметром $d/(2L_b)$: необхідно покласти $\operatorname{th}(d/2L_b) \approx d/(2L_b)$. Тоді струм (14.164) можна представити у вигляді

$$j_r = \frac{eD_p p_n}{L_b} \left(e^{\frac{eV_E}{kT}} - 2 \right) \frac{d}{2L_b} = \frac{ep_n d}{\tau_p} \left(e^{\frac{eV_E}{kT}} - 2 \right). \quad (14.165)$$

За припущення, що процеси рекомбінації не впливають на формування струму, маємо: струм $j_p = -eD_p(dp/dx)$ є сталим, тобто сталою є похідна dp/dx . Іншими словами, можна стверджувати, що

$$\frac{dp}{dx} = -\frac{p(0) - p(d)}{d}. \quad (14.166)$$

Як видно із (14.155), при $eV_C \gg kT$, $\Delta p_2(d) = -p_n$, тобто $p(d) = 0$. Таким чином, маємо

$$dp(x) = -\frac{p(0)}{d} dx \rightarrow p(x) = -\frac{p(0)}{d} \int_d^x dx = \frac{p(0)}{d} (d - x). \quad (14.167)$$

Звідси, враховуючи, що згідно із (14.151) $p(0) = p_n e^{eV_E/kT}$, дістанемо

$$p(x) = p_n e^{eV_E/kT} (1 - x/d). \quad (14.168)$$

Отже в області бази концентрація неосновних носіїв лінійно змінюється із координатою, зменшуючись від величини $p(0) = p_n e^{eV_E/kT}$ на емітерному p - n -переході до нуля на колекторному p - n -переході.

Відношення діркового струму колектора до струму емітера характеризує ефективність транзистора Ця величина зазвичай позначається через β .

$$\beta = \frac{j_{p,e}}{j_{p,c}}. \quad (14.169)$$

У випадку $eV_E \gg kT$ маємо із (14.160–14.161)

$$j_{p,c} \cong \frac{eD_p p_n}{L_b} e^{eV_E/kT} \operatorname{csch}\left(\frac{d}{L_b}\right), \quad (14.170)$$

$$j_{p,e} \cong \frac{eD_p p_n}{L_b} e^{eV_E/kT} \operatorname{cth}\left(\frac{d}{L_b}\right). \quad (14.171)$$

Підставляючи ці вирази до (14.169), отримаємо

$$\beta = \operatorname{csh}\left(\frac{d}{L_b}\right) \cong 1 - \frac{1}{2} \left(\frac{d}{L_b}\right)^2. \quad (14.172)$$

Параметр $1-\beta$ характеризує рекомбінаційні втрати при проходженні неосновними носіями області бази

$$1 - \beta = \frac{j_r}{j_{p,e}} \cong \frac{1}{2} \left(\frac{d}{L_b}\right)^2. \quad (14.173)$$

Отриманий результат ще раз переконує, що для найефективнішої роботи транзистора варто застосовувати такі технологічні прийоми, за яких параметр d/L_b буде мінімальний. Наприклад, у германієвому транзисторі, де дифузійна довжина дірок становить ~ 0.5 мм за товщини бази ~ 0.1 мм, отримуємо $1 - \beta \approx 0,1$, тобто тільки десять відсотків дірок, що утворюють струм, не доходять до колектора.

Повний струм емітера, крім діркової компоненти, що створюється дірками, які втікають до бази, включає ще струм, який утворюється

електронами, що йдуть із бази. Оскільки концентрація носіїв на p - n -переході залежить від прикладеної напруги, то остання у прямому напрямку збільшує концентрацію електронів на межі поділу **емітер-база**.

$$n_e^{(b)} = n_e e^{eV_E/kT}, \quad (14.174)$$

де n_e – рівноважна концентрація електронів в емітері. Електрони $n_e^{(b)}$ дифундують углиб емітера та створюють електронний струм в емітері. Якщо товщина емітерної області перевищує дифузійну довжину електронів, то через p - n -перехід потече електронний струм

$$j_{n,e} = \frac{eD_n n_e}{L_n} (e^{eV_E/kT} - 1). \quad (14.175)$$

Таким чином, повний струм через емітер запишемо у вигляді

$$j_{tot,e} = j_{p,e} + j_{n,e} \quad (14.176)$$

або

$$j_{tot,e} = \frac{eD_p p_n}{L_D} \left[(e^{eV_E/kT} - 1) \operatorname{cth} \left(\frac{d}{L_D} \right) + (1 - e^{-eV_C/kT}) \operatorname{csch} \left(\frac{d}{L_D} \right) \right] + \frac{eD_n n_e}{L_n} (e^{eV_E/kT} - 1). \quad (14.177)$$

Аналогічно обчислимо повний струм через колектор

$$j_{p,c} = \frac{eD_p p_n}{L_b} \left[(e^{eV_E/kT} - 1) \cdot \operatorname{csch} \left(\frac{d}{L_b} \right) + (1 - e^{-eV_C/kT}) \cdot \operatorname{cth} \left(\frac{d}{L_b} \right) \right] - \frac{eD_n n_e}{L_n} (e^{eV_C/kT} - 1). \quad (14.178)$$

Визначимо коефіцієнт підсилення за струмом при включенні приладу, як показано на рис. 14.23, коли на вхід емітера подається слабкий сигнал, а вихідний сигнал знімається із навантажувального опору, що включений до колекторного ланцюга. Таким чином, коефіцієнт підсилення за струмом є відношенням диференціалу колекторного струму до диференціалу емітерного струму за сталого потенціалу на колекторі, тобто

$$\alpha = \left(\frac{\partial j_{tot,c}}{\partial j_{tot,e}} \right)_{V_C}. \quad (14.179)$$

Зручніше визначати не відношення диференціалів, а відношення похідних

$$\alpha = \frac{\partial j_{tot,c} / \partial V_e}{\partial j_{tot,e} / \partial V_e} \Big|_{V_C = \text{const}}. \quad (14.180)$$

Використовуючи (14.178–14.179), обчислимо похідні

$$\left. \frac{\partial j_{tot,c}}{\partial V_E} \right|_{V_C=\text{const}} = \frac{eD_p p_n}{L_b} \cdot \text{csch}\left(\frac{d}{L_b}\right) \cdot \frac{e}{kT} e^{eV_E/kT}, \quad (14.181)$$

$$\left. \frac{\partial j_{tot,e}}{\partial V_E} \right|_{V_C=\text{const}} = \left[\frac{eD_p p_n}{L_b} \cdot \text{cth}\left(\frac{d}{L_b}\right) + \frac{eD_n n_e}{L_n} \right] \frac{e}{kT} e^{eV_E/kT}. \quad (14.182)$$

Підставляючи ці вирази до (14.180), дістанемо

$$\alpha = \left[\text{ch}\left(\frac{d}{L_b}\right) + \frac{D_n L_b n_e}{D_p L_n p_n} \text{sh}\left(\frac{d}{L_b}\right) \right]^{-1}. \quad (14.183)$$

Легко бачити, що всі величини, які входять до знаменника правої частини отриманого виразу, є додатними. Це означає, що коефіцієнт підсилення за струмом при включенні транзистора за схемою (рис. 14.23) завжди буде меншим від одиниці, тобто відсутність підсилення. Але цю величину можна зробити якомога ближчою до одиниці. Зокрема, як видно із (14.183), для цього товщина бази має бути набагато меншою за дифузійну довжину дірок, а також база має бути сильно легованою, тобто мають виконуватись нерівності $d \ll L_b$ і $p_n \gg n_e$. Оскільки довжини дифузії і коефіцієнти дифузії електронів і дірок є величинами одного порядку, із (14.183) отримаємо

$$\alpha = \left[\text{ch}\left(\frac{d}{L_b}\right) \right]^{-1} \approx 1 - \frac{1}{2} \left(\frac{d}{L_b}\right)^2. \quad (14.184)$$

Оцінка для германієвого приладу із параметрами, які використовувались вище ($D_p \sim 0,5$ мм, $d \sim 0,1$ мм), дає $\alpha = 0,98$. Отримана оцінка показує, що за такого ввімкнення приладу до електричного кола підсилення за струмом не відбувається, тобто струм емітера майже дорівнює струму колектора. Цікаво, чи відбувається при цьому підсилення за потужністю? Коефіцієнт підсилення за потужністю w можна визначити як відношення потужності, що розсіюється колекторним ланцюгом, до потужності, яка розсіюється емітером. Нехай r_C і r_E – диференціальні опори колектора та емітера, відповідно. Тоді

$$w = \frac{j_C^2 r_C}{j_E^2 r_E} \cong \frac{r_C}{r_E}. \quad (14.185)$$

Обернений диференційний опір колектора можна визначити як

$$(r_C)^{-1} = \left(\frac{\partial j_C}{\partial V_C} \right)_{V_E=\text{const}}. \quad (14.186)$$

Оцінимо цю величину, беручи до уваги залежність товщини бази від прикладеної колекторної напруги. Для цього згадаємо, що, згідно із

(14.118) товщина p - n -переходу залежить від прикладеної напруги як $L = \text{const} \sqrt{\varphi_C + V}$. Таким чином, можна записати залежність товщини бази від прикладеної напруги на колекторі

$$d(V_C) = d_0 \sqrt{\varphi_C + V_C} / \sqrt{\varphi_C}, \quad (14.187)$$

де d_0 – товщина бази за нульової напруги на колекторі. Тоді можна записати

$$\frac{\partial j_C}{\partial V_C} = \frac{\partial j_C}{\partial d} \frac{\partial d}{\partial V_C} \approx \frac{j_C}{d} \frac{\partial d}{\partial V_C}. \quad (14.188)$$

Використовуючи (14.187), маємо

$$\frac{\partial d}{\partial V_C} = \frac{d_0}{2\sqrt{\varphi_C} \sqrt{\varphi_C + V_C}}, \quad (14.189)$$

звідки дістаємо оцінку

$$(r_C)^{-1} \approx \frac{j_C}{2(\varphi_C + V_C)}. \quad (14.190)$$

З іншого боку, обернений диференційний опір емітера визначається як

$$(r_E)^{-1} = \left(\frac{\partial j_E}{\partial V_E} \right)_{V_C = \text{const}}. \quad (14.191)$$

Формулу (14.182) можна наближено представити у вигляді

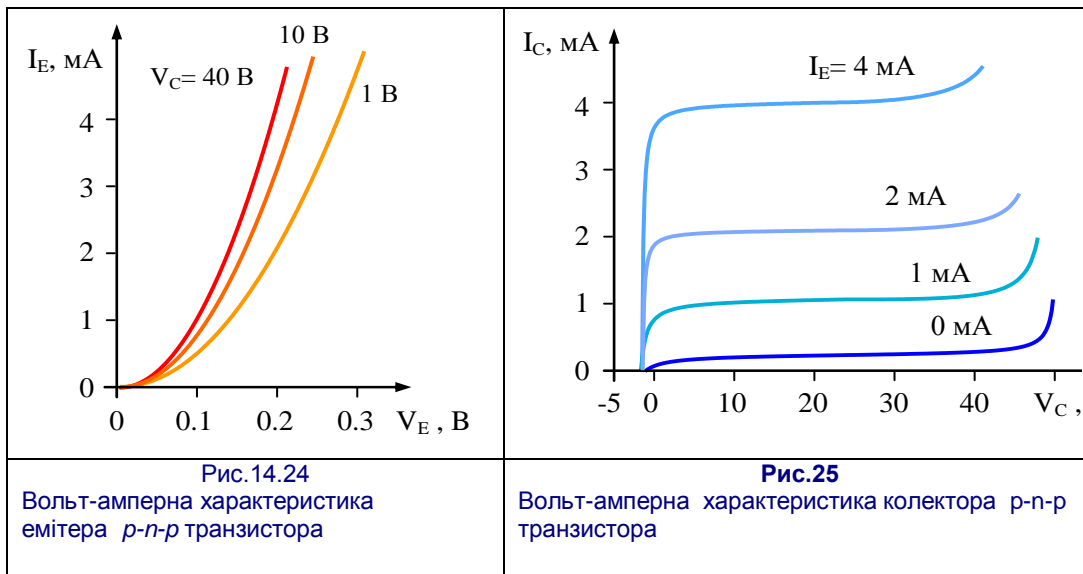
$$\left(\frac{\partial j_E}{\partial V_E} \right)_{V_C = \text{const}} \approx j_E \frac{e}{kT}. \quad (14.192)$$

Таким чином, із (14.185) дістаємо

$$\omega \approx \frac{2e(\varphi_C + V_C)}{kT}. \quad (14.193)$$

За кімнатної температури $e/kT \sim 26 \cdot 10^{-3} \text{ В}^{-1}$. Типові значення бар'єру на p - n -переході $\varphi_C \cong 0,25 \text{ В}$ та потенціалу на колекторі $V_C \cong 4 \text{ В}$. За таких параметрів отримуємо, що коефіцієнт підсилення становить $3 \cdot 10^2$.

На рис. 14.24 та 14.25 подано вольт-амперні характеристики типового транзистора. Емітерні характеристики є прямими, а колекторні – оберненими характеристиками p - n -переходу. Для колекторних характеристик струм емітера є параметром. За зростання колекторної напруги до великих значень струм колектора лишається сталим i , оскільки $\alpha \approx 1$, $I_C \approx I_E$. Якщо у транзисторі база є дуже тонкою та слабо легованою, то за достатньо високих напруг на колекторі настає пробій.



■ 14.9. ПОЛЬОВИЙ ТРАНЗИСТОР

У попередньому розділі ми розглянули ефект поля – зміну поверхневої провідності напівпровідника під дією поперечного електричного поля. Цей ефект широко використовується у польових транзисторах. Польовий транзистор являє собою тонку напівпровідникову пластинку (канал) із двома омичними контактами – стоком і виток, між якими розташовують третій електрод – затвор. Електрична напруга, що подається між затвором і стоком чи виток, спричиняє появу у підзатворній області каналу електричного поля, яке впливає на концентрацію носіїв заряду поблизу затвору, наслідком чого є зміна електричного опору каналу. Існують два основних типи польових транзисторів: до першого – належать польові транзистори, в яких роль затвору відіграє *p-n*-перехід або бар'єр Шотткі; до другого – належать прилади, в яких металічний електрод затвору відділений від каналу тонким шаром діелектрика – польовий транзистор з ізольованим затвором або транзистор на основі структури **метал-діелектрик-напівпровідник**. На рис. 14.26 подано схему польового транзистора із затвором, що відокремлений *p-n*-переходом, на рис. 14.27 – схему польового транзистора іншого типу – транзистор на структурі **метал-діелектрик-напівпровідник**. Такі прилади часто називають МДН-транзисторами.

Розглянемо детальніше роботу польового транзистора на прикладі транзистора із затвором на *p-n*-переході. У такому приладі, одну із

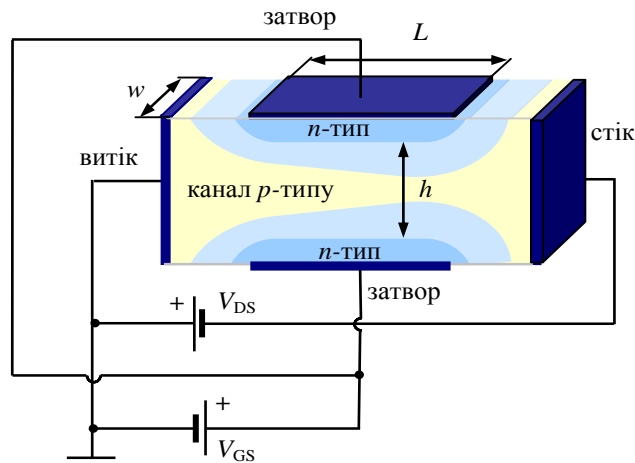


Рис. 14.26. Схема польового транзистора з затвором, що відокремлений р-n-переходом

можливих схем якого подано на рис. 14.26, омичні контакти до лівої та правої граней напівпровідникового кристалу – витік і стік. Область квазінейтрального об'єму, що розташована між збідненими областями p - n -переходів, утворює канал, а сильно леговані n^+ області зверху та знизу відіграватимуть роль затвора польового транзистора. Прикладення напруги $V_{GS} > 0$ до затвора, що забезпечує зворотне зміщення p - n -переходу, спричиняє розширення збідненої області (на рис. 14.26 подано світло-блакитним) p - n -переходу до кристалу, оскільки затвор легований істотно сильніше за область каналу ($N_D \gg N_A$). Цей процес зменшує поперечний переріз каналу, що збільшує його опір. Прикладення напруги у напрямку **витік–стік** V_{DS} викличе створення струму в

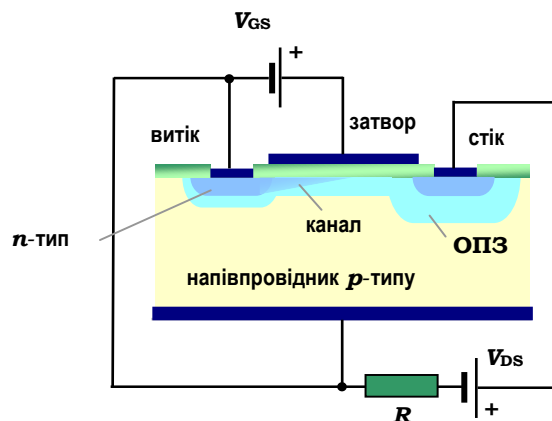


Рис. 14.27. Схема польового МДН-транзистора. Шар діелектрика позначено зеленим

ланцюгу каналу приладу. При цьому, знак напруги V_{DS} потрібно вибирати так, щоб напруга також викликала зворотне зміщення p - n -переходу на затворі, тобто V_{DS} повинна мати знак, протилежний до V_{GS} . Таким чином, польовий транзистор із затвором у вигляді p - n -переходу являє собою електричний опір, величину якого можна регулювати зовнішньою напругою. Зауважимо: якщо у біполярному транзисторі, про який ішлося раніше, електричний струм у приладі забезпечувався за рахунок неосновних носіїв, то у польовому транзисторі електричний струм утворюють основні носії заряду. Для визначення вольт-амперної характеристики транзистора виберемо систему координат так, щоб її вісь OY була спрямована вздовж каналу, вісь OX – вздовж ширини каналу, а вісь OZ – углиб каналу. Нехай довжиною, шириною та висотою каналу в умовах відсутності на транзисторі напруги є L , w та h , відповідно. Прикладення до затвору напруги $V_{GS} > 0$, а до стоку – напруги іншого знаку $V_{DS} < 0$ приведе до поширення збідненої області p - n -переходу на величину

$$\Delta l_f = l_f(V_{GS}, V_{DS}) - l_f(0, 0). \quad (14.194)$$

На основі (14.117), вважаючи, що енергія відраховується від рівня φ_C , можна записати, що товщина ОПЗ p - n -переходу, на який подано зворотне зміщення $V = V_{GS} - V_{DS}$,

$$\Delta l_f \cong \sqrt{\frac{\varepsilon(V_{GS} - V_{DS}(y))}{2\pi e N_D}}. \quad (14.195)$$

Оскільки напруга **витік-стік** розподілена вздовж каналу, то зміна ширини області збіднення p - n -переходу завдяки залежності $V_{DS}(y)$ від координати y , буде різною вздовж довжини каналу. Це приведе до того, що висота каналу також буде різною вздовж каналу

$$h(y) = h - 2\Delta l_f = h - 2\sqrt{\frac{\varepsilon(V_{GS} - V_{DS}(y))}{2\pi e N_D}}. \quad (14.196)$$

Позначимо через V_{G0} напругу на затворі, за якої у квазірівноважних умовах збіднені шари p - n -переходів змикаються, тим самим зануляючи висоту каналу. Таким чином, за умови $h(y) = 0$ із (14.196) дістаємо

$$V_{G0} = \frac{2\pi e N_D}{\varepsilon} \frac{h^2}{4}. \quad (14.197)$$

Тоді (14.196) можна переписати у термінах напруг

$$h(y) = h \left(1 - \sqrt{\frac{V_{GS} - V_{DS}(y)}{V_{G0}}} \right). \quad (14.198)$$

Якщо через канал протікає струм I_{DS} , то на елементі довжини каналу dy падає напруга

$$dV_{DS}(y) = I_{DS} dR, \quad (14.199)$$

де $dR = \frac{\rho}{wh} \frac{dy}{h(y)}$, ρ – питомий опір матеріалу. Із врахуванням (14.198)

запишемо величину електричного опору елемента довжини каналу

$$dR = \frac{\rho}{wh} \left(1 - \sqrt{\frac{V_{GS} - V_{DS}(y)}{V_{G0}}} \right)^{-1} dy. \quad (14.200)$$

Питома провідність виражається через рухливість і рівноважну щільність носіїв $\rho^{-1} = e\mu_p p_0$. Тоді маємо

$$\frac{wh}{\rho} = we\mu_p p_0 h = e\mu_p Q_p^{(0)}, \quad (14.201)$$

де $Q_p^{(0)} = ep_0 h$ – заряд дірок у каналі, що приходиться на одиницю площі каналу. Підставляючи (14.198) до (14.200) із врахуванням співвідношення (14.201), отримаємо рівняння

$$I_{DS}(y) dy = e\mu_p Q_p^{(0)} \left(1 - \sqrt{\frac{V_{GS} - V_{DS}}{V_{G0}}} \right) dV_{DS}. \quad (14.202)$$

Для інтегрування цього рівняння зазначимо, що на перерізі $y = 0$, напруга **витік-стік** є нульовою, а на перерізі $y = L$ спостерігається падіння напруги V_{DS} . Тоді інтеграл (14.202) запишемо як

$$\int_0^L I_{DS} dy = e\mu_p Q_p^{(0)} \int_0^{V_{DS}} dV_{DS} \left(1 - \sqrt{\frac{V_{GS} - V_{DS}}{V_{G0}}} \right). \quad (14.203)$$

Інтегрування дає залежність струму стоку I_{DS} від напруги на затворі та витоку

$$I_{DS} = e\mu_p Q_p^{(0)} \frac{1}{L} \left(V_{DS} - \frac{2}{3} \frac{V_{GS}^{3/2} - (V_{GS} - V_{DS})^{3/2}}{V_{G0}^{1/2}} \right). \quad (14.204)$$

Звідси, зокрема видно, що за невеликих значень напруги **витік-стік** в області плавного каналу $V_{DS} \ll V_{GS}$ струм через канал

$$I_{DS} = \frac{w}{L} \mu_p Q_p^{(0)} V_{DS}. \quad (14.205)$$

Як видно із (14.198), за напруг на затворі $V_{GS} < V_{G0}$ завжди можна визначити таке значення напруги на витоку V_{DS}^* , коли поблизу стоку відбудеться змикання каналу, тобто

$$h(y = L, V_{GS}, V_{DS}^*) = 0. \quad (14.206)$$

Ця напруга називається напругою відсічки. Її легко знайти із (14.198)

$$V_{DS}^* = V_{GS} - V_{G0}. \quad (14.207)$$

За зростання напруги **витік-стік** V_{DS} точка відсічки переміщується від витіку до стоку. При цьому спостерігається незалежність струму витіку від напруги на витіку та ефект модуляції довжини каналу. Для визначення струму в області відсічки необхідно підставити (14.207) до (14.204)

$$I_{DS} = e n \mu_p Q_p^{(0)} \frac{1}{L} \left(V_{GS} \left(1 - \frac{2}{3} \sqrt{\frac{V_{GS}}{V_{G0}}} \right) - \frac{1}{3} V_{G0} \right). \quad (14.208)$$

На рис. 14.28 подано вольт-амперні характеристики польового транзистора із затвором у вигляді p - n -переходу. За нульової напруги на затворі канал транзистора відкритий і величина струму через нього є максимальною. Швидкодія польового транзистора із p - n -переходом визначається зарядкою бар'єрних ємностей C_G затворних p - n -переходів через опір каналу R_C . Величина часу заряджання визначається зви-

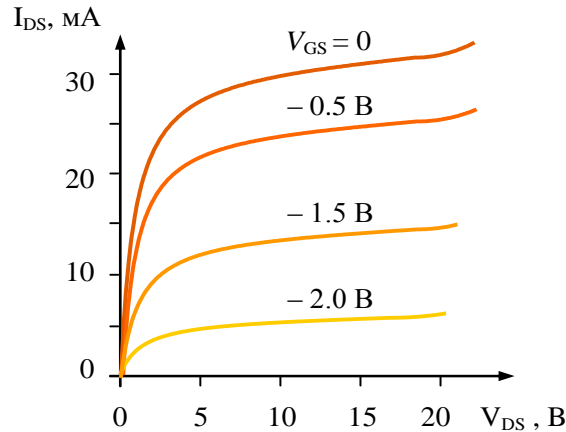


Рис.14.28. Вольт-амперні характеристики польового транзистора з p - n -переходом.

чайно, як $\tau = V_{GS} R_C$. Оскільки ємність затвору та електричний опір каналу визначаються як $C_G = \frac{\epsilon L w}{2\pi \Delta l_f}$ та $R_C = \frac{\rho L}{w(h - 2\Delta l_f)}$, то для часу заряду бар'єрних ємностей маємо

$$\tau = \frac{\epsilon L^2 \rho}{2\pi \Delta l_f (h - 2\Delta l_f)}. \quad (14.209)$$

Цей вираз досягає мінімального значення

$$\tau_{\min} = \frac{4\varepsilon\rho}{\pi} \left(\frac{L}{h}\right)^2 \quad (14.210)$$

за ширини області збіднення $\Delta l_f = h/4$. Таким чином, гранична частота польового транзистора із затвором у вигляді p - n -переходу визначається формулою

$$f_{Cr} = \frac{1}{\tau} = \frac{\pi}{4\varepsilon\rho} \left(\frac{h}{L}\right)^2. \quad (14.211)$$

Звідси видно, що гранична частота підвищується для напівпровідників із низькими значеннями питомого опору та за збільшенням геометричного параметру h/L . Однак цей параметр не має бути більшим за одиницю. Наприклад, за $h = L$ для транзистора, канал якого виготовлено із кремнію, де $\varepsilon = 11,8$ і питомий опір $\rho \approx 1$ Ом·см, гранична частота становить ~ 5 ГГц.

■ 14.10. Задачі

1. Подайте енергетичні діаграми контакту **метал–власний напівпровідник** залежно від співвідношення між роботами виходу металу та напівпровідника. Оцініть зміну провідності приконтактного шару напівпровідника.

2. Визначити хід потенціалу у збагаченому приповерхневому шарі напівпровідника n -типу. Вважаючи всі донори повністю іонізованими, знайти товщину збагаченого шару.

Розв'язок. У даному випадку границя напівпровідника характеризуватиметься вигином зон **уніз**. Це означає, що поверхневий потенціал, відрахований від дна зони провідності в об'ємі, має бути від'ємним. Таким чином, потенціал має задовольняти рівняння Пуассона

$$\frac{d^2\varphi}{dx^2} = \frac{4\pi e}{\varepsilon} (n(x) - n_0). \quad (14.212)$$

Вважаючи, що напівпровідник є невідродженим, перепишемо рівняння у вигляді

$$\frac{d^2\varphi}{dx^2} = \frac{4\pi en_0}{\varepsilon} (e^{e\varphi/kT} - 1). \quad (14.213)$$

При інтегруванні рівняння застосуємо метод, що використовувався при виведенні (14.81). Дістанемо

$$\frac{d}{dx} \left(\frac{d\varphi}{dx}\right)^2 = \frac{d\varphi}{dx} \frac{8\pi en_0}{\varepsilon} (e^{e\varphi/kT} - 1). \quad (14.214)$$

звідки маємо рівність диференціалів

$$d\left(\frac{d\varphi}{dx}\right)^2 = d\varphi \frac{8\pi en_0}{\varepsilon} (e^{e\varphi/kT} - 1). \quad (14.215)$$

Це рівняння зручно інтегрувати, переходячи до нових змінних: $\varphi \rightarrow Y = e\varphi/kT$. Використовуючи граничні умови $Y = 0$, $dY/dx = 0$ при $x \rightarrow \infty$ та $Y = Y_S$ при $x = 0$, маємо

$$\left(\frac{d\varphi}{dx}\right)^2 = 2L_S^{-2}(e^Y - 1 - Y) \rightarrow d\varphi = \sqrt{2}L_S^{-1}(e^{e\varphi/kT} - 1 - e\varphi/kT)^{1/2} dx,$$

звідки

$$\int_{e\varphi/kT}^{e\varphi_S/kT} dt \frac{1}{\sqrt{e^t - t - 1}} = \sqrt{2} \frac{x}{L_S}. \quad (14.216)$$

Якщо товщину шару збагачення L_e визначити із умови $e\varphi(L_e) = kT$, то за умови $e\varphi_S \gg kT$ інтегрування у (14.216) можна проводити у межах від 1 до ∞ . Таким чином інтеграл стає просто числом

$$L_e = \frac{L_S}{\sqrt{2}} c \quad \text{із} \quad c = \int_1^{\infty} \frac{dx}{\sqrt{e^x - x - 1}}.$$

Для уявлення про величину шару збагачення за сильного збагачення оцінимо інтеграл зверху

$$c \leq \int_1^{\infty} e^{-x/2} dx = 2e^{-1/2} \approx 1,63.$$

Бачимо, що товщина шару збагачення приблизно дорівнює дебаєвській довжині екранування в об'ємі.

3. Визначити електронний струм в p - та n -областях p - n -переходу та загальний струм через p - n -перехід при прямому зміщенні в умовах низького рівня інжекції.

Розв'язок. Позначимо щільність струму дірок у p -області p - n -переходу через j_p^p , щільність струму дірок, що інжектуються в n -область, – через j_p^n . Аналогічно, щільність струму електронів в n -області – j_n^n і щільність струму електронів, інжектуються в p -область – j_n^p . В n -області сумарний струм утворюють електрони та інжектуються із p -області дірки

$$j_{(p)} = j_p^n(x) + j_n^n(x). \quad (14.217)$$

Аналогічно, у p -області виникає струм інжектуються електронів, який разом зі струмом основних носіїв (дірок) утворює сумарний струм

$$j_{(n)} = j_n^p(x) + j_p^p(x). \quad (14.218)$$

За інжекції неосновних носіїв відбувається інтенсивна рекомбінація з основними носіями. Якщо рекомбінація у перехідному шарі є малою, то на його границях струми основних носіїв мають занулятися, тобто

$$j_p^p(-x_1) = j_n^n(x_2) = 0. \quad (14.219)$$

У випадку, якщо рекомбінація у перехідному шарі є значною, то сума струмів основних носіїв на границях дорівнюватиме струму рекомбінації усередині перехідного шару

$$j_p^p(-x_1) + j_n^n(x_2) = j_r. \quad (14.220)$$

Струм рекомбінації j_r дорівнює числу рекомбінуючих електронів і дірок в одиницю часу всередині перехідного шару товщиною $d = x_1 + x_2$ та одиничної площі, помноженому на заряд електрона.

Таким чином, повна щільність струму через перехід є сумою струмів $j_{(p)}$ та $j_{(n)}$, із врахуванням (14.218), можна записати у вигляді

$$j = j_p^n(x_2) + J_p^p(-x_1) + j_r. \quad (14.221)$$

Іншими словами, визначення повного струму через p - n -перехід зводиться до обчислення струмів неосновних носіїв на границях переходу і рекомбінаційного струму всередині.

Розглянемо випадок слабкої інжекції. Це означає, що час життя та довжину дифузії неосновних носіїв можна вважати сталими. Вважатимемо також, що довжини p - та n -областей є більшими за довжини дифузії, тому що інжектвані носії устигають рекомбінувати під час руху через прилад. З іншого боку, товщину перехідного шару вважатимемо такою малою, щоб рекомбінаційним струмом у (14.221) можна було знехтувати. Крім того, якщо струми через перехід не дуже великі, то при обчисленні струмів $j_p^p(-x_1)$ та $j_n^n(x_2)$ можна знехтувати дрейфовою складовою, беручи до уваги лише дифузію.

Оцінимо параметри системи, за яких таке припущення є справедливим. Розглянемо розподіл дірок в n -області. Як було показано раніше (див. (11.74)), концентрація надлишкових носіїв визначається формулою

$$\delta p = (p(x_2) - p_n) e^{-(x-x_2)/L_p}, \quad (14.222)$$

де L_p – довжина затягування дірок до n -області. Таким чином, для дифузійної компоненти струму дірок у площині x_2 можна записати

$$j_p^{(diff)}(x_2) = -eD_p \left. \frac{d\delta p}{dx} \right|_{x=x_2}.$$

Застосовуючи найпростішу оцінку для градієнта, формулу можна переписати

$$j_p^{(diff)}(x_2) = \frac{eD_p}{L_p} (p(x_2) - p_n). \quad (14.223)$$

З іншого боку, дрейфовий струм у цій же площині визначається електричним полем у площині $x = x_2$

$$j_p^{(dr)}(x_2) = ep(x_2)\mu_p E. \quad (14.224)$$

Зазвичай концентрація дірок на границі області переходу сильно перевищує їхню рівноважну концентрацію в n -області. Це означає, що другим доданком у дужках правої частини (14.230) можна знехтувати. Тоді відношення дифузійного струму до дрейфового

$$\left. \frac{j_p^{(diff)}}{j_p^{(dr)}} \right|_{x=x_2} \cong \frac{D_p}{L_p \mu_p E}. \quad (14.225)$$

Звідси видно, що дрейфовою складовою струму можна знехтувати, якщо поле на p - n -переході задовольнятиме нерівності

$$E \ll \frac{kT}{eL_p}. \quad (14.226)$$

При запису цього виразу використано співвідношення Ейнштейна. Зазначимо, як впливає із 11.5, в умовах дії не дуже сильного електричного поля довжина затягування L_p наближено дорівнює довжині дифузії.

Для тонких переходів концентрація дірок у перерізі $x = x_2$ пов'язана із напругою на переході ϕ співвідношенням $p(x_2) = p_n e^{e\phi/kT}$. Тоді із (14.223) маємо

$$j_p(x_2) = \frac{eD_p}{L_p} (e^{e\phi/kT} - 1). \quad (14.227)$$

Для електронного струму у перерізі $x = -x_1$ аналогічно дістанемо

$$j_n(-x_1) = \frac{eD_n}{L_n} (e^{e\phi/kT} - 1). \quad (14.228)$$

Тоді сумарна щільність струму через одиницю площини p - n -переходу має вигляд (14.119) $j = j_s (e^{e\phi/kT} - 1)$ зі струмом насичення

$$j_s = \frac{eD_p p_n}{L_p} + \frac{eD_n n_p}{L_n}. \quad (14.229)$$

Список літератури

1. Смит Р. Полупроводники. – М.: Мир, 1982.
2. Зи С.М.. Физика полупроводниковых приборов. – М.: Мир, 1973.
3. Бонч-Бруевич В.Л., Калашников С.Г. Физика полупроводников. – М.: Наука, 1977.
4. Пека Г.П., Стрїха В.І. Поверхневі та контактні явища у напівпровідниках, К.: Либідь, 1992.
5. Kasap S.O., Principles of Electronic Materials and Devices, McGraw-Hill, 2005.

Розділ 15

ОПТИКА НАПІВПРОВІДНИКІВ

Оптичні властивості напівпровідників становлять великий клас явищ, цікавих не тільки з погляду фундаментальної науки, але й щодо практичного застосування. Наприклад, ціла галузь сучасної електроніки – оптоелектроніка – заснована на використанні оптичних властивостей напівпровідників. Усі сучасні системи оптичного запису інформації базуються на використанні напівпровідникових лазерів. Останнім часом інтенсивний розвиток набувають нові принципи запису та обробки інформації, що ґрунтуються на використанні так званих ефектів ближнього поля. При цьому як активні матеріали у таких системах використовуються напівпровідники. Таким чином, вивчення оптичних властивостей є важливим елементом сучасної фізики напівпровідників.

■ 15.1. Поглинання світла напівпровідниками

При вивченні поглинання світла вимірюють зменшення інтенсивності світла у процесі проходження світлом визначеної області зразка. Розглянемо зразок напівпровідника, одна із поверхонь якого є плоскою (рис. 15.1). Нехай інтенсивність світла на цій поверхні – I . Завдяки поглинанню світла матеріалом інтенсивність випромінювання при проходженні шару dx речовини зменшується на величину dI . Кількість енергії, що поглинута речовиною, пропорційна кількості енергії, що падає на цей шар і товщині шару

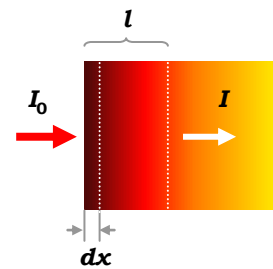


Рис. 15.1. Поглинання світла шаром речовини

$$-dI = \alpha I dx. \quad (15.1)$$

Проінтегруємо рівняння від 0 до l і знайдемо енергію, що поглинув в одиницю часу зразок напівпровідника шаром товщиною l ,

$$I = I_0 e^{-\alpha l}. \quad (15.2)$$

Формула (15.2) виражає так званий закон Бугера. Коефіцієнт α , що має фізичний сенс енергії, яка поглинається речовиною із пучка одиничної інтенсивності випромінювання за одиницю часу у шарі одиничної товщини, називається коефіцієнтом поглинання. Закон Бугера справедливий при поглинанні не дуже інтенсивних пучків випроміню-

вання, коли нелінійні процеси не виявляються. Коефіцієнт поглинання залежить від довжини хвилі світла та поглинаючої речовини. Геометрія експерименту, подана на рис. 15.2, майже ніколи не реалізується, оскільки зазвичай вимірювання поглинання відбувається на тонких плівках, і за взаємодії світла із передньою та задньою стінками плівки відбувається відбивання світла. Таким чином, виникає проблема акуратного визначення інтенсивності світла, що поглинається тонкою плівкою. Для розв'язання цієї проблеми розглянемо тонку плівку, що розташована в однорідному ізотропному середовищі. Вважатимемо, що за відбиття світла від межі розподілу **середовище-плівка** його інтенсивність змінюється в R разів, тобто, якщо I_0 – інтенсивність падаючого на

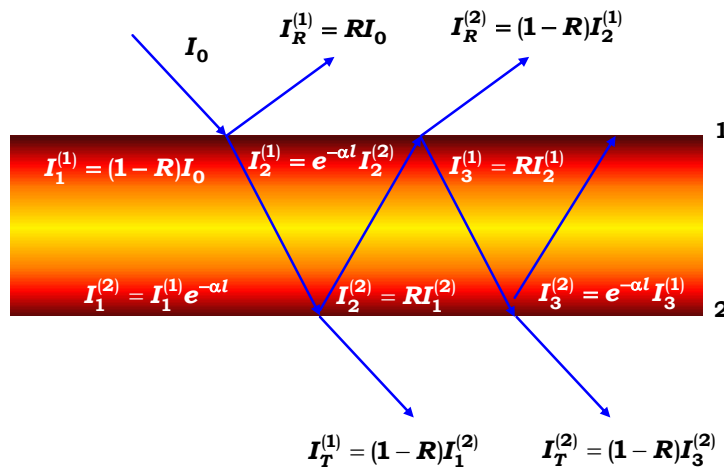


Рис. 15.2. Інтенсивності світла за багаторазового внутрішнього відбиття у поглинаючій плівці

плівку світла, а I_R – інтенсивність відбитого від межі розподілу випромінювання, то

$$R = I_R / I_0. \tag{15.3}$$

Величина R називається коефіцієнтом відбиття. Нехай на плівку падає випромінювання інтенсивністю I_0 . Тоді до плівки проникне випромінювання з інтенсивністю $I_1^{(1)} = (1-R)I_0$. До другої межі розподілу плівки це світло за рахунок поглинання прийде, згідно із законом Бугера, ослабленим в $e^{-\alpha l}$ разів. Його інтенсивність $I_2^{(1)} = (1-R)e^{-\alpha l}I_0$. Частина його $I_2^{(2)} = R(1-R)e^{-\alpha l}I_0$ відіб'ється у напрямку першої межі розподілу, а частина інтенсивністю $I_T^{(1)} = (1-R)I_2^{(1)}$ пройде до зовніш-

нього середовища. Відбите світло дійде до першої межі розподілу ослабленим в $e^{-\alpha l}$ разів і характеризуватиметься інтенсивністю $I_2^{(1)} = I_2^{(2)} e^{-\alpha l}$ (див. рис. 15.2); далі частина його пройде до зовнішнього середовища через першу межу розподілу, інша частина відіб'ється у напрямку межі розподілу 2. Таким чином, інтенсивність світла, що пройде півку після нескінченної кількості перевідбиттів від меж розподілу, матиме вигляд ряду

$$I_R^{(F)} = (1 - R)^2 e^{-\alpha l} I_0 + R^2 (1 - R)^2 e^{-3\alpha l} I_0 + R^4 (1 - R)^2 e^{-5\alpha l} I_0 + \dots \quad (15.4)$$

Цю формулу можна переписати у вигляді

$$I_R^{(F)} = (1 - R)^2 e^{-\alpha l} \left\{ 1 + R^2 e^{-2\alpha l} + R^4 e^{-4\alpha l} + \dots \right\} I_0. \quad (15.5)$$

У фігурних дужках стоїть нескінченний ряд геометричної прогресії із першим членом 1 і знаменником $R^2 e^{-2\alpha l}$. Таким чином, інтенсивність світла, що пройшло через поглинаючу півку,

$$I_R^{(F)} = \frac{(1 - R)^2 e^{-\alpha l}}{1 - R^2 e^{-2\alpha l}} I_0. \quad (15.6)$$

Формула (15.6) показує, що поглинання тонкої півки визначається коефіцієнтом поглинання матеріалу α і характеристикою межі розподілу – коефіцієнтом відбиття.

■ 15.2. Коефіцієнт поглинання світла

Раніше ми з'ясували, що при вивченні процесу поглинання світла напівпровідниками необхідно враховувати можливість перевипромінювання світла різними поверхнями зразка. Аналіз поглинання світла напівпровідниковою півкою показав, що важливу роль при цьому відіграє важлива характеристика матеріалу – коефіцієнт поглинання світла α , що має фізичний сенс енергії, яка поглинається речовиною із пучка одиничної інтенсивності випромінювання за одиницю часу у шарі одиничної товщини. Виходячи із простих квантово-механічних міркувань, обчислимо коефіцієнт поглинання та обговоримо особливості поглинання світла напівпровідником на краю смуги поглинання. Оптичні властивості напівпровідників базуються на можливості поглинання та емісії фотонів напівпровідниками. При цьому відбуваються три типи процесів (рис. 15.3). Перший процес – поглинання фотона із генерацією електрон-діркової пари. Енергія фотона, що поглинається системою, переходить в енергію збудженого

електрона, який переходить із валентної зони до зони провідності (рис. 15.3 а). Зворотній процес – це емісія фотона внаслідок рекомбінації електрон-діркової пари, коли в результаті фотопереходу електрон переходить із зони провідності до валентної зони. При цьому можуть відбуватись спонтанна та фотостимульована емісія (рис. 15.3 б, в).

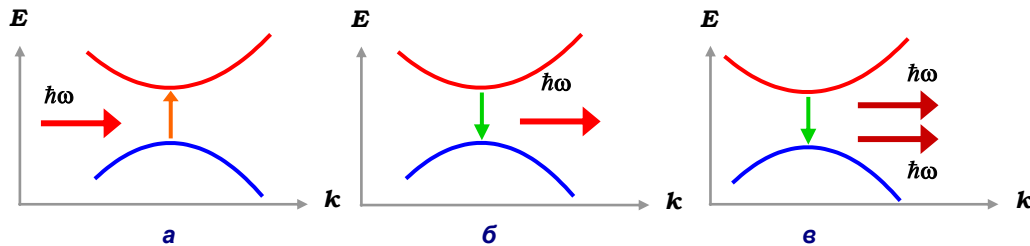


Рис. 15.3. Поглинання світла (а) та фотоемісія (б, в) у напівпровідниках

Розглянемо процес поглинання фотона зовнішнього випромінювання електронним газом. Якщо при поглинанні напівпровідником кванта світла відбувається збудження електрона із валентної зони до зони провідності, таке поглинання називають *фундаментальним* або *власним*. При вивченні особливостей власного поглинання напівпровідника зрозумілим стає значення будови його енергетичних зон. З погляду процесів поглинання світла напівпровідники можна поділити на два основних класи. Перший із них характеризується тим, що мінімум енергії у зоні провідності, що відповідає хвильовому вектору \mathbf{k}_{\min} і максимум енергії у валентній зоні, що визначається хвильовим вектором \mathbf{k}_{\max} , розташовані в одній точці зони Бриллюена, тобто, $\mathbf{k}_{\min} = \mathbf{k}_{\max}$. Зазвичай це точка $\mathbf{k} = 0$. Як приклад таких напівпровідників можна навести антимонід індію. Інший клас напівпровідникових речовин характеризується тим, що екстремуми зони провідності та валентної зони розташовані у різних точках зони Брилюена, тобто у таких напівпровідниках де $\mathbf{k}_{\min} \neq \mathbf{k}_{\max}$. До цього класу належить більшість напівпровідників, наприклад германій і кремній. Зрозуміло, що переходи електронів через заборонену зону відбуватимуться спершу між енергетичними станами, що відповідають максимуму валентної зони та мінімуму зони провідності, тобто для прямозонних напівпровідників, за значень хвильового вектора $\mathbf{k} \approx 0$. Розглянемо детальніше процес поглинання світла у напівпровіднику, користуючись простими квантово-механічними уявленнями.

Якщо електрон перебуває у полі зовнішнього випромінювання, то його квантовий рух описується рівнянням Шредингера

$$\left[\frac{(\mathbf{p} + e\mathbf{A})^2}{2m} + V_{cryst} \right] \Psi = i\hbar \frac{\partial \Psi}{\partial t}, \quad (15.7)$$

де V_{cryst} – кристалічний потенціал. Збуренням, що вносить до системи зовнішнє світло, є

$$\hat{V} = \frac{e}{2m} (\mathbf{A} \cdot \mathbf{p} + \mathbf{p} \cdot \mathbf{A} + e\mathbf{A}^2). \quad (15.8)$$

Обмежуючись лінійними оптичними процесами, вважатимемо поле слабким. Це означає, що останнім доданком у (15.8) можна знехтувати. Перший доданок у (15.8) не викликає сумнівів. Пам'ятаємо, що $\mathbf{p} = -i\hbar\nabla$, тобто є векторним оператором диференціювання, тоді

$$\mathbf{A} \cdot \mathbf{p}\Psi = -i\hbar\mathbf{A} \cdot (\nabla\Psi). \quad (15.9)$$

Обчислимо дію другого доданку в операторі (15.8) на хвильову функцію

$$\mathbf{p} \cdot \mathbf{A}\Psi = -i\hbar\nabla \cdot (\mathbf{A}\Psi) = -i\hbar[(\nabla \cdot \mathbf{A})\Psi + \mathbf{A} \cdot (\nabla\Psi)]. \quad (15.10)$$

Оскільки для поперечної хвилі $\nabla \cdot \mathbf{A} = 0$, то згідно із (15.9) та (15.10), маємо, що оператор збурення (15.8) зводиться до

$$\hat{V} = 2\mathbf{A} \cdot \mathbf{p}. \quad (15.11)$$

Якщо на електрон діє поле амплітудою E_0 із хвильовим вектором \mathbf{Q} і частотою ω , то вектор-потенціал запишеться як

$$\mathbf{A} = (eE_0 / i\omega) (e^{i(\mathbf{Q}\mathbf{R} - \omega t)} - e^{-i(\mathbf{Q}\mathbf{R} - \omega t)}). \quad (15.12)$$

Підставляючи цей вираз до (15.11), отримаємо потенціал збурення у вигляді

$$\hat{V}(\mathbf{R}, t) = (eE_0 / i\omega) (e^{i(\mathbf{Q}\mathbf{R} - \omega t)} - e^{-i(\mathbf{Q}\mathbf{R} - \omega t)}) (\mathbf{e} \cdot \mathbf{p}), \quad (15.13)$$

де \mathbf{e} – орт поляризації зовнішнього випромінювання. Типові значення довжин хвиль оптичного діапазону становлять кілька сотень нанометрів, що набагато більше за характерні розміри кристалічної елементарної комірки напівпровідника, яка є характерним мірилом просторової зміни електронної хвильової функції. Отже ми можемо знехтувати імпульсом фотона у (15.13). Це є так зване електродипольне наближення. Таким чином, темп переходу електронів зі стану i до стану j за збудження напівпровідника зовнішнім випромінюванням, задається правилом Фермі

$$W_{ji} = \frac{2\pi}{\hbar} \left(\frac{eE_0}{m\omega} \right)^2 |\langle j | \mathbf{e} \cdot \mathbf{p} | i \rangle|^2 \delta(E_j - E_i - \hbar\omega). \quad (15.14)$$

Енергія, що поглинається завдяки такому переходу, є добутком цього темпу переходу та енергії фотона, що поглинається $\hbar\omega$. Крім того, ми

маємо домножити цю величину на Фермі-фактор $F(E_i)$, що визначає заповнення вихідного стану i , та $[1 - F(E_j)]$, що відповідає умові, за якої кінцевий стан j незаповнений. Таким чином, електронна система поглинає в одиницю часу енергію

$$P_+ = 2 \frac{2\pi}{\hbar} \hbar\omega \sum_{i,j} \left(\frac{eE_0}{m\omega} \right)^2 |\langle j | \mathbf{e} \cdot \mathbf{p} | i \rangle|^2 F(E_i) [1 - F(E_j)] \delta(E_j - E_i - \hbar\omega). \quad (15.15)$$

Фактор 2 перед знаком суми відповідає подвійному виродженню за спіном при переходах, коли спін зберігається. Крім поглинання фотона, можуть відбуватись процеси емісії (оскільки поле має як доданок з експонентою $e^{-i\omega t}$ так і $e^{i\omega t}$). Цей внесок має вигляд, аналогічний до P_+ , але із частотою $-\omega$

$$P_- = -2 \frac{2\pi}{\hbar} \hbar\omega \sum_{i,j} \left(\frac{eE_0}{m\omega} \right)^2 |\langle j | \mathbf{e} \cdot \mathbf{p} | i \rangle|^2 F(E_i) [1 - F(E_j)] \delta(E_j - E_i + \hbar\omega).$$

Цей вираз можна привести до вигляду, аналогічного (15.15). Для цього необхідно замінити індекси підсумовування, і, таким чином, аргумент у δ -функції привести до такого самого вигляду, як і в (15.15)

$$P_- = -2 \frac{2\pi}{\hbar} \hbar\omega \sum_{i,j} \left(\frac{eE_0}{m\omega} \right)^2 |\langle j | \mathbf{e} \cdot \mathbf{p} | i \rangle|^2 F(E_j) [1 - F(E_i)] \delta(E_j - E_i - \hbar\omega). \quad (15.16)$$

Ми використали, що $|\langle i | \mathbf{e} \cdot \mathbf{p} | j \rangle|^2 = |\langle j | \mathbf{e} \cdot \mathbf{p} | i \rangle|^2$. Таким чином, повний темп зміни енергії при поглинанні світла дорівнюватиме сумі $P = P_+ - P_-$. Беручи до уваги рівність

$$F(E_i) [1 - F(E_j)] - F(E_j) [1 - F(E_i)] = F(E_i) - F(E_j)$$

і використовуючи (15.15) та (15.16), можна записати цю суму у вигляді

$$P = \frac{4\pi}{\hbar} \hbar\omega \left(\frac{eE_0}{m\omega} \right)^2 \sum_{i,j} |\langle j | \mathbf{e} \cdot \mathbf{p} | i \rangle|^2 [F(E_i) - F(E_j)] \delta(E_j - E_i - \hbar\omega). \quad (15.17)$$

Згадаємо визначення коефіцієнта поглинання світла α як енергію зовнішнього випромінювання одиничної амплітуди, що поглинається одиницею об'єму системи в одиницю часу. Таким чином, величина

$$\alpha = 4\pi\omega \sum_{i,j} |\langle j | \mathbf{e} \cdot \mathbf{D} | i \rangle|^2 [F(E_i) - F(E_j)] \delta(E_j - E_i - \hbar\omega) \quad (15.18)$$

і визначає коефіцієнт поглинання світла. При запису (15.18) було використано, що $\mathbf{p} = -i\omega m \mathbf{r}$ і принцип відповідності. Крім того, введено оператор дипольного моменту $\mathbf{D} = e\mathbf{r}$.

■ 15.3. Механізми поглинання світла напівпровідниками

Процеси поглинання світла напівпровідниками можна класифікувати за характером витрати енергії поглинутих фотонів. На цій підставі можна виділити такі механізми поглинання світла об'ємом напівпровідника:

- ♦ **міжзонне поглинання**, за якого енергія фотона витрачається на генерацію електронно-діркової пари. При цьому поглинання відбувається суттєво різними шляхами у прямозонних і непрямозонних напівпровідниках;

- ♦ **екситонне поглинання**, за якого енергія поглинутого фотону витрачається на створення специфічного збудженого стану у напівпровіднику – екситону;

- ♦ **поглинання вільними носіями заряду**, за якого енергія поглинутого фотона витрачається на утворення високочастотних струмів у системі носіїв заряду у зонах – електронів у зоні провідності та дірок – у валентній зоні;

- ♦ **домішкове поглинання**, за якого енергія фотона витрачається на іонізацію або збудження домішкових центрів;

- ♦ **поглинання ґраткою**, за якого відбувається поглинання фотонів зі збудженнями фононної підсистеми (ґратки).

Розглянемо детально ці механізми поглинання світла напівпровідником.

15.3.1. Власне поглинання за прямих переходів

За міжзонного поглинання у напівпровіднику у формулі (15.18) за енергію початкового стану E_i маємо вважати енергію валентної зони $E_V(\mathbf{k}_v)$, а за енергію кінцевого стану E_j – енергію зони провідності $E_C(\mathbf{k}_c)$. Крім того, замість матричних елементів $\langle i | \mathbf{e} \cdot \mathbf{D} | j \rangle$ необхідно записати матричні елементи $\langle v, \mathbf{k}_v | \mathbf{e} \cdot \mathbf{D} | c, \mathbf{k}_c \rangle$, які визначаються на хвильових функціях електрона в кристалі

$$\Psi_{v, \mathbf{k}_v} = (1/\sqrt{V})u_{v, \mathbf{k}_{\max}}(\mathbf{r})e^{i \mathbf{k}_v \mathbf{r}}, \quad \Psi_{c, \mathbf{k}_c} = (1/\sqrt{V})u_{c, \mathbf{k}_{\min}}(\mathbf{r})e^{i \mathbf{k}_c \mathbf{r}},$$

де $u_{v, \mathbf{k}_{\max}}$ та $u_{c, \mathbf{k}_{\min}}$ – блохівські функції електронів із валентної і зони провідності, відповідно. Для прямозонних напівпровідників максимум валентної зони і мінімум зони провідності відповідають одному й тому

самому хвильовому вектору \mathbf{k}_0 . Таким чином, для матричного елемента запишемо

$$\begin{aligned} \langle v, \mathbf{k}_v | \mathbf{e} \cdot \mathbf{D} | c, \mathbf{k}_c \rangle &= (1/V) \int d\mathbf{r} u_{v, \mathbf{k}_0}^* e^{-i\mathbf{k}_v \cdot \mathbf{r}} (\mathbf{e} \cdot \mathbf{D}) u_{c, \mathbf{k}_0} e^{i\mathbf{k}_c \cdot \mathbf{r}} = \\ &= (1/V) \sum_{\mathbf{R}_i} e^{-i(\mathbf{k}_v - \mathbf{k}_c) \cdot \mathbf{R}_i} \int_{\Omega} d\mathbf{r} u_{v, \mathbf{k}_0}^* (\mathbf{e} \cdot \mathbf{D}) u_{c, \mathbf{k}_0}, \end{aligned} \quad (15.19)$$

де інтегрування ведеться за об'ємом елементарної комірки. Дуже важливим моментом при цьому є той факт, що ці інтеграли не залежать від номера комірки

$$\int_{\Omega} d\mathbf{r} u_{v, \mathbf{k}_v}^* (\mathbf{e} \cdot \mathbf{D}) u_{c, \mathbf{k}_c} = \Omega \cdot (\mathbf{e} \cdot \mathbf{D}_{vc}), \quad (15.20)$$

де \mathbf{D}_{vc} – матричний елемент оператора дипольного моменту на блохівських функціях. Застосуємо тепер правило підсумовування за комірками кристала

$$\frac{\Omega}{V} \sum_{\mathbf{R}_i} e^{-i(\mathbf{k}_v - \mathbf{k}_c) \cdot \mathbf{R}_i} = \delta_{\mathbf{k}_v, \mathbf{k}_c}. \quad (15.21)$$

Використовуючи (15.20) та (15.21), для матричного елемента (15.19) маємо

$$|\langle v, \mathbf{k}_v | \mathbf{e} \cdot \mathbf{D} | c, \mathbf{k}_c \rangle|^2 = |(\mathbf{e} \cdot \mathbf{D}_{cv})|^2 \delta_{\mathbf{k}_v, \mathbf{k}_c}. \quad (15.22)$$

У граничному випадку $V \rightarrow \infty$ можна зробити заміну $V \delta_{\mathbf{k}_v, \mathbf{k}_c} \rightarrow \delta(\mathbf{k}_c - \mathbf{k}_v)$.

Із (15.22), зокрема випливає, що фотопереходи відбуваються при задоволенні правил відбору

$$\mathbf{k}_v = \mathbf{k}_c. \quad (15.23)$$

Беручи до уваги (15.22), запишемо для коефіцієнта поглинання (15.18)

$$\begin{aligned} \alpha &= 4\pi\omega \left| (\mathbf{e} \cdot \mathbf{D}_{c,v}) \right|^2 \sum_{\mathbf{k}_v, \mathbf{k}_c} \delta_{\mathbf{k}_v, \mathbf{k}_c} [F(E_c(\mathbf{k}_c)) - F(E_v(\mathbf{k}_v))] \delta(E_c(\mathbf{k}_c) - E_v(\mathbf{k}_v) - \hbar\omega) = \\ &= 4\pi\omega \left| (\mathbf{e} \cdot \mathbf{D}_{c,v}) \right|^2 \sum_{\mathbf{k}} [F(E_c(\mathbf{k})) - F(E_v(\mathbf{k}))] \delta(E_c(\mathbf{k}) - E_v(\mathbf{k}) - \hbar\omega) \end{aligned}$$

Зазначимо, що правила відбору (15.23) відображають закон збереження імпульсу, а наявність дельта-функції під знаком суми – закон збереження енергії за процесів поглинання світла напівпровідником. Для подальшого аналізу коефіцієнта поглинання перейдемо від підсумовування за хвильовими векторами до інтегрування згідно із правилом

$$\sum_{\mathbf{k}} \dots \rightarrow \frac{V}{(2\pi)^3} \int d\mathbf{k} \dots$$

Маємо

$$\alpha = 4\omega |(\mathbf{e} \cdot \mathbf{D}_{c,v})|^2 \frac{V}{(2\pi)^2} \int d\mathbf{k} [F(E_c(\mathbf{k})) - F(E_v(\mathbf{k}))] \delta(E_c(\mathbf{k}) - E_v(\mathbf{k}) - \hbar\omega). \quad (15.24)$$

Запишемо закон збереження енергії за фотопереходу (тільки в цьому випадку інтеграл в правій частині (15.24) не є нулем)

$$E_c - E_v = \hbar\omega. \quad (15.25)$$

Беручи до уваги (15.23), запишемо $k_v = k_c = k$. Оскільки розглядаються прямі переходи, **то можна записати**

$$E_c - E_v = E_c^0 + \frac{\hbar^2 k^2}{2m_e} - \left(E_v^0 - \frac{\hbar^2 k^2}{2m_h} \right) = E_g + \frac{\hbar^2 k^2}{2} \left(\frac{1}{m_e} + \frac{1}{m_h} \right). \quad (15.26)$$

Підставляючи (15.26) до формули закону збереження (15.25) і вводячи зведену масу

$$m_r = \frac{m_e m_h}{m_e + m_h}, \quad (15.27)$$

отримаємо

$$k = \sqrt{2m_r(\hbar\omega - E_g)}. \quad (15.28)$$

Таким чином, можна вважати, що аргументи у функціях розподілу електронів у валентній зоні та зоні провідності залежать від частоти, а не від хвильового вектора. Тоді можна замість (15.24) записати

$$\alpha = 2\omega |(\mathbf{e} \cdot \mathbf{D}_{c,v})|^2 [F(E_c(\omega)) - F(E_v(\omega))] \frac{V}{(2\pi)^2} \int d\mathbf{k} \delta(E_c(\mathbf{k}) - E_v(\mathbf{k}) - \hbar\omega). \quad (15.29)$$

Введемо тепер ефективну (міжзонну) щільність станів, що визначає кількість станів електрон-діркової системи, яка взаємодіє із фотонами у частотному діапазоні $[\omega, \omega + d\omega]$

$$\rho(\omega) = \frac{V}{(2\pi)^2} \int d\mathbf{k} \delta(E_c(\mathbf{k}) - E_v(\mathbf{k}) - \hbar\omega). \quad (15.30)$$

Неважко побачити із точністю до перепозначень, що ця щільність станів є точно такою самою, яку ми вводили у розд. 8 (див. 8.2). Дійсно, зважаючи на те, що аргумент дельта-функції дорівнює нулю лише у випадку виконання закону збереження енергії

$$\hbar\omega = E_g + \frac{\hbar^2 k^2}{2m_e^*} + \frac{\hbar^2 k^2}{2m_h^*} = E_g + \frac{\hbar^2 k^2}{2m_r},$$

де $m_r = (1/m_e^* + 1/m_h^*)^{-1}$, отримаємо, що міжзонна щільність станів у випадку квадратичного закону дисперсії

$$\rho(\omega) = \frac{(2m_r)^{3/2}}{2\pi^2\hbar^2} (\hbar\omega - E_g)^{1/2}, \quad \hbar\omega > E_g. \quad (15.31)$$

У результаті маємо, що коефіцієнт поглинання світла за його взаємодії із напівпровідником

$$\alpha = 4\omega |(\mathbf{e} \cdot \mathbf{D}_{c,v})|^2 \rho(\omega) [F(E_c(k)) - F(E_v(k))], \quad (15.32)$$

із хвильовим вектором k , що визначається (15.28). Підставляючи сюди щільність станів (15.31), отримуємо частотну залежність коефіцієнта поглинання

$$\alpha(\omega) = P \cdot \omega \cdot (\hbar\omega - E_g)^{1/2} [F(E_c) - F(E_v)], \quad (15.33)$$

де P – матеріальна константа, що визначається через матричний елемент оператора дипольного моменту, а енергії E_c та E_v задовольняють закону збереження (15.25). Таким чином, для прямих міжзонних переходів поведінка коефіцієнту поглинання на краю смуги поглинання, а саме в області $\omega \geq E_g/\hbar$, визначається щільністю станів поблизу E_g .

Поглинання світла із частотою, меншою за ширину забороненої зони, не відбувається. А край смуги поглинання із низькочастотного боку має бути різким. Зрозуміло, що у реальних ситуаціях через наявність

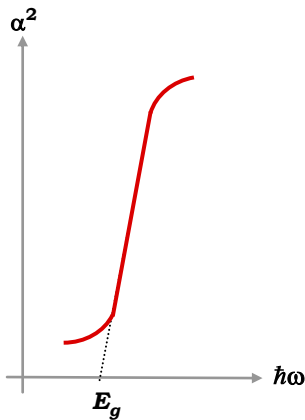


Рис. 15.4. Залежність квадрата коефіцієнта поглинання від енергії фотона для прямих міжзонних переходів

дефектів цей край буде розмитий, і графік залежності квадрата коефіцієнта поглинання від частоти падаючого випромінювання матиме вигляд, що поданий на рис. 15.4. Оскільки залежність $\alpha^2(\omega)$ буде лінійною, подовження цієї прямої до її перетину із віссю $\hbar\omega$ дозволяє визначити ширину забороненої зони для прямих переходів. При обчисленні коефіцієнта поглинання для реальних напівпровідників варто брати до уваги переходи між всіма парами енергетичних зон, в яких наявні стани, розділені енергією $\hbar\omega$. Наприклад, у напівпровідниках зі структурою цинкової обманки (GaAs) для енергій фотона, що незначно перевищують ширину забороненої зони, необхідно враховувати чотириразове виродження валентної зони при $\mathbf{k} = 0$, проводячи підсумовування за зонами легких і важких дірок.

15.3.2. Власне поглинання за непрямих переходів

Як згадувалось вище, більшість напівпровідників характеризується тим, що максимальне та мінімальне значення валентної зони та зони провідності рознесені у зоні Брилюена на деякий вектор \mathbf{q} . Зрозуміло, що тепер для виконання умов збудження електрона із валентної зони до зони провідності з енергіями $\hbar\omega$, що незначно перевищують ширину забороненої зони E_g , набагато складніше задовольнити закон збереження імпульсу. Для цього у процесі поглинання має бути додатково задіяний інший, крім фотона та електрона, агент, яким у кристалах зазвичай виступають фонони. Отже непрямі (невертикальні) переходи відбуваються із поглинанням або висиланням фононів. Такі переходи визначають поглинання, розташоване із довгохвильового боку від границі власного поглинання за прямих переходів. У напівпровідниках із непрямыми переходами при обчисленні коефіцієнта поглинання для врахування взаємодії електронів як з фотонами, так і з фононами, необхідно застосовувати теорію збурень другого порядку. Як наслідок, повний розгляд ймовірності переходу виявляється набагато складнішим за той, що ми проводили раніше. Тому обмежимося коротким викладенням ідеї такого розгляду і зупинимось на якісному характері ефектів поглинання.

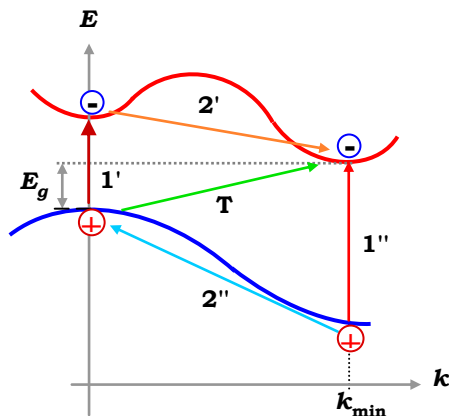


Рис. 15.5. Власне поглинання світла за непрямих переходів

Зрозуміло, що для сумарного процесу закон збереження енергії має виконуватись. Припустимо, що процес переходу електрона із максимуму валентної до мінімуму зони провідності відбувається із поглинанням або висиланням одного фонону. Такий процес може відбуватись двома шляхами.

У напівпровідниках із непрямыми долинами можливі непрямі переходи із будь-якого зайнятого стану валентної зони до будь-якого вільного стану зони провідності. Головною умовою такого процесу є виконання законів збереження енергії та імпульсу. Розглядаючи поглинання за непрямих переходів, вважатимемо, що перехід (T на рис. 15.5) відбуватиметься через ряд проміжних станів, які можна вважати віртуальними, тобто такими, що мають нескінченно малий час життя. Це припущення дозволяє нехтувати законом збереження енергії на кожному із проміжних етапів. Але, зрозуміло, що для сумарного процесу закон збереження енергії має виконуватись.

Спочатку електрон, що перебуває у стані $k = 0$ поблизу стелі валентної зони, завдяки збудженню світлом, переходить до зони провідності у стан, що характеризується тим самим хвильовим вектором (перехід 1'). У валентній зоні, завдяки такому переходу, з'являється дірка. Електрон, що перебуває у стані $k = 0$ у зоні провідності, має більшу енергію, ніж енергія дна зони провідності. Тому він за дуже малий проміжок часу перейде у стан, що відповідає мінімуму зони провідності (перехід 2' до стану k_{\min}), поглинаючи чи випромінюючи при цьому фонон.

Інший шлях починається з того, що електрон із глибокого рівня у валентній зоні вертикально переходить до стану k_{\min} зони провідності (перехід 1''). Глибоко розташована дірка із валентної зони при цьому перейде до більш енергетично вигідного стану валентної зони $k = 0$ (перехід 2''). Припустимо, що початкова енергія електрона у валентній зоні – E , а енергія електрона у кінцевому стані – E' . Енергія фонона, який бере участь у додатковому розсіюванні електрона за оптичного поглинання, позначимо через E_{ph} . Тоді у випадку переходу із поглинанням фонона закон збереження енергії вимагає умови

$$\hbar\omega = E' - E - E_{ph}. \quad (15.34)$$

У випадку переходу із випромінюванням фонону закон збереження енергії має вигляд

$$\hbar\omega = E' - E + E_{ph}. \quad (15.35)$$

Аналогічно тому, як визначався коефіцієнт поглинання за вертикальних переходів, можна вважати, що коефіцієнт поглинання світла визначається як інтеграл за всіма можливими парами станів, які розділені енергетичним зазором

$$E' - E = \hbar\omega \pm E_{ph} \quad (15.36)$$

від ймовірності переходу. Ця ймовірність пропорційна добутку матричного елемента на дельта-функцію, що визначає закон збереження енергії за поглинання. При цьому, на відміну від попереднього випадку (див. (15.22)), у першому наближенні цей матричний елемент можна вважати незалежним від \mathbf{k} і \mathbf{k}' , тобто, якщо відлік енергії вести від дна зони провідності, то

$$\alpha = A \int d\mathbf{k} \int d\mathbf{k}' \delta(E - E' - \hbar\omega \pm E_{ph}), \quad (15.37)$$

де A – стала. У наближенні сферичних зон $d^3k \sim E^{1/2}dE$, а $d^3k' \sim (-E_g - E')^{1/2}dE'$. Тоді

$$\alpha \sim A \int dE(E)^{1/2} \int dE'(-E_g - E') \cdot \delta(E - E' - \hbar\omega \pm E_{ph}). \quad (15.38)$$

Звідси, знімаючи інтегрування за допомогою дельта-функції, дістанемо

$$\alpha \sim A \int_0^{\hbar\omega \mp E_{ph} - E_g} dE(E)^{1/2}(-E_g - E + \hbar\omega \mp E_{ph})^{1/2}. \quad (15.39)$$

Для обчислення цього інтеграла перейдемо до нової змінної $x = E / E_{\mp}$, де $E_{\mp} = \hbar\omega \mp E_{ph} - E_g$

$$\alpha \sim A(E_{\mp})^2 \int_0^1 dx(x)^{1/2}(E_{\mp} - x)^{1/2}. \quad (15.40)$$

Інтеграл є бета-функцією від аргументів $3/2, 3/2$, яку можна виразити через гамма-функції $B(3/2, 3/2) = [\Gamma(3/2)\Gamma(3/2)]/\Gamma(3) = \pi/8$. Тоді

$$\alpha \sim A \frac{\pi}{8} (\hbar\omega \mp E_{ph} - E_g)^2. \quad (15.41)$$

Процес поглинання світла із поглинанням фонона пропорційний функції розподілу фононів N_q , а процес, що проходить із випромінюванням фонона – пропорційний $N_q + 1$. Таким чином, за частоти зовнішнього випромінювання, що задовольняє умові $\hbar\omega > E_g - E_{ph}$, поглинання світла за непрямих переходів відбувається із додатковим поглинанням фонона, і коефіцієнтом поглинання є

$$\alpha \sim N_q(\hbar\omega + E_{ph} - E_g)^2. \quad (15.42)$$

Коефіцієнт поглинання світла із висиланням фонона

$$\alpha \sim (N_q + 1)(\hbar\omega - E_{ph} - E_g)^2. \quad (15.43)$$

Такий процес може відбуватись, якщо частота світла задовольняє умову $\hbar\omega > E_g + E_{ph}$. Більш того, у випадку $\hbar\omega > E_g + E_{ph}$ поглинання світла може відбуватись як із поглинанням, так й із висиланням фонона, тому для коефіцієнта поглинання маємо

$$\alpha \sim N_q(\hbar\omega + E_{ph} - E_g)^2 + (N_q + 1)(\hbar\omega - E_{ph} - E_g)^2. \quad (15.44)$$

Всюди у цих формулах

$$N_q = \frac{1}{e^{E_{ph}/kT} - 1}. \quad (15.45)$$

Із (15.42–15.44), зокрема впливає, що на відміну від прямих переходів, де слабка залежність від температури зумовлена залежністю $E_g(T)$, непрямі переходи завдяки участі фононів демонструють додаткову експоненціальну залежність від температури завдяки залежності $N_q(T)$. За низьких температур при значеннях коефіцієнтів поглинання

$\sim 10 \text{ см}^{-1}$ у спектрі поглинання спостерігається структура, пов'язана із розсіюванням на повздовжніх і поперечних оптичних фононах, а також із міждолинними переходами із додатковим розсіюванням на акустичних фононах. Як випливає із (15.42) та (15.43), графік залежності $\alpha^{1/2}$ від $\hbar\omega$ має дві прямолінійних ділянки (рис. 15.6). Пряма 1 характеризує поглинання світла із висиланням фонона, і якщо її подовжити до перетину із віссю $\hbar\omega$, перетинає вісь у точці $\hbar\omega = E_g + E_{ph}$. Пряма 2, що характеризує поглинання світла із додатковим поглинанням фонона, перетинає вісь $\hbar\omega$ у точці $\hbar\omega = E_g - E_p$. Довжина відрізка між

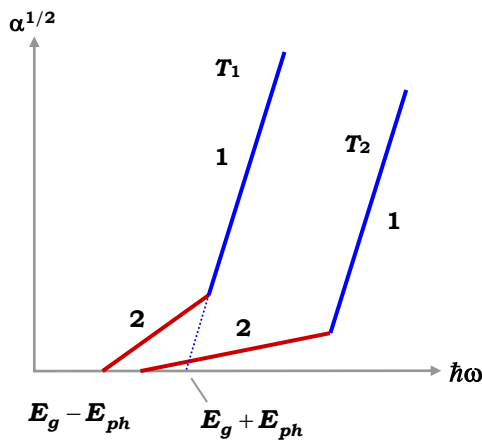


Рис. 15.6. Залежність $\alpha^{1/2}$ від частоти зовнішнього випромінювання за різних температура ($T_1 > T_2$)

цими точками дорівнює $2E_{ph}$. Зі зниженням температури графік залежності $\alpha^{1/2}$ зсувається у бік більших значень $\hbar\omega$, що відбиває температурну залежність ширини забороненої зони, а нахил прямої 2 прямує до нуля, оскільки за низьких температур збуджується дуже мала кількість фононів, і переходи із поглинанням фонона стають малоімовірними. Характерні значення коефіцієнта поглинання для прямих міжзонних переходів становлять $10^4 - 10^5 \text{ см}^{-1}$, а для непрямих міжзонних переходів — $\alpha \cong 10^{-1} - 10^3 \text{ см}^{-1}$.

15.3.3. Екситонне поглинання

Поглинання світла напівпровідником може відбуватись шляхом такого збудження електрона валентної зони, за якого він не переходить до зони провідності, а утворює із діркою зв'язаний стан, який можна розглядати як деяку квазічастинку. Оскільки електрон і дірка характеризуються напівцілим спіном та є ферміонами з антипаралельними спінами, то нова квазічастинка характеризуватиметься нульовим спіном, тобто підлягатиме статистиці Бозе. Така частинка називається екситоном, і, оскільки зв'язаний стан електрона та дірки має меншу енергію за енергію забороненої зони, то екситонне поглинання має спостерігатись за енергій фотона, менших від E_g .

Якщо розміри екситону суттєво перевищують сталу ґратки, то взаємодію електрона та дірки можна уявити як взаємодію двох точкових зарядів, що розташовані у деякому середовищі, яке характеризується діелектричною функцією ϵ . Для визначення екситонного стану припустимо, що \mathbf{r}_n та \mathbf{r}_p – радіус-вектори координат електрона та дірки, що характеризуються ефективними масами m_n^* і m_p^* . Рівняння Шредингера електрон-діркової пари запишемо у вигляді

$$\left\{ -\frac{\hbar^2}{2m_n^*} \nabla_{\mathbf{r}_n}^2 - \frac{\hbar^2}{2m_p^*} \nabla_{\mathbf{r}_p}^2 - V_C(|\mathbf{r}_n - \mathbf{r}_p|) \right\} \Psi(\mathbf{r}_n, \mathbf{r}_p) = E_{exc} \Psi(\mathbf{r}_n, \mathbf{r}_p), \quad (15.46)$$

де

$$V_C(|\mathbf{r}_n - \mathbf{r}_p|) = \frac{e^2}{\epsilon |\mathbf{r}_n - \mathbf{r}_p|} \quad (15.47)$$

– енергія кулонівської взаємодії. Тепер зручно ввести радіус-вектор центру тяжіння електрон-діркової пари

$$\mathbf{R} = \frac{m_n^* \mathbf{r}_n + m_p^* \mathbf{r}_p}{m_n^* + m_p^*}, \quad (15.48)$$

радіус-вектор відносного положення електрона та дірки

$$\mathbf{r} = \mathbf{r}_n - \mathbf{r}_p, \quad (15.49)$$

ефективну сумарну та зведену маси, відповідно

$$M^* = m_n^* + m_p^*, \quad m_r^* = \frac{m_n^* m_p^*}{m_n^* + m_p^*}. \quad (15.50)$$

Тоді (15.46) зведеться до

$$\left\{ -\frac{\hbar^2}{2M^*} \nabla_{\mathbf{R}}^2 - \frac{\hbar^2}{2m_r^*} \nabla_{\mathbf{r}}^2 - V_C(|\mathbf{r}|) \right\} \Psi(\mathbf{R}, \mathbf{r}) = E_{exc} \Psi(\mathbf{R}, \mathbf{r}). \quad (15.51)$$

Розв'язок його шукаємо у вигляді

$$\Psi(\mathbf{R}, \mathbf{r}) = e^{i\mathbf{k}_{exc}\mathbf{R}} \varphi(\mathbf{r}), \quad (15.52)$$

що означає: рух центру екситону відбувається як рух вільної частинки з ефективною масою M^* і описується плоскою хвилею $\exp(i\mathbf{k}_{exc}\mathbf{R})$ із хвилевим вектором екситону $\mathbf{k}_{exc} = \mathbf{k}_n + \mathbf{k}_p$. Хвилева функція $\varphi(\mathbf{r})$ описує збуджені стани екситона та задовольняє рівнянню

$$\left\{ \frac{\hbar^2}{2m_r^*} \nabla_{\mathbf{r}}^2 - V_C(|\mathbf{r}|) \right\} \varphi(\mathbf{r}) = E'_{exc} \varphi(\mathbf{r}), \quad (15.53)$$

де $E'_{exc} = E_{exc} - \frac{\hbar^2 k_{exc}^2}{2M^*}$.

Власні значення (15.53) мають сенс енергії зв'язку екситону та являють собою воднеподібну серію

$$E'_{exc} = -\frac{W}{n^2}, \quad W = \frac{e^4 m_r^*}{2\hbar^2 \epsilon^2}. \tag{15.54}$$

Квантове число $n \geq 1$ описує екситонні стани напівпровідника. Отже повна енергія екситона

$$E_{exc} = \frac{\hbar^2 k_{exc}^2}{2M^*} + \frac{W}{n^2}. \tag{15.55}$$

Зауважимо, що вираз енергії зв'язку екситона формально збігається із виразом для енергії іонізації мілких донорів. Оскільки зведена маса електрон-діркової пари менша за ефективну масу електрона, то енергії зв'язку екситону буде меншою за енергію іонізації донора.

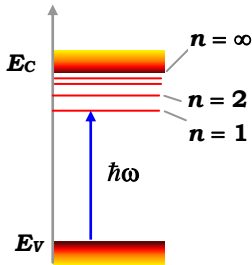


Рис. 15.7. Діаграма енергетичних станів екситону

Для визначення особливостей екситонного поглинання світла приймемо за початок відліку енергії дно зони провідності та вважатимемо, що йому відповідає початок неперервного спектру збудження екситону ($n = \infty$). Тоді енергетична діаграма екситонних збуджень матиме вигляд воднеподібного набору рівнів (рис. 15.7). Оскільки енергія основного стану екситону лежить нижче дна зони провідності, то нижче краю власного поглинання спостерігатиметься серія дискретних рівнів поглинання. На рис. 15.8 подано такий спектр поглинання для закису міді. Оскільки в та-

ких екситонах дипольні переходи заборонені, то лінії спектру поглинання починаються зі значення квантового числа $n = 2$. Інша ситуація виникає у випадку екситонного поглинання в арсеніді галію. Тут дозволений перехід із $n = 1$. Але, оскільки енергія основного стану екситону дуже близько розташована від дна зони провідності, то зазвичай експериментатори спостерігають один пік. Інші лінії дискретного спектру часто замиваються та зливаються із краєм власного поглинання.

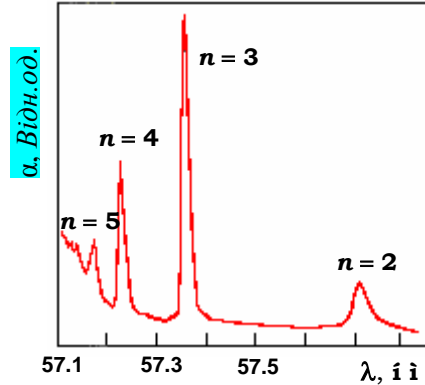


Рис. 15.8. Спектр екситонного поглинання у Cu_2O_3 при $T = 4\text{K}$

На рис. 15.9 подано спектри екситонного поглинання в GaAs за низьких температур. Підвищення температури спричиняє замивання екситонного рівня аж до повного зникнення піка екситонного поглинання. Видно, що за експериментального дослідження як екситонне, так і власне поглинання світла напівпровідниками виявляються на краю смуги поглинання одночасно. Це означає, що спектр поглинання, обчислений без урахування екситонного компонента, має бути модифікований: у випадку прямозонного напівпровідника (15.33) має бути доповнена доданком

$$\Delta\alpha_{exc} = \frac{2\pi(\Delta E_{exc}^1)^{1/2}(\hbar\omega - E_g)^{-1/2}}{1 - \exp\left[2\pi(\Delta E_{exc}^1)^{1/2}(\hbar\omega - E_g)^{-1/2}\right]}, \quad (15.56)$$

де ΔE_{exc}^1 – енергетична відстань між стелею валентної зони та основним станом екситону.

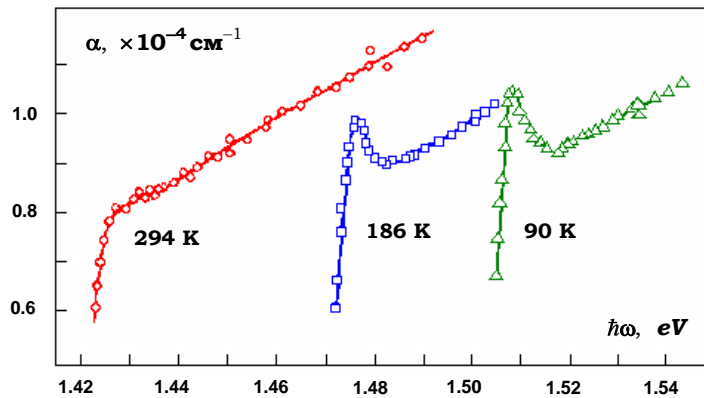


Рис. 15.9. Екситонне поглинання в арсеніді галію

На рис. 15.10 подано спектри поглинання на краю смуги поглинання без урахування та із врахуванням екситонної компоненти. У непрямозонних напівпровідниках утворення екситонів супроводжується виникненням сходинки на краю власного поглинання, що зумовлюється випромінюванням і поглинанням фотонів. При цьому можливі переходи за участю кількох фононів, висилання та поглинання яких можливі у різноманітних комбінаціях. Це зумовлює виникнення кількох сходинок на спектрі поглинання. Насправді, завдяки температурному розмиттю ліній, замість сходинок у спектрах екситонного поглинання спостерігаються точки перегину (рис. 15.11).

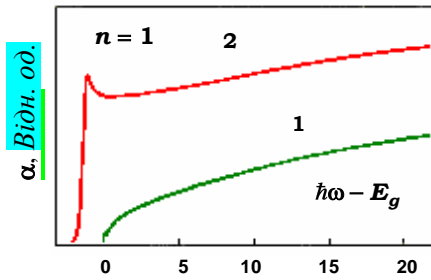


Рис. 15.10. Спектри поглинання, обчислені для:
 1 – прямих міжзонних переходів без урахування **екситонних** факторів;
 2 – із врахуванням екситонних ефектів

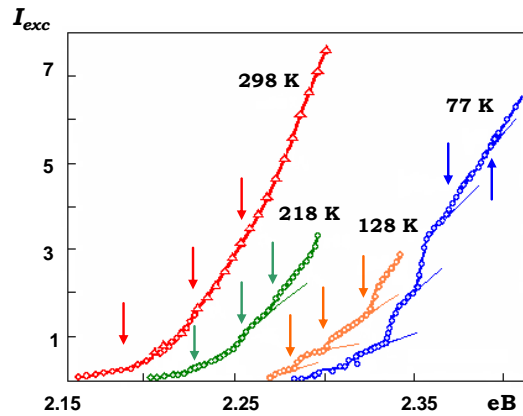


Рис. 15.11. Спектри екситонного поглинання світла у фосфіді галію. Сходинки та точки перегину позначено стрілками

15.3.4. Поглинання вільними носіями заряду

Ще один механізм поглинання світла у напівпровідниках може бути зумовлений поглинанням фотону із переходом носіїв з одного стану до іншого в межах тієї самої зони (рис. 15.12). Такі внутрішньозонні переходи відбуваються із порушенням правил відбору. Дійсно, оскільки виконується закон збереження імпульсу, ці переходи відбуваються разом із поглинанням фону або висиланням фону або коли електрон додатково розсіюється на іонізованих домішках і тим самим компенсує дефіцит імпульсу при переході зі стану E_1 до стану E_2 . Розглянемо детальніше процес поглинання фотона, що спричиняє внутрішньодолинні переходи із поглинанням чи висиланням фону. Оскільки у процесі поглинання беруть участь три частинки – електрон, фотон і фонон, то варто розглядати квантово-механічну задачу у другому наближенні теорії збурень, де необхідно припускати наявність проміжних (віртуальних) станів \mathbf{k}'' між початковим \mathbf{k} і кінцевим \mathbf{k}' станами електрона. Перехід зі стану \mathbf{k} до стану \mathbf{k}' може відбуватись кількома шляхами:

- ♦ спочатку система поглинає фотон з енергією $\hbar\omega$, а потім поглинає чи висилає фонон з енергією $\hbar\omega_{\mathbf{q}}$;
- ♦ поглинанню фотона передують поглинання або висилання електронном фону.

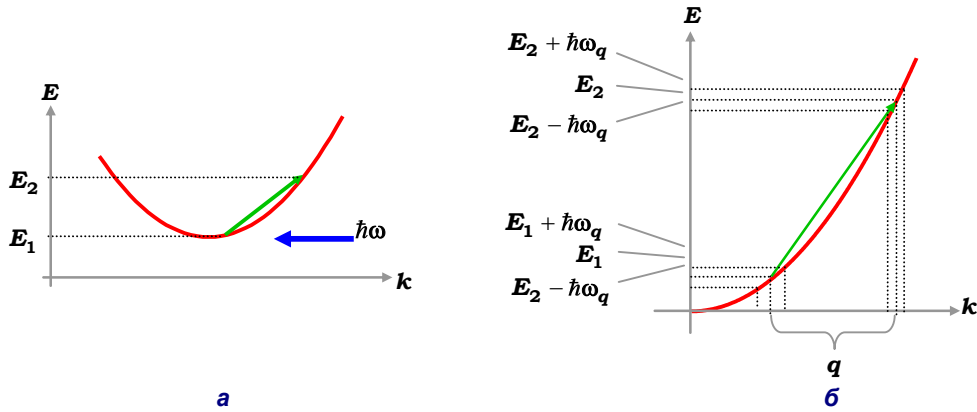


Рис. 15.12. Поглинання фотона із переходом вільного електрона зі стану E_1 у стан E_2 зони провідності, $E_2 - E_1 = \hbar\omega$ (а). Зміна імпульсу носіїв заряду як наслідок зміни їхньої енергії за внутрішньо-долинних переходів за рахунок поглинання світла $\hbar(\mathbf{k}' - \mathbf{k}) = \hbar\mathbf{q}$ (б)

Крім того, на відміну від класичного випадку, необхідно розглядати також процеси індукованого випромінювання фотонів. При цьому, як того вимагає закон збереження імпульсу, цей процес також відбувається із поглинанням або висиланням фонону. Ураховуючи ці процеси, у другому порядку теорії збурень ймовірність переходу електрону зі стану \mathbf{k} до стану \mathbf{k}' із поглинанням фотона можна записати у вигляді

$$W_{\mathbf{k}\mathbf{k}'}^{\pm} = \frac{2\pi}{\hbar} \frac{|P_{\mathbf{k}\mathbf{k}'}|^2 |H_{\mathbf{k}''\mathbf{k}'}|_{\pm}^2}{\hbar^2 \omega^2} \delta(E' - E - \hbar\omega \pm \hbar\omega_q), \quad (15.57)$$

де $\hbar\omega$ – енергія фотона, що поглинається; $P_{\mathbf{k}\mathbf{k}''}$ – матричний елемент, що описує процеси висилання або поглинання фонона; $H_{\mathbf{k}''\mathbf{k}'} -$ матричний елемент, що описує процес поглинання фотона. Ймовірність переходу електрону зі стану \mathbf{k} до стану \mathbf{k}' з індукованим випромінюванням фотона має вигляд

$$\tilde{W}_{\mathbf{k}\mathbf{k}'}^{\pm} = \frac{2\pi}{\hbar} \frac{|P_{\mathbf{k}\mathbf{k}'}|^2 |H_{\mathbf{k}''\mathbf{k}'}|_{\pm}^2}{\hbar^2 \omega^2} \delta(E' - E + \hbar\omega \pm \hbar\omega_q). \quad (15.58)$$

Оскільки взаємодія зонного електрона з полем світлової хвилі визначається оператором (див. (15.9–15.13))

$$H_{e-photon} = i \frac{e\hbar}{m^*} \mathbf{A} \nabla, \quad (15.59)$$

то матричний елемент $H_{\mathbf{k}'\mathbf{k}}$, що визначений на блохівських функціях,

$$H_{\mathbf{k}'\mathbf{k}} = -\frac{eA}{2m^*} \mathbf{e} \mathbf{p}_{\mathbf{k}'\mathbf{k}}. \quad (15.60)$$

У цьому виразі через \mathbf{e} позначено орт поляризації світла, через A – його амплітуду ($A = E_0 / (-i\omega)$) (порівн. із (15.12)), а через

$$\mathbf{p}_{\mathbf{k}'\mathbf{k}} = -i\hbar \int u_{\mathbf{k}'}^* \nabla u_{\mathbf{k}} \quad (15.61)$$

– матричний елемент оператора імпульсу.

Тепер для обчислення коефіцієнта поглинання світла можна поступити таким чином. Нехай електрон переходить зі стану \mathbf{k} до стану \mathbf{k}' , поглинаючи фотон, що рухається у середовищі, яке характеризується діелектричною функцією ϵ . Швидкістю фотону при цьому є $c/\sqrt{\epsilon}$. Число переходів в одиниці об'єму за одиницю часу, підсумоване за всіма кінцевими станами, є швидкістю переходів. Позначимо її як I_{\pm}/V , де знак плюс показує, що перехід відбувається із випромінюванням фонуна, а мінус – із його поглинанням. Якщо число фотонів в об'ємі V дорівнює N , а потік фотонів – $Nc/(\sqrt{\epsilon}V)$, то швидкість електронних переходів в одиниці об'єму за поглинання світла буде добутком потоку фотонів на коефіцієнт поглинання світла, тобто

$$I_{\pm} = \alpha N(c/\sqrt{\epsilon}). \quad (15.62)$$

З іншого боку, електронна підсистема напівпровідника, поглинаючи світло, поглинає енергію, що характеризується потоком

$$S = \hbar\omega Nc/(\sqrt{\epsilon}V). \quad (15.63)$$

Тоді маємо, що коефіцієнт поглинання визначається швидкістю електронних переходів і потоком енергії поглинутого світла

$$\alpha = \frac{I_{\pm} \hbar\omega}{SV}. \quad (15.64)$$

Потік енергії (15.63) можна оцінити як модуль усередненого за часом вектора Пойтинга $\mathbf{S} = \mathbf{E} \times \mathbf{H}$. Для електромагнітного поля, що характеризується вектор-потенціалом (15.13), маємо електричну та магнітну компоненти: $\mathbf{H} = A(\mathbf{e} \times \mathbf{q}) \sin(\mathbf{q}\mathbf{r} - \omega t)$ та $\mathbf{E} = \omega A \mathbf{e} \sin(\mathbf{q}\mathbf{r} - \omega t)$, відповідно. Тоді для середнього за часом вектора Пойтинга дістанемо $\bar{\mathbf{S}} = (\mathbf{e}(\mathbf{e}\mathbf{q}) - \mathbf{q}) \omega A^2 / 2$. У випадку поперечної електромагнітної хвилі $|\mathbf{q}| = \omega\sqrt{\epsilon}/c$. Тоді

$$S = |\bar{\mathbf{S}}| = \omega^2 A^2 \sqrt{\epsilon} / (2c). \quad (15.65)$$

Підставляючи цей вираз до (15.64), маємо

$$\alpha = \frac{2cI_{\pm}\hbar}{\omega A^2 \sqrt{\epsilon} V}. \quad (15.66)$$

Для визначення коефіцієнту поглинання згадаємо, що інтеграл зіткнень має **сенс** швидкості переходу електрона із поглинанням фотона. Його можна визначити стандартно

$$I_{\pm} = \frac{2V}{(2\pi)^6} \int d\mathbf{k} d\mathbf{k}'' \cdot W_{\mathbf{k}\mathbf{k}''}^{\pm} f(\mathbf{k})(1 - f(\mathbf{k}'')). \quad (15.67)$$

У цій формулі проводиться інтегрування за початковими станами електрона, оскільки ці стани залежать від стану фононів, на яких розсіявся електрон у процесі поглинання фотона. Нехай електрони напівпровідника визначається електронною температурою T_e та є не-виродженими. Тоді, вважаючи, що ймовірність знайдення електроном вільного стану, до якого б він міг потрапити після поглинання фотона, мало відрізняється від одиниці, (15.67) можна записати

$$I_{\pm} = \frac{Ve^2 A^2 n}{2(2\pi)^5 \hbar \omega^2 (m^*)^2 N_C} \int d\mathbf{k} d\mathbf{q} |P_{\mathbf{k}\mathbf{k}''}^{\pm}|^2 (\mathbf{e} \cdot \mathbf{q})^2 \delta(E' - E - \hbar\omega \pm \hbar\omega_{\mathbf{q}}) e^{-(E/kT_e)}. \quad (15.68)$$

Ми скористались наближенням $\mathbf{p}_{\mathbf{k}\mathbf{k}''}/m^* \approx \hbar(\mathbf{k}'' - \mathbf{k})/m^* = \hbar\mathbf{q}/m^*$. Для подальшої оцінки I_{\pm} візьмемо до уваги, що матричний елемент $P_{\mathbf{k}\mathbf{k}''}^{\pm}$ в основному залежить від переданого імпульсу – імпульсу фонона \mathbf{q} . Це означає, що інтегрування за імпульсом початкового стану електрону можна провести безпосередньо, і (15.68) можна записати у вигляді

$$I_{\pm} = \frac{Ve^2 A^2 n}{2(2\pi)^5 \hbar \omega^2 (m^*)^2 N_C} \int d\mathbf{q} |P_{\mathbf{q}}^{\pm}|^2 (\mathbf{e} \cdot \mathbf{q})^2 J_{\mathbf{q}}. \quad (15.69)$$

А величина $J_{\mathbf{q}}$ визначається інтегруванням за імпульсом \mathbf{k}

$$\begin{aligned} J_{\mathbf{q}} &= \int k^2 dk \sin \vartheta d\vartheta d\varphi \delta\left(\left(\frac{\hbar^2}{2m^*}\right)(q^2 + 2kq \cos \vartheta) - \hbar\omega \pm \hbar\omega_{\mathbf{q}}\right) e^{-\frac{\hbar^2 k^2}{2m^* k T_e}} = \\ &= -\frac{2\pi m^*}{\hbar^2 q} \int_{k_m}^{\infty} k dk \cdot e^{-\frac{\hbar^2 k^2}{2m^* k T_e}} = \frac{2\pi (m^*)^2 k T_e}{\hbar^4 q} e^{-\frac{\hbar^2 k_m^2}{2m^* k T_e}} \end{aligned} \quad (15.70)$$

Межі інтегрування за модулем хвильового вектора k визначались із умов, за якими аргумент δ -функції дорівнює нулю,

$$k = \frac{m^*}{\hbar^2 q \cos \vartheta} \left(\hbar\omega \mp \hbar\omega_{\mathbf{q}} - \frac{\hbar^2 q^2}{2m^*} \right). \quad (15.71)$$

При цьому нижня межа інтегрування відповідає значенню $\vartheta = 0$, що означає

$$k_m = \frac{m^*}{\hbar^2 q} \left(\hbar\omega \mp \hbar\omega_{\mathbf{q}} - \frac{\hbar^2 q^2}{2m^*} \right). \quad (15.72)$$

Верхня межа відповідає значенню $\vartheta = \pi/2$. Тобто $k = \infty$.

Для подальших оцінок необхідно провести інтегрування за переданим фонону імпульсом. Зручно перейти до нової змінної

$$\zeta = \frac{\hbar^2 q^2}{4m^* kT_e} \quad (15.73)$$

і ввести позначення

$$z_{\pm} = \frac{\hbar\omega \pm \hbar\omega_{\mathbf{q}}}{2kT_e}. \quad (15.74)$$

У цих позначеннях величина $J_{\mathbf{q}}$ із (15.70) має вигляд

$$J_{\mathbf{q}} = \frac{\pi}{\hbar^3} \sqrt{(m^*)^3 kT_e} \frac{e^{z_{\pm}}}{\sqrt{\zeta}} e^{-(\zeta+z_{\pm}^2/\zeta)/2}. \quad (15.75)$$

Підставимо цей вираз до (15.69)

$$I_{\pm} = \frac{Ve^2 A^2 n \sqrt{(m^*)^3 kT_e}}{2(2\pi)^5 \hbar^4 \omega^2 (m^*)^2 N_C} \int d\mathbf{q} |P_{\mathbf{q}}^{\pm}|^2 (\mathbf{e} \cdot \mathbf{q})^2 \frac{e^{z_{\pm}}}{\sqrt{\zeta}} e^{-(\zeta+z_{\pm}^2/\zeta)/2}. \quad (15.76)$$

При обчисленні цього інтеграла зручно вибрати за полярну вісь напрямком орту поляризації світла \mathbf{e} . Тоді можна провести інтегрування за кутами

$$2\pi \int_0^{\pi} q^2 \cos^2 \vartheta \sin \vartheta d\vartheta = \frac{4}{3} \pi q^2. \quad (15.77)$$

У результаті дістанемо

$$I_{\pm} = \frac{8}{3} \frac{Ve^2 A^2 n}{(2\pi)^3 \hbar^9 \omega^2 m^* N_C} (m^* kT_e)^3 \int \zeta d\zeta |P_{\zeta}^{\pm}|^2 e^{z_{\pm}} e^{-(\zeta+z_{\pm}^2/\zeta)/2}. \quad (15.78)$$

Для подальшого аналізу необхідно використовувати конкретну модель електрон-фононної взаємодії, що дає внесок до процесів розсіювання електронів на фононах, тобто конкретизувати вигляд матричного елемента $P_{\mathbf{k}\mathbf{k}}$. Для прикладу розглянемо поглинання світла із розсіюванням електронів на акустичних фононах. Модуль матричного елемента при поглинанні та висиланні фонона не залежить від переданого

імпульсу (q), тобто – і від ζ , а саме $P_{\mathbf{k}\mathbf{k}}^{\pm} \sim \sqrt{kT}$. Таким чином, необхідно обчислити інтеграл

$$L = \int_0^{\infty} \zeta d\zeta e^{z_{\pm}} e^{-(\zeta+z_{\pm}^2/\zeta)/2}. \quad (15.79)$$

Для цього скористаємось визначенням модифікованої функції Бесселя

$$K_2(z) = \frac{1}{2z^2} \int_0^{\infty} \zeta d\zeta e^{-(\zeta+z^2/\zeta)/2}, \quad z > 0, \quad (15.80)$$

тобто із використанням (15.80), отримуємо

$$L = 2z_{\pm}^2 e^{z_{\pm}} K_2(|z_{\pm}|). \quad (15.81)$$

Скористаємось тим, що енергія фонона є малою порівняно з енергією поглинутого фотона, $\hbar\omega_q \ll \hbar\omega$. Це дає $z_{\pm} \approx \hbar\omega/(2kT_e)$. Тоді, підставляючи отриманий вираз до (15.78), із співвідношення (15.66) знайдемо

$$\alpha_a \sim \frac{1}{\sqrt{\varepsilon}} \frac{ne^2}{\omega^3} (m^*)^{1/2} kT(kT_e)^{3/2} \left(\frac{\hbar\omega}{2kT_e} \right)^2 \exp\left(\frac{\hbar\omega}{2kT_e} \right) K_2\left(\frac{\hbar\omega}{2kT_e} \right). \quad (15.82)$$

Аналогічно для процесів з індукованою емісією фотонів отримаємо із (15.82) просто заміною $\hbar\omega \rightarrow -\hbar\omega$

$$\alpha_e \sim \frac{1}{\sqrt{\varepsilon}} \frac{ne^2}{\omega^3} (m^*)^{1/2} kT(kT_e)^{3/2} \left(\frac{\hbar\omega}{2kT_e} \right)^2 \exp\left(-\frac{\hbar\omega}{2kT_e} \right) K_2\left(\frac{\hbar\omega}{2kT_e} \right). \quad (15.83)$$

Повний коефіцієнт поглинання світла із розсіюванням на акустичних фононах $\alpha = \alpha_a - \alpha_e$, або, використовуючи (15.82–15.83),

$$\alpha \sim \frac{1}{\sqrt{\varepsilon}} \frac{ne^2}{\omega^3} (m^*)^{1/2} kT(kT_e)^{3/2} \left(\frac{\hbar\omega}{2kT_e} \right)^2 \operatorname{sh}\left(\frac{\hbar\omega}{2kT_e} \right) K_2\left(\frac{\hbar\omega}{2kT_e} \right). \quad (15.84)$$

У випадку теплової рівноваги $T_e = T$ за таких температур, що $2kT \gg \hbar\omega$, тобто при $z \ll 1$, маємо $\operatorname{sh}(z) \approx z$, $K_2(z) \approx 2/z^2$. У цьому випадку із (15.84) отримуємо

$$\alpha \sim \frac{1}{\sqrt{\varepsilon}} \frac{ne^2}{\omega^2} (m^*)^{1/2} (kT)^{3/2}. \quad (15.85)$$

Таким чином, поглинання світла вільними носіями за розсіювання електронів на акустичних фононах в умовах високих температур зростає за зростання електронної температури та довжини хвилі світла. Поглинання світла вільними носіями в умовах низьких температур $2kT_e \ll \hbar\omega$ не залежить від електронної температури та характеризується коефіцієнтом **поглинання**

$$\alpha \sim \frac{1}{\sqrt{\epsilon}} \frac{ne^2}{\omega^{3/2}} (m^*)^{1/2} kT. \tag{15.86}$$

Іншими словами, у цьому випадку його залежність від довжини хвилі $\sim \lambda^{3/2}$ світла слабша, ніж у класичному (високотемпературному) випадку, де $\alpha \sim \lambda^2$. На рис. 15.13 подано залежність коефіцієнту поглинання від довжини світла за різних температур. Тут C_a – коефіцієнт пропорційності. Штриховими лініями позначено асимптотики коефіцієнту поглинання у класичному – довгохвильовому, і квантовому – короткохвильовому – випадках. На рис. 15.14 подано залежність коефіцієнту поглинання світла від довжини хвилі в арсеніді індію n -типу за кімнатної температури для різних значень концентрації електронів.

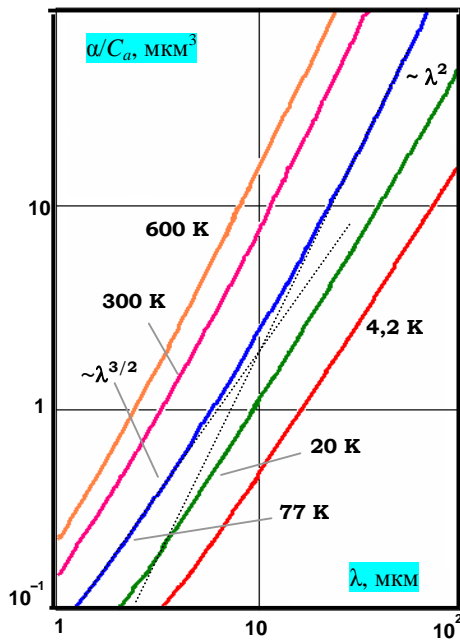


Рис. 15.13. Залежність коефіцієнта поглинання світла вільними носіями від довжини хвилі за розсіювання на акустичних фонах

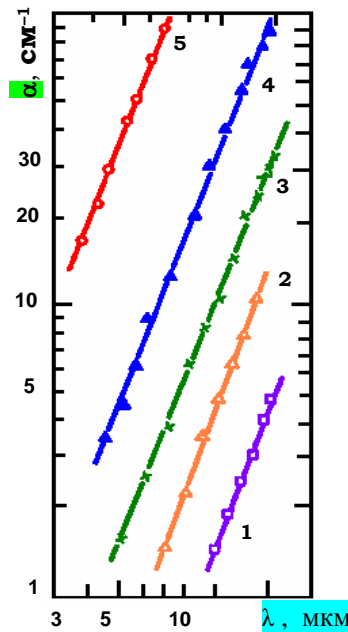


Рис. 15.14. Залежність коефіцієнта поглинання світла вільними носіями в n -InAs від довжини хвилі світла для різних значень концентрацій електронів:
 $n=1: 2,8 \cdot 10^{16}$; $n=2: 1,4 \cdot 10^{17}$;
 $n=3: 2,5 \cdot 10^{17}$; $n=4: 7,8 \cdot 10^{17}$;
 $n=5: 3,9 \cdot 10^{18} \text{ (CM}^{-3}\text{)}$

Оскільки, згідно із (15.56), поглинання світла вільними носіями **ви-**
значається величиною I_{\pm} ((15.78)), яка у свою чергу **залежить від** мат-

ричного елемента розсіювання електрона на коливаннях ґратки (такий вигляд матиме формула, якщо розглядати розсіювання на домішках і дефектах), то можна стверджувати, що α визначається часом релаксації. Отже залежність коефіцієнта поглинання світла вільними носіями заряду від довжини хвилі світла визначатиметься механізмом розсіювання носіїв. Наприклад, при розсіюванні на акустичних фононах $\alpha \sim \lambda^{3/2}$ розсіювання на оптичних фононах приведе до залежності $\alpha \sim \lambda^{5/2}$, а розсіювання іонізованими домішками дає $\alpha \sim \lambda^{7/2}$. Зрозуміло, що у реальних ситуаціях усі ці механізми дають внесок до процесу поглинання світла вільними носіями. Тоді можна описати залежність коефіцієнту поглинання формулою

$$\alpha = A_1 \lambda^{3/2} + A_2 \lambda^{5/2} + A_3 \lambda^{7/2} \quad (15.87)$$

із феноменологічними константами A_i . У більшості ж випадків домінуватиме той чи інший механізм розсіювання, причому домінування великою мірою визначається концентрацією носіїв. Таким чином, часто поведінку коефіцієнта поглинання визначають простим співвідношенням (C – феноменологічна константа)

$$\alpha = C \cdot \lambda^p, \quad (15.88)$$

де показник ступеня p залежить від матеріалу та концентрації носіїв. Наприклад, для германію n -типу в інтервалі концентрацій електронів від $0,5 \cdot 10^{17} \text{ см}^{-3}$ до $5 \cdot 10^{17} \text{ см}^{-3}$ величина p становить 2. Для n -GaAs у цьому самому інтервалі концентрацій електронів $p \approx 3$.

15.3.5. Домішкове поглинання

Поглинання світла напівпровідниками може визначатись іонізацією або збудженням домішок. Таким чином, можливі два принципово різних типи процесів домішкового поглинання світла. По-перше, **можливе поглинання фотона** нейтральним атомом домішки, і перехід її із основного стану до збудженого за механізмом **донор-зона провідності** або **акцептор-валентна зона**. По-друге, – за механізмом **валентна зона-іонізований донор** або **іонізований акцептор-зона провідності**. Ймовірність процесів першого типу значно залежить від частоти падаючого світла та стає максимальною, коли енергія поглинутого фотона дорівнює енергетичній відстані між основним станом домішки та її першим збудженим станом, тобто – енергії іонізації. Оскільки зазвичай енергія іонізації домішки є набагато меншою за ширину забороненої зони, лінії домішкового поглинання такого типу добре розді-

лені зі спектрами поглинання типу **зона-зона** та екситонного поглинання. Домішкове поглинання у цьому випадку проходить таким

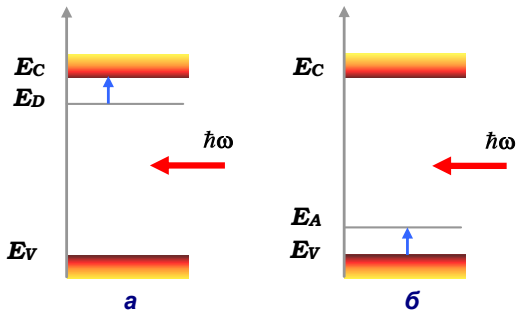


Рис. 15.15. Поглинання світла через іонізацію домішок: а – донора, б – акцептора

чином: нейтральний атом донорної домішки, поглинаючи фотон, збуджується. Електрон покидає атом домішки та переходить до зони провідності (рис. 15.15 а). При цьому енергія фотона $\hbar\omega = \Delta E_D$, де ΔE_D – енергія іонізації донора. Аналогічно може проходити процес домішкового поглинання за збудження нейтрального акцептора. Електрон із валентної зони поглинає фотон і переходить на збуджений рівень

акцептора, тим самим іонізуючи його (рис. 15.15 б). У цьому випадку $\hbar\omega = \Delta E_A$, ΔE_A – енергія іонізації акцептора. Оскільки домішки зазвичай іонізуються за досить низьких температур, домішкове поглинання з іонізацією нейтральних домішок можна спостерігати лише в умовах низьких температур. Іонізовані домішки можуть давати додатковий внесок до поглинання **світла. Домішкові** центри можуть мати енергетичні рівні збудження, енергію яких можна оцінити із формули

$$W_{exc} = E_H \frac{Z^2}{\varepsilon^2} \left(\frac{m^*}{m} \right) \frac{1}{n^2}, \quad n = 2, 3, \dots \quad (15.89)$$

де E_H – енергія іонізації атома водню, що дорівнює 13,5 еВ. Збуджений стан, що відповідає іонізації домішки, визначається значенням головного квантового числа $n = 1$. Таким чином, оптичне збудження, що приводить до переведення домішки із основного стану до збудженого, супроводжуватиметься поглинанням світла, спектр якого характеризуватиметься профілем із кількох ліній. На рис. 15.16 подано спектри поглинання кремнію, що легований бором і миш'яком до рівня $N_B = 1,8 \cdot 10^{16} \text{ см}^{-3}$ та $N_{As} = 1,5 \cdot 10^{15} \text{ см}^{-3}$, відповідно, при температурі 4,2 К. На спектрах поглинання простежуються набори ліній, що відповідають збудженому стану іонізованих домішок.

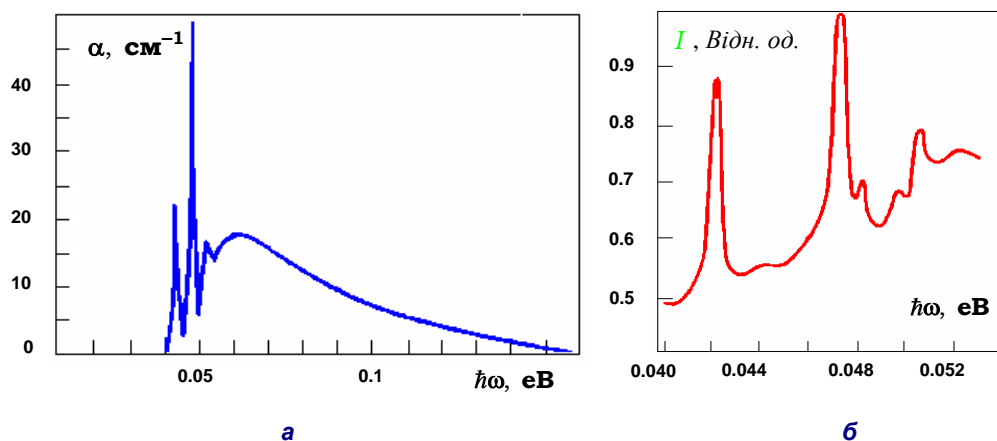


Рис. 15.16. Домішкове поглинання кремнію, легованого: *а* – бором; *б* – миш'яком

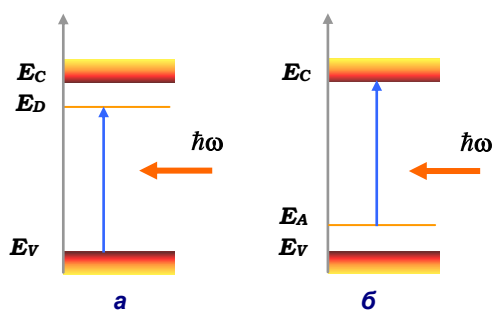


Рис. 15.17. Схема, що пояснює

Механізми поглинання:

- а* – валентна зона-іонізований донор,
б – іонізований акцептор-зона провідності

мілкі рівні характеризуються слабкою локалізацією, електрони, що беруть участь у переходах, невизначені за імпульсом $\Delta p \sim 1/a_B$ (рис. 15.18). Тут a_B – радіус Бора домішкового атома

$$a_B = \varepsilon \cdot \left(\frac{m^*}{m} \right)^{-1} a_H, \quad (15.90)$$

де ε – діелектрична функція напівпровідника, $a_H = 0,5 \cdot 10^{-8}$ [см] – радіус Бора атома водню. Через m^* позначено ефективну масу електрона із зони провідності, якщо розглядається поглинання через мілкий донор, або – ефективну масу дірки у валентній зоні, якщо поглинання відбувається через неглибокий акцепторний рівень. Оскільки ефективна маса дірок зазвичай є більшою за ефективну масу електро-

Процеси поглинання світла за механізмом **валентна-зона іонізований донор** або **іонізований акцептор-зона провідності** схематично подано на рис. 15.17. У випадку мілких домішок край смуги поглинання світла за таким механізмом розташований поблизу краю власного поглинання. Дійсно, у цьому випадку $\hbar\omega_{\tilde{n}} = E_g - E_{D(A)} \approx E_g$.

Крім цього, завдяки тому, що

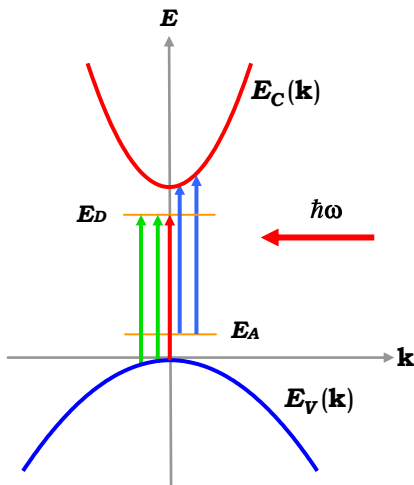


Рис. 15.18. Розкид значень хвильового вектора, що зумовлює виникнення смуг домішкового поглинання

нів, то поглинання світла за механізмом **іонізований акцептор-зона провідності** характеризується спектрами із дещо ширшими лініями. Дійсно, порівнюючи профілі поглинання кремнію, легovanого миш'яком (донорна домішка) рис. 15.16 б та бором (акцепторна домішка) рис. 15.19, можна бачити, що лінії на другому рисунку ширші. Для оцінки коефіцієнту поглинання розглянемо перехід із неглибокого акцепторного рівня до зони провідності. Ймовірність такого переходу пропорційна концентрації іонізованих акцепторів

$$N_A^- = \frac{N_A}{g_A e^{(E_A - E_F)/kT} + 1} \quad (15.91)$$

Згідно із висновками, що зроблені раніше, можна стверджувати: коефіцієнт поглинання пропорційний щільності станів, до яких переходить частинка після поглинання фотону, і числу заповнення вихідного стану. Отже, зважаючи на (15.91), можна записати

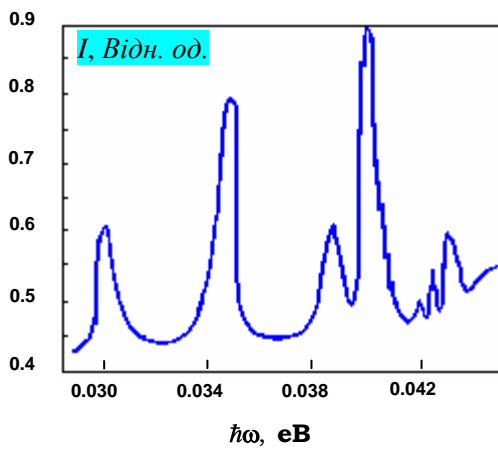


Рис. 15.19. Спектр поглинання кремнію, легovanого бором при \$T = 4,2\$ К

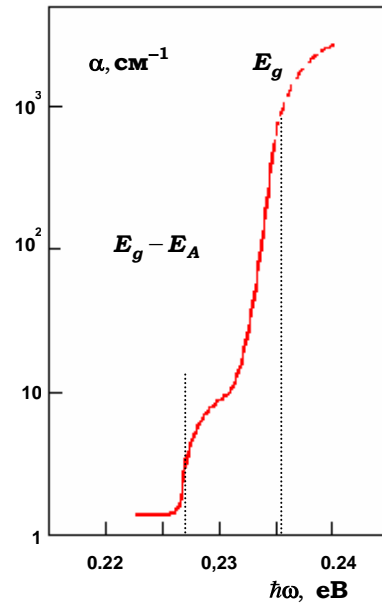


Рис. 15.20. Домішкове поглинання світла в InSb за \$T = 10\$ К, \$E_g - E_A \approx 0,008\$ еВ

$$\alpha \sim \frac{1}{\hbar\omega} \frac{N_A \sqrt{\hbar\omega - (E_g - \Delta E_A)}}{g_A e^{(E_A - E_F)/kT} + 1}. \quad (15.92)$$

Таким чином, хід залежності $\alpha(\omega)$ за домішкового поглинання такий самий, як й у випадку власного поглинання, але край смуги поглинання зсунутий на величину E_A . На рис. 15.20 подано хід домішкового поглинання світла в акцепторному InSb на краю смуги поглинання.

15.3.6. Поглинання ґраткою

Як відомо, у кристалах, де в одній елементарній комірці розташовано більше одного атома, можливі специфічні коливання ґратки, в яких змінюється відносна відстань між атомами в елементарній комірці – оптичні коливання ґратки. У випадку, коли такі кристали характеризуються міжатомними зв'язками із великим ступенем іонності, такі коливання являють собою коливання електричного дипольного моменту, що може бути збуджений змінним електричним полем світлової хвилі. До таких кристалів належать, наприклад напівпровідникові бінарні сполуки типу $A^{III}B^V$, саме в них можливе поглинання світла через механізм збудження зовнішнім оптичним випромінюванням коливань електричних диполів. При цьому найбільш ефективно таке поглинання відбувається за збігу частоти світла із частотою власних коливань електричних диполів. Такі частоти лежать в інфрачервоній області спектра. Оскільки ґратка може поглинати енергію електромагнітного випромінювання тільки за певних значень енергії фонона, спектр поглинання ґраткою матиме вигляд набору піків, що лежать в інфрачервоній області. Положення піків визначається сукупністю оптичних та акустичних гілок коливань ґратки. При цьому має виконуватись закон збереження імпульсу, тобто імпульс поглинутого фотона має дорівнювати сумарному імпульсу фононів, що беруть участь у процесі поглинання. За фіксованої частоти імпульс фотона $q_{phot} = 2\pi\hbar/\lambda_{phot}$ суттєво перевищує відповідний імпульс фонона $q_{phon} = 2\pi\hbar/\lambda_{phon}$. Це означає, що при поглинанні фотона має бути емітовано два чи більше фононів, що спричиняє складну картину спектрів поглинання ґраткою. На рис. 15.21 для прикладу подано спектр поглинання ґраткою арсеніду галію за різних температур. Як видно, спектр поглинання має складну структуру, яка, зокрема при $T = 293K$ містить пік поглинання при $\hbar\omega_p = 0,0648 eV$, обумовлений висиланням двох поперечних оптичних фононів з енергією

$\hbar\omega_{TO}^{(1)} = 0,0324$ еВ. Пік, відповідний частоті $0,0631$ еВ, обумовлений висиланням комбінації із двох поперечних оптичних фотонів частотою

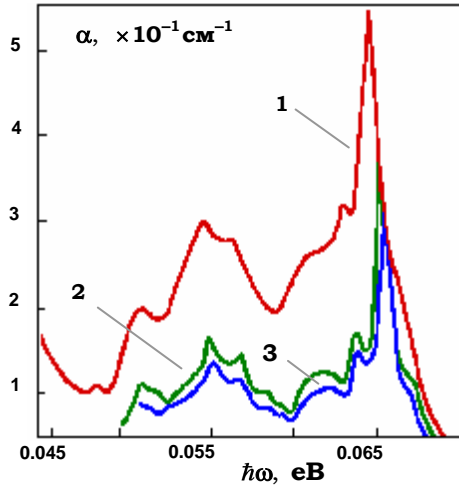


Рис. 15.21. Спектр поглинання ґраткою в GaAs за різних значень температури: 1 – 293 К; 2 – 77 К; 3 – 20 К

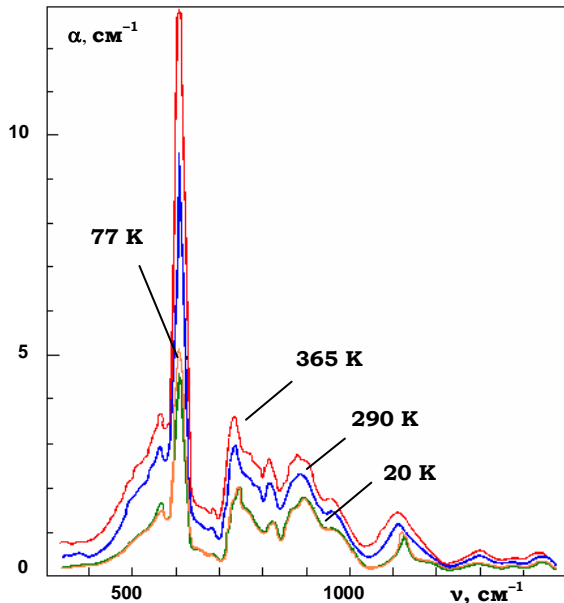


Рис. 15.22. Поглинання світла ґраткою в Si, обумовлене багатофононими процесами за різних температур

$\hbar\omega_{TO}^{(2)} = 0,0316$ еВ, поперечного оптичного фонона $\hbar\omega_{TO}^{(1)} = 0,0324$ еВ і повздовжнього оптичного фонона $\hbar\omega_{LO} = 0,0288$ еВ. Піку з енергією $\hbar\omega_p = 0,548$ еВ відповідає комбінація поперечного оптичного фонона $\hbar\omega_{TO}^{(2)} = 0,0316$ еВ і повздовжнього акустичного фонона $\hbar\omega_{LA} = 0,0232$ еВ.

В елементарних напівпровідниках атоми кристалу є нейтральними, тобто не мають дипольних моментів.

Однак у цих напівпровідниках можна спостерігати кілька слабких смуг поглинання, що зумовлені взаємодією падаючого фотону із кількома фононами. Це поглинання викликано тим, що під дією зовнішнього електромагнітного поля світлової хвилі атоми таких напівпровідників набувають індукованого дипольного моменту, який взаємодіє із випромінюванням. Завдяки цій взаємодії виникають індуквані світлом коливання ґратки. Таке поглинання ґраткою лежить в інфрачервоній області спектра. Наприклад, для кремнію цей діапазон становить 7 – 30 мкм.

На рис. 15.22 подано спектр поглинання ґраткою світла у кремнії за різних температур. Основний пік ($\hbar\omega_p = 0,0756$ еВ тут відповідає поглинанню світла із висиланням комбінації поперечного оптичного фонона $\hbar\omega_{TO} = 0,0598$ еВ і поперечного акустичного фонона $\hbar\omega_{TA} = 0,0158$ еВ.

■ 15.4. Власне поглинання у напівпровідниках в умовах зовнішніх впливів

Зовнішні впливи, такі як температура, гідростатичний тиск та одновісні напруження, електричні та магнітні поля, спричиняють зміни електронних властивостей напівпровідників. Це, у свою чергу, є причиною зовнішніх впливів на процеси оптичного поглинання у напівпровідниках. Зокрема, дії зовнішніх впливів викликають зміни у спектрах власного поглинання напівпровідників. Вивчення таких впливів на спектри оптичного поглинання дозволяє детальніше вивчити особливості енергетичної структури напівпровідника та шляхи її контрольованих змін, що необхідно як для цілей фундаментальної науки, так і для розв'язку практичних задач, пов'язаних із розробкою, удосконаленням нових електронних пристроїв і пошуком нових принципів їхньої побудови. Обговоримо дію деяких зовнішніх впливів на власне поглинання світла напівпровідниками. Почнемо із впливу температури.

15.4.1. Вплив температури на власне поглинання світла

Як відомо, підвищення температури для більшості напівпровідників спричиняє зменшення ширини забороненої зони E_g . Залежність E_g від температури можна описати емпіричною формулою

$$E_g(T) = E_g^0 - \frac{\alpha T^2}{\beta + T}, \quad (15.93)$$

де E_g^0 – ширина забороненої зони при $T = 0$ К, α та β – константи. Хід залежності ширини забороненої зони у германії, кремнії та арсеніді

галію подано на рис. 15.23. На рис. 15.24 подано край смуги поглинання у кремнії за температур рідкого азоту та кімнатної. Видно, що згідно зі зменшенням ширини забороненої зони, край смуги зсувається до довгохвильової області. Причому на цьому рисунку ясно видно перехід від непрямого поглинання до прямого (область перегину кривої при $\alpha \approx 10^2 \text{ см}^{-1}$), коли енергії фотона вистачає для збудження електронних переходів у центрі зони Брилюена.

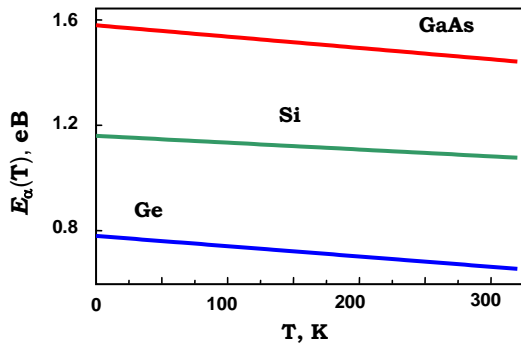


Рис. 15.23. Залежність ширини забороненої зони деяких напівпровідників від температури

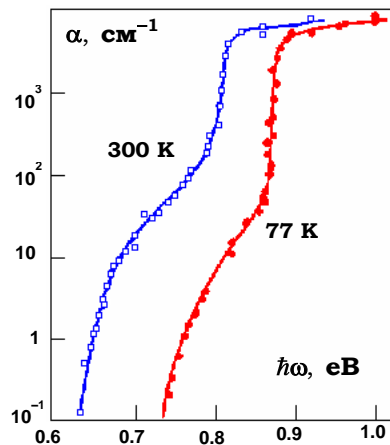


Рис. 15.24. Край смуги поглинання світла у германії

15.4.2. Вплив тиску на власне поглинання світла

Дія гідростатичного тиску на кристал викликає зменшення відстаней між атомами кристалічної ґратки. Це є причиною зміни внутрішньокристалічного потенціалу, що очевидно приводить до зміни зонної структури. Не дуже значні тиски, зокрема у прямозонних напівпровідниках, спричиняють збільшення ширини забороненої зони: в умовах гідростатичного тиску край смуги поглинання у прямозонних напівпровідниках зсуватиметься у короткохвильовий бік. Якщо тиск підвищується, поглинання світла прямозонними напівпровідниками стає складнішим. Наприклад, в арсеніді галію після досягнення максимуму при $P \approx 6 \cdot 10^4 \text{ атм}$ край смуги поглинання починає зсуватись у довгохвильовий бік (рис. 15.25). Така поведінка краю поглинання за гідростатичного тиску пояснюється тим, що за невисоких тисків дно зони провідності піднімається щодо стелі валентної зони, яка також розташована у точці Г. Але одночасно із цим долини, розташовані на висі $\langle 100 \rangle$ та еквівалентних їй опускаються зі зростанням тиску. За

тиску $\approx 6 \cdot 10^4$ атм. виявляється, що долини розташовані нижче за основну долину, і напівпровідник починає поглинати як непрямозонний із шириною забороненої зони меншою за $E_g(P=0)$. У результаті на

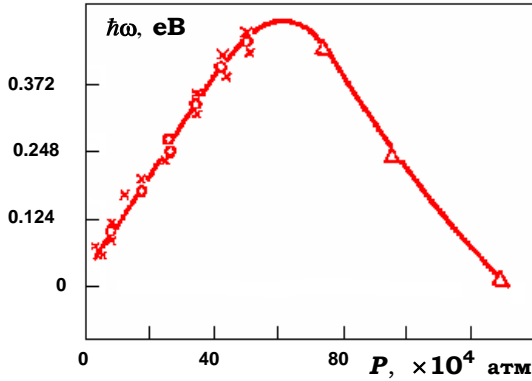


Рис. 15.25. Залежність краю поглинання світла в GaAs від гідростатичного тиску при 300 К

на кривій залежності $\hbar\omega(P)$ виникає область спаду. Аналогічно у непрямозонних напівпровідниках поведінка краю смуги поглинання під дією гідростатичного тиску є досить складною та різною для різних матеріалів. Наприклад, германій при збільшенні гідростатичного тиску демонструє зсув краю смуги поглинання до короткохвильової області, а кремній, навпаки, – до довгохвильової.

15.4.3. Вплив електростатичного поля на власне поглинання світла. Ефект Франца–Келдиша

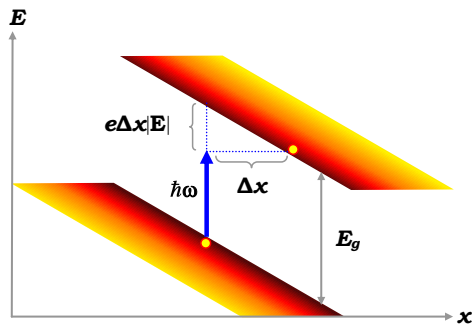


Рис. 15.26. Діаграма енергетичних зон та оптичні переходи за наявності електричного поля

Під дією сталого однорідного електричного поля у напівпровіднику відбувається нахил енергетичних зон, причому кут нахилу пропорційний величині електростатичного поля. Таким чином, при переході електрона під дією зовнішнього світла із валентної зони до зони провідності можливе виявлення ефекту тунелювання. Це спричинить зсув краю смуги поглинання до довгохвильової області, оскільки перехід електрона із

валентної до зони провідності тепер може відбутися за поглиннанні світла з енергією, меншою за ширину забороненої зони. Дійсно, як видно із рис. 15.26

$$e |\mathbf{E}| \Delta x = E_g - \hbar\omega, \quad (15.94)$$

для переходу електрона з валентної до зони провідності в умовах дії електричного поля \mathbf{E} електрону досить подолати трикутний потенціальний бар'єр висотою $\hbar\omega$ та ефективною шириною $(E_g - \hbar\omega)/e|\mathbf{E}|$.

При цьому ймовірність такого переходу є ненульовою за будь-яких значень частоти світла, що поглинається. Однак, ця ймовірність швидко зменшується за збільшення різниці $E_g - \hbar\omega$. Таким чином, за наявності електричного поля у напівпровіднику може відбуватись поглинання світла з енергією, меншою за енергію забороненої зони. Вплив електричного поля на край поглинання напівпровідника було теоретично незалежно досліджено Францем і Келдишем у 1958 р. Для обчислення коефіцієнту електропоглинання вчені застосували формулу ймовірності тунелювання через трикутний бар'єр, використовуючи наближені вирази інтегралів Ейрі для випадку невеликих полів та енергій фотонів $\hbar\omega \approx E_g$. Хід їхніх міркувань був приблизно таким: Як вихідна позиція використовувалось рівняння Шредингера для електрона, що перебуває у полі трикутного потенціалу

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} - eEx\psi = \varepsilon\psi, \quad (15.95)$$

де ε – кінетична енергія електрона за руху у напрямку Ox . Вводячи нову безрозмірну змінну

$$\xi = -\frac{1}{l} \left(x + \frac{\varepsilon}{eE} \right), \quad (15.96)$$

де

$$l = \left(\frac{\hbar^2}{2meE} \right)^{1/3}, \quad (15.97)$$

рівняння (15.95) можна привести до вигляду

$$\frac{d^2\psi}{d\xi^2} = \xi\psi. \quad (15.98)$$

Розв'язком цього рівняння є функція Ейрі

$$\Phi(\xi) = \frac{1}{\sqrt{\pi}} \int_0^\infty \cos(u\xi + u^3/3) du. \quad (15.99)$$

За великих значень аргументу ξ цю функцію можна представити у вигляді

$$\Phi(\xi) \approx \frac{1}{2} \xi^{-1/4} e^{-(2/3)\xi^{3/2}}. \quad (15.100)$$

Ймовірність перебування електрона у нескінченно малому околі точки x : $W = |\Phi(\xi)|^2 dx$. Тоді можна знайти коефіцієнт поглинання світла за тунелювання електрона через трикутний бар'єр, оскільки він буде пропорційним ймовірності W , проінтегрованої за шириною забороненої зони

$$\alpha \sim \int |\Phi(\xi)|^2 d\xi. \quad (15.101)$$

Виконуючи інтегрування частинами та беручи до уваги (15.98), тобто, $d^2\Phi/d\xi^2 = \xi\Phi$, інтеграл у правій частині (15.101) можна оцінити так

$$\alpha \sim \int \Phi^2 d\xi = \Phi^2 \xi - \left(\frac{d\Phi}{d\xi} \right)^2. \quad (15.102)$$

У результаті для коефіцієнту поглинання світла для прямих дозволених переходів ($\hbar\omega < E_g$) було отримано вираз

$$\alpha(\omega, \mathbf{E}) = A \cdot \frac{e|\mathbf{E}|}{\hbar\omega \cdot (E_g - \hbar\omega)} \cdot \exp \left[-\frac{4(2m_r)^{1/2}(E_g - \hbar\omega)^{1/2}}{3\hbar e|\mathbf{E}|} \right], \quad (15.103)$$

де $m_r = m_e^* m_h^* (m_e^* + m_h^*)^{-1}$ – приведена ефективна маса носіїв заряду, A – коефіцієнт, залежний від m_r і матричного елемента електродипольного переходу. Таким чином, на краю смуги поглинання з'являється "хвіст", що експоненційно спадає у довгохвильовий бік.

За енергій фотонів, що перевищують ширину забороненої зони, тобто $\hbar\omega > E_g$, електропоглинання стає складнішим: коефіцієнт поглинання для прямих дозволених переходів **має вигляд**

$$\alpha(\omega, \mathbf{E}) = A(\hbar\omega - E_g)^{1/2} [1 + G(\omega, |\mathbf{E}|)], \quad (15.104)$$

де $G(\omega, |\mathbf{E}|)$ – осцилююча функція енергії фотона та електричного поля, що відбиває властивості інтегралів Ейрі. Ці осциляції у поглинання зменшуються за мірою збільшення відстані $\hbar\omega - E_g$. Зі збільшенням електричного поля амплітуда та період осциляцій збільшуються. Наприклад, оцінки показують, що для полів $E \approx 5 \cdot 10^6 \text{ В [см}^{-1}\text{]}$ період осциляцій досягає величини 0,05 еВ, а амплітуда зміни α становить близько $10^3 [\text{см}^{-1}]$. На практиці процеси розсіювання носіїв заряду, особливо на фонах, призводитимуть до швидкішого затухання осциляцій, ніж передбачає теорія. За малих коефіцієнтів поглинання в області експоненційного краю, що описується формулою

$$\alpha = \alpha_0 \exp[\beta(\hbar\omega - E_0)], \quad (15.105)$$

дія електричного поля спричиняє простий зсув краю поглинання.

$$\alpha = \alpha_0 \exp\left[\beta(\hbar\omega - E_0 + \beta^2 e^2 \hbar^2 E^2 / 12m_r)\right]. \quad (15.106)$$

За експериментальних досліджень електропоглинання для спостереження зсуву краю смуги поглинання необхідно використовувати високоомні зразки для зменшення ролі теплових ефектів: необхідно використовувати широкозонні напівпровідники або виконувати вимірювання за низьких температур. Вимірювання ефекту Франца–Келдиша було виконано Моссом в 1961 р. на зразках GaAs, зокрема було показано, що зсув краю смуги поглинання добре описується формулою (15.106). Ефект Франца–Келдиша краще вивчати експериментально, прикладаючи змінне електричне поле до зразка та проводячи синхронне детектування сигналу, що викликається світлом, яке пройшло скрізь зразок. Наприклад, до зразку у вигляді тонкої пластинки ($h \sim 10^{-1}$ см) прикладають змінне електричне поле напрутою $\sim 10^3$ В, тобто $E \sim 10^4$ В [см $^{-1}$], та освітлюють зразок монохроматичним світлом. При цьому вимірюють інтенсивність світла, що пройшло зразок, детектуючи сигнал на частоті, що дорівнює подвійній частоті змінного

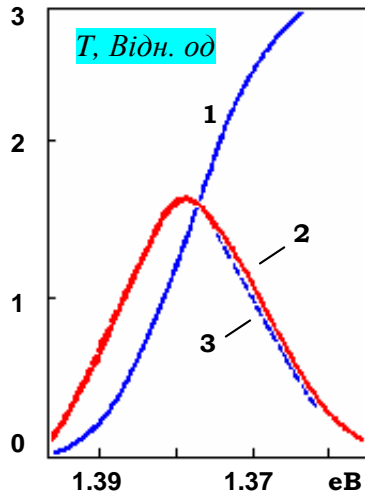


Рис. 15.27. Прозорість GaAs (1) поблизу краю поглинання. Середньоквадратична змінна складової прозорості, що виникає за дії на зразок змінного електричного поля (2). Крива (3) – результат диференціювання кривої (1)

електричного поля. Сигнал буде пропорційний похідній коефіцієнта поглинання по $\hbar\omega$, тобто нахилу кривої пропускання (як функції $\hbar\omega$) за відсутності електричного поля (крива 3 рис. 15.27 є похідною кривої 1, що вимірюється у нульовому полі). Таким чином, на кожному краї смуги поглинання спостерігатимуться "диференціальні" піки, що дає можливість використовувати цю методику для визначення положення краю, який не спостерігається явно за прямих вимірювань поглинання.

15.4.4. Міжзонне магнетопоглинання

Розглянемо поглинання світла напівпровідником, в якого екстремуми валентної зони та зони провідності розташовані в одній точці $\mathbf{k} = 0$ і на який діє зовнішнє магнітне поле B . Вважатимемо зони не виродженими (рис. 15.28 а). Дія магнітного поля приводить до того, що рух електронів у напрямку, перпендикулярному полю, стає квантованим так, що закон дисперсії носіїв трансформується у набір вкладених одна в одну парабол (рис. 15.28 б). Магнітне поле не тільки спричиняє квантування енергії електронів в напівпровідникові, але й зсуває дно зони провідності угору на величину $\hbar\omega_e/2$, а стелю валентної зони – згори на величину $\hbar\omega_h/2$, де $\omega_e = eB/m_e^*$ і $\omega_h = eB/m_h^*$ – циклотронні частоти електронів і дірок, відповідно. Таким чином, під дією магнітного поля ширина забороненої зони напівпровідника зростає на величину

$$\Delta E_g = \frac{e\hbar B}{2m_r}. \quad (15.107)$$

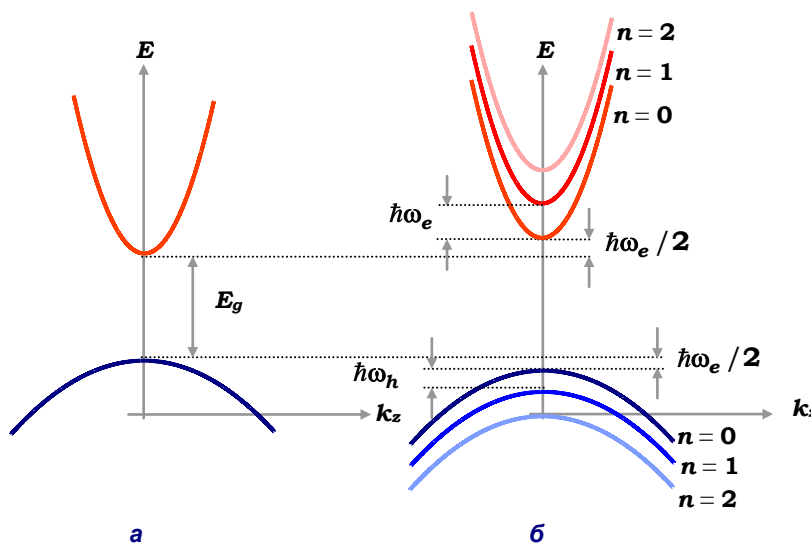


Рис. 15.28. Енергетична діаграма двозонної моделі напівпровідника:
а – без дії; б – під дією магнітного поля

За прямих переходів між підзонами, що виникли у результаті квантування Ландау, поглинання світла електронами напівпровідника, що розташований у магнітному полі, відбувається за умови

$$\hbar\omega_n = E_g + (n + 1/2)(\hbar\omega_e + \hbar\omega_h) + \beta B m_i (g_e + g_h), \quad (15.108)$$

де $\beta = e\hbar/2m$ – магнетон Бора; g_e та g_h – ефективні g -фактори електронів і дірок напівпровідника, відповідно; $m_i = \pm 1/2$ – спінове квантове число. При цьому такі переходи можуть відбуватись тільки при виконанні правил відбору, що залежать від поляризації світла, яке поглинається,

$$\Delta n = 0, \Delta m_i = \pm 1 \quad \text{– для світла поляризованого так, що } \mathbf{E} \perp \mathbf{B}, \quad (15.109)$$

$$\Delta n = 0, \Delta m_i = 0 \quad \text{– для світла поляризованого так, що } \mathbf{E} \parallel \mathbf{B}. \quad (15.110)$$

Оскільки розглядаються прямі міжзонні переходи, то для визначення коефіцієнта поглинання можна скористатись (15.32) і записати $\alpha(\omega, B) \sim \rho(\omega, B)$. Записуючи щільність станів в явному вигляді (див. (8.90)), отримуємо, що

$$\alpha_{n,m_i}(\omega, B) = A(\omega_e + \omega_h) \sum_n (\hbar\omega - \hbar\omega_n)^{-1/2}, \quad (15.111)$$

де A – коефіцієнт, що залежить від показника заломлення матеріалу. Згадаємо, що коефіцієнт поглинання у нульовому полі

$$\alpha(\omega, B = 0) = A(\hbar\omega - E_g)^{1/2}. \quad (15.112)$$

Із порівняння (15.111) та (15.112) видно, що замість звичайного квазінеперервного спектру поглинання, що має місце за відсутності магнітного поля, наявність поля спричинить появу серії піків за енергій, що відповідають максимумам щільності станів. Типову поведінку експериментальної кривої магнетопоглинання подано на рис. 15.29. Зазначимо, що особливості в $\alpha(B)$ (формула (15.111)) зникають, якщо врахувати ефекти розширення рівнів (обмежена величина часу релаксації τ), що пов'язані із дисипативними процесами в матеріалі. Добре роз-

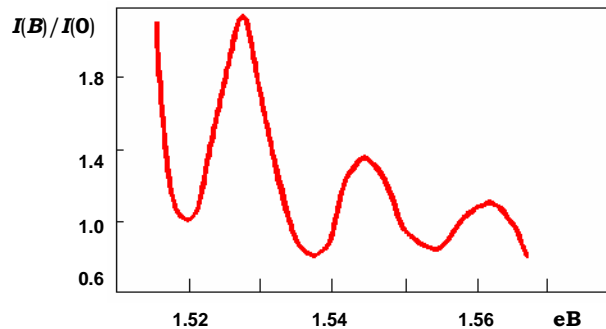


Рис. 15.29. Осциляції магнетопоглинання світла в GaAs при температурі 77 К ($B = 0,8$ Тесла)

ділені піки спостерігатимуться тільки у випадках, коли умова $\omega_{\hbar}\tau \gg 1$, ($\omega_c = \omega_e, \omega_h$) виконується принаймні для однієї із зон.

Із правилами відбору (15.109-110) пов'язані поляризаційні ефекти. Наприклад, для світла, що розповсюджується паралельно магнітному полю ($\mathbf{E} \perp \mathbf{B}$), перехід $\Delta m_i = +1$ відповідає правій круговій поляризації, а перехід $\Delta m_i = -1$ – лівій круговій поляризації щодо напрямку вздовж розповсюдження світла. У випадку, коли світло, що поглинається, розповсюджується перпендикулярно магнітному полю ($\mathbf{k} \perp \mathbf{B}$, \mathbf{k} – хвильовий вектор світла) обидва типи дозволених переходів $\Delta m_i = 0$ ($\mathbf{E} \parallel \mathbf{B}$) і $\Delta m_i = +1$ ($\mathbf{E} \perp \mathbf{B}$) відповідають плоскополяризованому випромінюванню.

■ 15.5. Циклотронний резонанс

Якщо носій заряду рухається під кутом до напрямку магнітного поля \mathbf{B} , то його траєкторія являтиме спіраль, що навивається вздовж напрямку магнітного поля, а кругова частота обертання визначається циклотронною частотою (див. (5.37))

$$\omega_c = \frac{eB}{m^*}. \quad (15.113)$$

Для оцінки циклотронної частоти у напівпровідниках зручно користуватись формулою

$$\omega_c = 56\pi \frac{m}{m^*} B \text{ [ГГц]}. \quad (15.114)$$

Якщо ефективна маса носіїв наближено дорівнює масі вільного електрона, а магнітне поле становить $\sim 10^{-1}$ Тесла, то циклотронна частота відповідає надвисокочастотному діапазону. Таким чином, якщо електромагнітна хвиля надвисокочастотного діапазону, поляризована під деяким кутом до напрямку \mathbf{B} , падає на зразок, вона поглинається найефективніше на частоті $\omega = \omega_c$. Це явище називається циклотронним резонансом і може використовуватись для визначення деталей зонної структури напівпровідника. Розглянемо цей ефект детальніше. Для цього запишемо рівняння руху носія заряду у напівпровіднику у змінному електричному та сталому магнітному полях, вважаючи що розподіл носіїв за енергіями має вигляд дуже вузького піку

$$\frac{d(m^* \mathbf{v})}{dt} + \frac{m^* \mathbf{v}}{\tau_m} = e[\mathbf{E} + (\mathbf{v} \times \mathbf{B})], \quad (15.115)$$

де τ_m – час релаксації. Якщо у декартовій системі координат вибрати вісь OZ так, щоб її напрямок збігався із напрямком магнітного поля, то, переходячи до Фур'є-трансформант за часом, і враховуючи, що $\frac{d}{dt} f \rightarrow i\omega f$, із (15.115) дістанемо

$$\begin{aligned} i\omega m^* v_x + m^* v_x \tau_m^{-1} &= eE_x + ev_y B, \\ i\omega m^* v_y + m^* v_y \tau_m^{-1} &= eE_y - ev_x B, \\ i\omega m^* v_z + m^* v_z \tau_m^{-1} &= eE_z. \end{aligned} \quad (15.116)$$

Оскільки

$$ev_k n = j_k = \sigma_{kl} E_l, \quad (15.117)$$

то із останнього рівняння системи (15.116) знайдемо

$$\sigma_{zz} = \frac{\sigma_0 \tau_m^{-1}}{i\omega + \tau_m^{-1}}, \quad (15.118)$$

де $\sigma_0 = ne^2 \tau_m / m^*$. Два перші рівняння системи при цьому перепишемо у вигляді

$$\begin{aligned} j_x - \frac{\omega_C}{i\omega + \tau_m^{-1}} j_y &= \frac{\sigma_0 \tau_m^{-1}}{i\omega + \tau_m^{-1}} E_x, \\ \frac{\omega_C}{i\omega + \tau_m^{-1}} j_x + j_y &= \frac{\sigma_0 \tau_m^{-1}}{i\omega + \tau_m^{-1}} E_y. \end{aligned} \quad (15.119)$$

Звідси знаходимо компоненти провідності

$$\sigma_{xx} = \sigma_{yy} = \sigma_0 \tau_m^{-1} \frac{i\omega + \tau_m^{-1}}{(i\omega + \tau_m^{-1})^2 + \omega_C^2}, \quad (15.120)$$

$$\sigma_{xy} = -\sigma_{yx} = \sigma_0 \tau_m^{-1} \frac{\omega_C}{(i\omega + \tau_m^{-1})^2 + \omega_C^2}. \quad (15.121)$$

Нехай на зразок напівпровідника **у магнітному** полі падає електромагнітна хвиля із правою круговою поляризацією, тобто компоненти електричного поля у такій хвилі пов'язані співвідношенням $E_y = -iE_x$.

Тоді, використовуючи (15.120) і (15.121), можна записати, що ефективна

електропровідність (відгук на хвилю правополяризованого електромагнітного поля)

$$\sigma_{\pm} = \frac{j_x}{E_x} = \sigma_{xx} + \frac{\sigma_{xy} E_y}{E_x} = \sigma_0 \tau_m^{-1} \frac{(i\omega + \tau_m^{-1}) - i\omega_C}{(i\omega + \tau_m^{-1}) + \omega_C^2} = \frac{\sigma_0 \tau_m^{-1}}{(i\omega + \tau_m^{-1}) + i\omega_C}. \quad (15.122)$$

Аналогічно для лівополяризованої електромагнітної хвилі

$$\sigma_{\mp} = \frac{j_y}{E_y} = \sigma_{yy} + \frac{\sigma_{yx} E_x}{E_y} = \sigma_0 \tau_m^{-1} \frac{(i\omega + \tau_m^{-1}) + i\omega_C}{(i\omega + \tau_m^{-1}) + \omega_C^2} = \frac{\sigma_0 \tau_m^{-1}}{(i\omega + \tau_m^{-1}) - i\omega_C}. \quad (15.123)$$

Розглянемо геометрію, коли електричне поле електромагнітної хвилі перпендикулярне до сталого магнітного поля \mathbf{B} (геометрія Фарадея), тоді потужність, що поглинається одиничним об'ємом напівпровідника, визначатиметься дисипативною функцією

$$P(\omega) = \frac{1}{2} \operatorname{Re}(\mathbf{jE}^*). \quad (15.124)$$

Оскільки часова залежність електромагнітного поля падаючого світла має вигляд $E_j = E_j^0 e^{i\omega t}$, то згідно із (15.124), потужність енергії електромагнітного поля кожної поляризації, що поглинається носіями напівпровідника,

$$P_{\pm}(\omega) = \frac{|E_j^0|^2}{2} \operatorname{Re} \frac{\sigma_0 \tau_m^{-1}}{\tau_m^{-1} + i(\omega \pm \omega_C)} = \frac{|E_j^0|^2}{2} \frac{\sigma_0 \tau_m^{-2}}{\tau_m^{-2} + (\omega \pm \omega_C)^2}. \quad (15.125)$$

Лінійно-поляризовану хвилю можна представити як суму двох право- та лівополяризованих хвиль кругової поляризації, тобто енергія лінійно-поляризованої хвилі, що поглинається напівпровідником за одиницю часу, визначатиметься формулою

$$P(\omega) = P_+(\omega) + P_-(\omega). \quad (15.126)$$

Підставляючи до цієї формули вирази для P_+ і P_- із (15.125), отримаємо

$$\begin{aligned} P(\omega) &= \frac{\sigma_0}{2} |E_j^0|^2 \left\{ \frac{1}{1 + (\omega + \omega_C)^2 \tau_m^2} + \frac{1}{1 + (\omega - \omega_C)^2 \tau_m^2} \right\} = \\ &= \sigma_0 |E_j^0|^2 \frac{1 + (\omega^2 + \omega_C^2) \tau_m^2}{\left[1 + (\omega^2 - \omega_C^2) \tau_m^2 \right]^2 + 4\omega_C^2 \tau_m^2}. \end{aligned} \quad (15.127)$$

Явище циклотронного резонансу настає, якщо $\omega_C^2 \tau_m^2 \gg 1$ і $\omega = \omega_C$. У цьому випадку

$$P(\omega_C) = \frac{\sigma_0}{2} |E_j^0|^2. \quad (15.128)$$

В експериментах зазвичай частоту утримують сталою, змінюючи величину магнітного поля, тобто для аналізу умов резонансу зручно як незалежну змінну використовувати величину ω_C/ω . При цьому результати варто представляти у безрозмірному вигляді, нормуючи величину енергії $P(\omega_C)$, що поглинається в одиницю часу зразком, який розташований у магнітному полі, до енергії, що поглинається в нульовому магнітному полі $P_0 = P(\omega_C = 0)$. На рис. 15.30 подано залежності $P(\omega_C)/P_0$ від ω_C/ω за різних значень часу релаксації. Видно, що за зменшення $\omega\tau_m$ резонансний пік розпливається, і при $\omega\tau_m < 1$ резонанс взагалі зникає (крива 4 на рисунку). Зауважимо, що за значень $\omega\tau_m \approx 1$ лінія, яка відповідає резонансному поглинанню поля, сильно розширюється та зсувається у бік менших частот (крива 3 на рисунку).

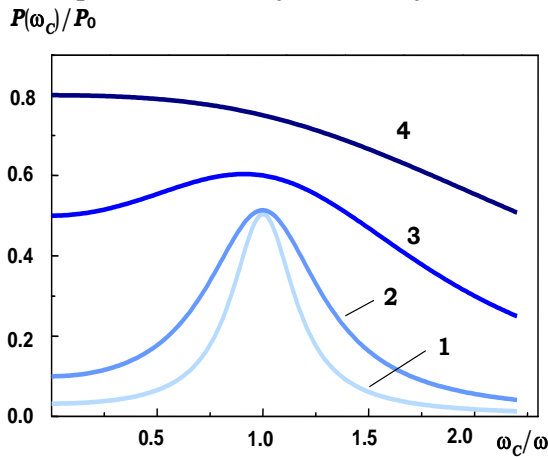


Рис. 15.30. Залежність потужності електромагнітного поля, поглинутої напівпровідником, від ω_C/ω за різних значень часу релаксації:
1 – $\omega\tau_m = 5,5$; **2** – $\omega\tau_m = 3$; **3** – $\omega\tau_m = 1$;
4 – $\omega\tau_m = 0,5$

Врахуємо тепер той факт, що електронні властивості напівпровідників визначаються тензором ефективних мас, що означає: умови поглинання поля напівпровідником будуть різними для різних співвідношень між орієнтаціями тестуючого поля, магнітного поля та кристалографічними осями зразку. Запишемо рівняння (15.116) в умовах циклотронного резонансу, коли $\omega = \omega_C$ і величинами E_i та τ_m^{-1} можна знехтувати. Вважаючи, що напівпровідник характеризується діагональним тензором ефективних мас, маємо

$$\begin{aligned} i\omega m_x^* v_x + e(v_z B_y - v_y B_z) &= 0, \\ i\omega m_y^* v_y + e(v_x B_z - v_z B_x) &= 0, \\ i\omega m_z^* v_z + e(v_y B_x - v_x B_y) &= 0. \end{aligned} \quad (15.129)$$

Введемо направляючі косинуси магнітного поля щодо кристалографічних осей зразку (схему експерименту з циклотронного резонансу подано на рис. 15.31. Напрямки обертання електронів і дірок удовж циклотронних орбіт – різні) так, що $B_x = B\alpha$, $B_y = B\beta$ і $B_z = B\gamma$. Зв'язок з полярним та азимутальним кутами задається стандартними формулами $\gamma = \cos \vartheta$, $\alpha = \sin \vartheta \cos \varphi$, $\beta = \sin \vartheta \sin \varphi$. Тоді із системи (15.129) отримаємо умову нетривіального розв'язку

$$\begin{vmatrix} i\omega_C m_x^* & -eB\gamma & eB\beta \\ eB\gamma & i\omega_C m_y^* & -eB\alpha \\ -eB\beta & eB\alpha & i\omega_C m_z^* \end{vmatrix} = 0. \quad (15.130)$$

Розв'язок цього рівняння дає зв'язок між резонансною частотою та компонентами тензора ефективних мас

$$\omega_C = eB \left(\frac{\alpha^2 m_x^* + \beta^2 m_y^* + \gamma^2 m_z^*}{m_x^* m_y^* m_z^*} \right)^{1/2}. \quad (15.131)$$

Звідси можна визначити циклотронну ефективну масу m_C^* :

$$\omega_C = eB / m_C^*$$

$$m_C^* = \left(\frac{m_x^* m_y^* m_z^*}{\alpha^2 m_x^* + \beta^2 m_y^* + \gamma^2 m_z^*} \right)^{1/2}. \quad (15.132)$$



Рис. 15.31. Схема експерименту з вимірювання резонансного поглинання електромагнітного поля напівпровідником під дією сталого магнітного поля.

Праворуч – схеми формування електронних і діркових орбіт

Формули (15.131–15.132) свідчать про те, що від орієнтації вектора магнітного поля щодо кристалографічних осей напівпровідника залежить резонансна частота. Це означає, що при вивченні залежності поглинутої потужності поля у кристала, може існувати кілька піків. На рис. 15.32 подано експериментальні залежності поглинутої потужності (відносні одиниці) електромагнітного поля частотою 23 ГГц кремнієвим зразком n -типу від відношення $u = \omega_C / \omega$ для трьох різних орієнтацій магнітного поля. Як видно, багатодолинність закону дисперсії кремнію виявляється у формування спектрів поглинання. Наприклад, коли магнітна індукція \mathbf{B} паралельна кристалографічному напрямку $\langle 001 \rangle$, крива поглинання має два яскраво виражених максимуми при $u_1^{\langle 001 \rangle} = 0,19$ та $u_2^{\langle 001 \rangle} = 0,43$. Два піки виявляються й в геометрії, коли магнітна індукція паралельна кристалографічному напрямку $\langle 110 \rangle$. У цьому випадку $u_1^{\langle 110 \rangle} = 0,28$ та $u_2^{\langle 110 \rangle} = 0,43$. У випадку, коли \mathbf{B} направлена паралельно напрямку $\langle 111 \rangle$, крива циклотронного резонансу має один пік при $u_1^{\langle 111 \rangle} = 0,20$. Таким чином, вивчення циклотронного резонансу дає можливість експериментально встановлювати деталі зонної структури напівпровідників, тобто визначати величину ефективної маси у різних напрямках. Зрозуміло, що накладення умови $\omega_C \tau_m > 1$ існування резонансного поглинання додає обмеження до умов проведення експерименту. Наприклад, для спостереження циклотронного резонансу у НВЧ діапазоні ($\sim 10^1$ – 10^2 ГГц) час релаксації τ_m має бути більшим за 10^{-10} с. Цього можна досягнути лише за гелієвих температур і лише у деяких напівпровідниках із дуже низькою концентрацією домішок для того, щоб розсіювання на домішках (нейт-

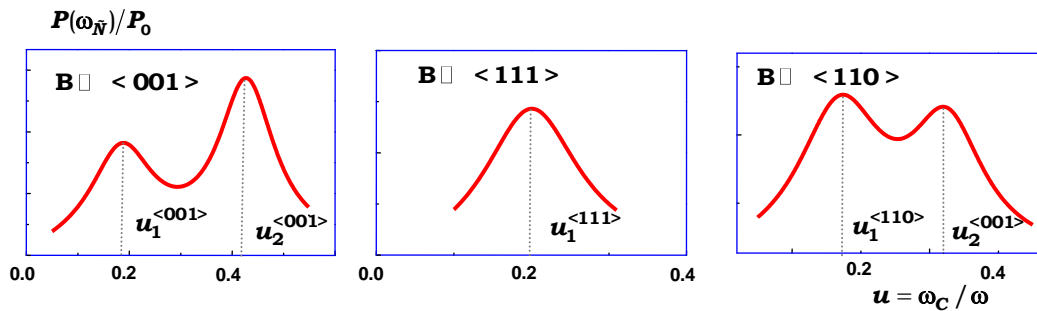


Рис. 15.32. Залежність енергії НВЧ поля, що поглинає зразок n -Si від відношення ω_C/ω для різних орієнтацій магнітного поля

ральних та іонізованих) не сильно знижувало величину τ_m . Іноді замість НВЧ-випромінювання використовують електромагнітні хвилі інфрачервоного діапазону в умовах дії сильного магнітного поля (імпульсного – до 10^2 Тесла, або з використанням надпровідникових магнітів – ~ 10 Тесла). У цьому випадку умови резонансу можна виконати за високих температур – аж до кімнатної.

■ 15.6. Магнітовідбиття вільними носіями та ефект Фарадея

Розглянемо взаємодію плоскої хвилі електромагнітного поля, що характеризується компонентами $\mathbf{E} = \mathbf{E}_0 e^{i(\omega t - \mathbf{q}\mathbf{r})}$ та $\mathbf{B} = \mathbf{B}_0 e^{i(\omega t - \mathbf{q}\mathbf{r})}$, із плазмою носіїв заряду напівпровідника, що знаходиться в сталому магнітному полі. У немагнітному напівпровіднику ці компоненти поля задовольняють рівнянням Максвелла

$$\mathbf{q} \times \mathbf{E} = (\omega / \tilde{n}) \mathbf{B}, \quad (15.133)$$

$$\mathbf{q} \times \mathbf{B} = -(\omega / c) \varepsilon \mathbf{E} + i \tilde{\sigma} \mathbf{E}. \quad (15.134)$$

Виключаючи із цих рівнянь вектор магнітної індукції, отримаємо

$$\left(q^2 - \varepsilon \frac{\omega^2}{c^2} + i \frac{\omega}{c} \tilde{\sigma} \right) \mathbf{E} = (\mathbf{q}\mathbf{E})\mathbf{q}. \quad (15.135)$$

Як ми бачили раніше, вплив сталого магнітного поля спричиняє перенормування компонентів тензора провідності $\tilde{\sigma}$. Але z -компонента цього тензора не залежить від магнітного поля. Цей факт дозволяє покласти $E_z = 0$, не обмежуючи розгляд можливих впливів магнітного поля на оптичні характеристики напівпровідникової плазми.

Для початку розглянемо геометрію Фарадея, коли хвиля електромагнітного поля розповсюджується вздовж напрямку сталого магнітного поля (яке направлено вздовж осі OZ), тобто розглянемо хвилю, в якій $q_x = q_y = 0$. Оскільки електромагнітна хвиля – поперечна, це означає, що права частина (15.135) дорівнюватиме нулю. Таким чином (15.135) переходить у систему лінійних алгебраїчних рівнянь щодо компонент вектора електричного поля хвилі

$$\left(q^2 - \varepsilon \frac{\omega^2}{c^2} + i \frac{\omega}{c} \sigma_{xx} \right) E_x + i \frac{\omega}{c} \sigma_{xy} E_y = 0, \quad (15.136)$$

$$-i \frac{\omega}{c} \sigma_{xy} E_x + \left(q^2 - \varepsilon \frac{\omega^2}{c^2} + i \frac{\omega}{c} \sigma_{xx} \right) E_y = 0 \quad (15.137)$$

із компонентами тензора провідності (15.120–15.121). Ця однорідна система рівнянь має нетривіальні розв'язки лише за умови рівності її детермінанту нулю, тобто

$$(q^2 - \varepsilon q_0^2 + i q_0 \sigma_{xx})^2 + (i q_0 \sigma_{xy})^2 = 0, \quad (15.138)$$

де $q_0 = \omega/c$. Дисперсійне рівняння (15.138) має два розв'язки

$$q_{\pm}^2 = \varepsilon q_0^2 - i q_0 \sigma_{xx} \pm q_0 \sigma_{xy}. \quad (15.139)$$

Далі зручно ввести ефективну електропровідність (див. (15.122) та (15.123))

$$\sigma_{\mp} = \sigma_{xx} \mp i \sigma_{xy}. \quad (15.140)$$

Тоді отримуємо дисперсійне рівняння для право- та лівополяризованої хвилі

$$\frac{q_{\pm}^2}{q_0^2} = \varepsilon - i \frac{\sigma_{\mp}}{q_0}. \quad (15.141)$$

Підставляючи (15.140) та використовуючи явний вигляд компонент тензора ефективної електропровідності (15.122–15.123), отримуємо закон дисперсії коливань зв'язаних збуджень електромагнітної хвилі та напівпровідникової плазми у геометрії Фарадея

$$q_{\pm}^2 (c/\omega)^2 = \varepsilon - \frac{i c \sigma_0 (\omega \tau_m)^{-1}}{\tau_m^{-1} + i (\omega \mp \omega_C)}. \quad (15.142)$$

Ліва частина рівняння являє собою квадрат показника заломлення середовища, де розповсюджується хвиля. Якщо показник заломлення дорівнює нулю, то відбувається повне відбиття світла. Цей ефект називається *плазмовим відбиттям*. Він може спостерігатись на частотах, що задовольняють рівнянню

$$\varepsilon - \frac{i c \sigma_0 (\omega \tau_m)^{-1}}{\tau_m^{-1} + i (\omega \mp \omega_C)} = 0. \quad (15.143)$$

Для розв'язку цього рівняння введемо ефективну плазмову частоту $\bar{\omega}_p = \omega_p / \sqrt{\varepsilon} = \sqrt{c\sigma_0 / \varepsilon\tau_m}$. Тоді, використовуючи наближення $\tau_m^{-1} \ll |\omega \mp \omega_C|$, отримаємо із (15.143) рівняння

$$1 - \frac{\bar{\omega}_p^2}{\omega(\omega \mp \omega_C)} = 0, \quad (15.144)$$

розв'язком якого є

$$\omega = \bar{\omega}_p \left(\sqrt{1 + \left(\frac{\omega_C}{2\bar{\omega}_p} \right)^2} \pm \frac{\omega_C}{2\bar{\omega}_p} \right). \quad (15.145)$$

Звідси, зокрема випливає, що у слабких магнітних полях, коли $(\omega_C / 2\bar{\omega}_p)^2 \ll 1$, плазмове відбиття відбувається на зсунутій частоті

$$\omega - \bar{\omega}_p = \pm \frac{\omega_C}{2} + \frac{1}{8} \frac{\omega_C^2}{\bar{\omega}_p}, \quad (15.146)$$

тобто край плазмового відбиття в магнітному полі зсувається для хвилі із лівою круговою поляризацією – у довгохвильовий бік, а для електромагнітної хвилі із правою круговою поляризацією – у короткохвильовий бік. Наприклад, у *n*-InSb, що в магнітному полі, $B = 2,5$ Тесла для концентрації електронів $n \approx 10^{18} \text{ см}^{-3}$ величина зсуву становить $0,65 \cdot 10^{13} \text{ с}^{-1}$. Оскільки у слабких магнітних полях, як випливає із (15.146), зсув магнітоплазмового відбиття дорівнює половині циклотронної частоти, то із вимірювань краю магнітоплазмового відбиття можна отримати відношення ефективної маси носіїв до маси вільного електрона

$$\frac{m^*}{m_e} = \frac{1}{2} \frac{\omega_C}{\omega - \bar{\omega}_p}. \quad (15.147)$$

Наприклад, для вищезазначених умов експерименту маємо, що ефективна маса електронів в InSb $m^* = 0,033m_e$, що суттєво відрізняється від значень, отриманих із вимірювань ефекту Холла. Різниця зумовлена непараболічністю зони провідності в InSb. З іншого боку, оскільки $\bar{\omega}_p$ визначається часом релаксації, то, вимірюючи частотний зсув магнітоплазмового відбиття, можна знайти також час релаксації. Для обговорюваних умов експерименту час релаксації імпульсу дорівнює $2,8 \cdot 10^{-13} \text{ с}$.

Оскільки фазова швидкість визначається співвідношенням $v_{ph} = \omega/q$, то із рівняння безпосередньо випливає, що фазова швидкість хвиль лівої і правої кругової поляризації, що розповсюджуються у напівпровіднику в магнітному полі, є різними. З іншого боку, лінійно поляризована електромагнітна хвиля являє собою суперпозицію двох хвиль кругової поляризації. Таким чином, якщо через зразок товщиною d , що перебуває в магнітному полі, проходить світло лінійної поляризації, то на виході зі зразку виявляється також лінійно-поляризоване світло. Але, оскільки право-, та лівополяризовані компоненти цієї хвилі розповсюджувалися із різною фазовою швидкістю, площина поляризації лінійно-поляризованої хвилі на виході виявиться повернутою на деякий кут ϑ_F , що залежить як від типу матеріалу, величини магнітного поля, так і від товщини зразку. Таке повертання площини поляризації лінійно поляризованого світла при проходженні через зразок напівпровідника, який перебуває під дією сталого магнітного поля, становить суть ефекту Фарадея. Кут повороту площини поляризації дорівнює середньому значенню кутів повороту векторів електричного поля право- та лівополяризованих хвиль, тобто

$$\vartheta_F = \frac{\vartheta_+ - \vartheta_-}{2}, \quad (15.148)$$

або

$$\vartheta_F = \frac{q_+ - q_-}{2} \cdot d. \quad (15.149)$$

Використовуючи (15.142), отримуємо

$$\vartheta_F = \frac{\omega}{2c} d \left[\left(\varepsilon - i \frac{c}{\omega} \sigma_+ \right)^{1/2} - \left(\varepsilon - i \frac{c}{\omega} \sigma_- \right)^{1/2} \right] \quad (15.150)$$

із провідностями σ_+ та σ_- , що визначаються (15.140). Вважаючи, що $\tau_m^{-1} \ll |\omega \mp \omega_C|$ та $\bar{\omega}_p^2 \ll \omega(\omega \pm \omega_C)$, із (15.150) отримаємо

$$\vartheta_F = \frac{\omega}{2c} \sqrt{\varepsilon} d \frac{\bar{\omega}_p^2 \omega_C}{\omega(\omega^2 - \omega_C^2)}. \quad (15.151)$$

Часто цю формулу переписують у вигляді, зручному для оцінок експериментальних вимірювань,

$$\vartheta_F = \frac{15.1^\circ}{\sqrt{\varepsilon}} \cdot \frac{n}{10^{18} \text{ см}^{-3}} \cdot \left(\frac{\lambda}{10 \text{ мкм}} \right)^2 \frac{B}{1 \text{ Тесла}} \cdot \frac{d}{1 \text{ см}} \cdot \left(\frac{m^*}{m_e} \right)^{-2} \left(1 - \frac{\omega_C^2}{\omega^2} \right)^{-1}, \quad (15.152)$$

де λ – довжина електромагнітної хвилі у вакуумі. Знак кута Фарадея є різним для електронів і дірок і залежить від того, за напрямком чи проти напрямку вектору магнітної індукції розповсюджується світло. Таким чином, кут повороту площини поляризації лінійно поляризованого світла при проходженні напівпровідникового середовища товщиною d , що перебуває у магнітному полі,

$$\vartheta_F = \theta dB, \quad (15.153)$$

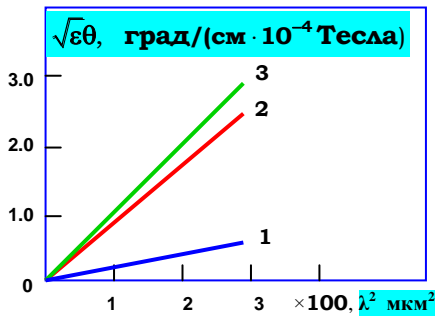


Рис. 15.33. Ефект Фарадея на вільних носіях в n -InSb.
 Пряма 1 відповідає температурі 300 К і концентрації електронів $n = 6 \cdot 10^{16} \text{ см}^{-3}$;
 пряма 2 – $T=300 \text{ К}$, $n = 6,56 \cdot 10^{17} \text{ см}^{-3}$;
 пряма 3 – $T=77 \text{ К}$, $n = 6,4 \cdot 10^{17} \text{ см}^{-3}$

де θ називається коефіцієнтом Верде. Кут повороту вважають позитивним, якщо площина поляризації повертається за годинниковою стрілкою за розповсюдження світла вздовж напрямку магнітного поля. На рис. 15.33 подано поведінку коефіцієнту Верде в електронному антимоніді індію як функції довжини хвилі світла за різних температур і ступенів легування.

Зауважимо, що, як випливає із (15.150), ефект Фарадея визначається механізмом формування відгуку напівпровідника на зовнішнє поле – механізмом формування дисперсії коефіцієнта заломлення або механізмів поглинання світла.

Оскільки у напівпровідниках існує кілька видів поглинання світла, то можна очікувати стільки ж видів повороту площини поляризації, якщо вважати, що кожен із видів поглинання забезпечує свою дисперсію показника заломлення. Вище ми розглянули ефект повороту площини поляризації, обумовлений дисперсією внаслідок поглинання вільними носіями. Якщо дисперсія показника заломлення зумовлена власним міжзонним поглинанням, то йдеться про міжзонні ефекти фарадеєвських поворотів.

Як зазначалось раніше, для зразків із обмеженою довжиною проходження світла характерними є процеси багаторазового відбиття світла від меж розподілу. Для аналізу експериментальних результатів за необхідності урахування процесів багаторазового перевідбиття світла від границь зразку, кут Фарадея можна отримати вимірюванням куту повороту площини поляризації ϑ_F^{exp} і застосуванням формули

$$\vartheta_F = \frac{\vartheta_F^{\text{exp}}}{1 + 2R_\infty^2 e^{-\alpha d} \cos(4\vartheta_F^{\text{exp}})}, \quad (15.154)$$

де α – коефіцієнт поглинання, R_∞ – коефіцієнт відбиття від плоскої межі розподілу напівпростору, що характеризується комплексним коефіцієнтом заломлення $N = n - ik$

$$R_\infty = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2}. \quad (15.155)$$

Крім геометрії Фарадея розглядають також геометрію Фогта, коли хвильовий вектор тестуючого випромінювання направлений перпендикулярно до вектора сталого магнітного поля. У цьому випадку існують ефекти, аналогічні магнетоплазмовому відбиттю та ефекту Фарадея. Але вони виходять за рамки цього курсу і тут не розглядаються.

■ 15.7. Поглинання світла у сильно легованих напівпровідниках

Експериментальні дослідження краю поглинання антимоніду індію показали, що в сильно легованих зразках n -типу край смуги поглинання значно зсувається до короткохвильової області порівняно із краєм поглинання у власному напівпровіднику. У той же час різниця між поглинанням світла власним напівпровідником і матеріалом p -типу була незначною. Пояснення цього ефекту було дано Бурштейном та Моссом: завдяки невеликій щільності станів у зоні провідності відносно невелика кількість електронів може заповнювати зону, що й впливатиме на край поглинання, оскільки для заповнення вільних станів у зоні провідності, що лежать не дуже близько до дна зони, необхідно надати електрону достатню енергію). Цей ефект експериментально спостерігався у багатьох матеріалах. Найбільш яскраво він виявляється в напівпровідниках із малою ефективною масою, тобто переважно у вузькозонних напівпровідниках n -типу. У вузькозонному матеріалі, наприклад в InSb, цей ефект легко спостерігати, оскільки край поглинання знаходиться у довгохвильовій області та достатньо лише незначного зміщення за шкалою енергій для отримання помітного зміщення за шкалою довжин хвиль.

Для обчислення коефіцієнту поглинання світла сильно легованим напівпровідником розглянемо прямі переходи у напівпровіднику n -типу, припускаючи, що енергетичні зони напівпровідника є пара-

болічними та їхній вигляд не змінюється при введенні донорних домішок великої концентрації. Нехай валентна зона повністю зайнята електронами, а стани зони провідності заповнені з ймовірністю $f_e(E)$, що визначається функцією розподілу Фермі

$$f_e(E) = \left(1 + \exp \frac{E - E_F}{kT} \right)^{-1}, \quad (15.156)$$

Коефіцієнт поглинання $\alpha(\omega)$ для вертикальних переходів, згідно із (15.15), пропорційний ймовірності того, що відповідний стан у зоні провідності не зайнятий, тобто

$$\alpha(\omega) = \alpha_0(\omega) [1 - f_e(E_c)], \quad (15.157)$$

де енергія E_c однозначно визначається енергією фотона, що поглинається,

$$\hbar\omega = E_g + \frac{\hbar^2 k^2}{2m_e^*} \left(1 + \frac{m_e^*}{m_h^*} \right). \quad (15.158)$$

Якщо за початок відліку енергії взяти дно зони провідності, то $E_c = \hbar^2 k^2 / 2m_e^*$. Тепер, користуючись (15.157), знайдемо E_c . Позначимо $E_F / kT = \zeta$ і $\alpha_0(\omega) / \alpha(\omega) = P$, тоді із (15.157) маємо

$$e^{E_c / kT} = e^\zeta \frac{P^{-1}}{1 - P^{-1}}, \quad (15.159)$$

звідси отримаємо $E_c = E_F - \ln(P - 1)$. (15.160)

Підставляючи цей вираз до (15.158), отримаємо вираз для енергії фотона, за якої коефіцієнт поглинання зменшується у P разів

$$\hbar\omega = E_g + \left\{ E_F - kT \ln \left[\frac{\alpha_0(\omega)}{\alpha(\omega)} - 1 \right] \right\} \left(1 + \frac{m_e^*}{m_h^*} \right). \quad (15.161)$$

Із цієї формули (уперше отримана Кайзером і Феном), зокрема впливає, що при $\alpha(\omega) \ll \alpha_0(\omega)$ і за припущення, що малими змінами величини $\alpha_0(\omega)$ можна знехтувати, графік залежності $\ln \alpha(\omega)$ від $\hbar\omega$ повинен мати нахил $kT(1 + m_e^* / m_h^*)$ (рис. 15.34). Іншими словами, вимірюючи зсув Бурштейна, можна визначати коефіцієнт поглинання

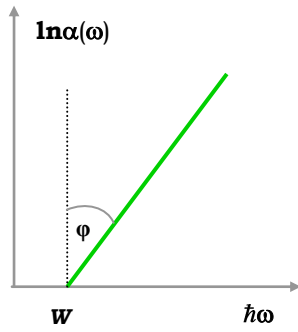


Рис. 15.34. Визначення коефіцієнта поглинання за зсувом Бурштейна

світла сильно легованим напівпровідником. При цьому $tg\varphi = kT(1 + m_e^*/m_h^*)$, а відрізок, що відсікає пряма на осі $\hbar\omega$, визначає енергію

$$W = E_g + E_F - kT(1 + m_e^*/m_h^*) \ln \alpha_0(\omega).$$

Формулу (15.161) можна використовувати для визначення ефективної маси електрона, порівнюючи криву поглинання сильно легованого напівпровідника (із визначеною концентрацією носіїв) із кривою для відносно чистого зразка. Із цих даних можна також отримати досить точне значення енергії Фермі E_F , оскільки відношення

m_e^*/m_h^* зазвичай досить мале, щоб можна було використовувати оціночну величину. Формула

(15.161) враховує лише оптичні переходи з однієї валентної зони. Насправді ситуація є дещо складнішою, оскільки доводиться брати до уваги переходи як із валентної зони легких, так і важких дірок. Співвідношення між зсувами порогу поглинання за нульової температури, що відповідають цим видам переходів, можна усвідомити при розгляді рис. 15.35 а. Для переходів між зоною важких дірок і зоною провідності поглинання світла матиме місце, починаючи з енергії кванту $\hbar\omega_1$, далі спостерігатиметься другий поріг з енергією $\hbar\omega_2 > \hbar\omega_1$. Ці зсунуті краї можна розрізнити експериментально лише за дуже низьких температур. Оскільки основний внесок до коефіцієнту поглинання дають переходи із зони важких дірок і цій зоні відповідає поріг за менших енергій фотона, то на практиці застосування (15.161) викликає лише несуттєві похибки при $\alpha(\omega) \ll \alpha_0(\omega)$. У напівпровідниках *p*-типу пороги поглинання розташовані у зворотному порядку (рис. 15.35 б). Край поглинання для зони легких дірок відповідає меншій енергії кванта $\hbar\omega_2$, ніж край поглинання для зони важких дірок ($\hbar\omega_1 > \hbar\omega_2$), оскільки рівень Фермі перетинає зону важких дірок завжди за більших значень імпульсу, а зважаючи на те, що крутизна зони провідності є великою, нерівність $\hbar\omega_1 > \hbar\omega_2$ виконуватиметься завжди при $m_{h,1}^* > m_{h,2}^*$. За мірою руху рівня Фермі вглиб валентної зони поглинання, пов'язане із переходами між верхньою парою валентних зон, зсувається у бік все більших енергій, лишаючись при цьому обмеженим енергією $\hbar\omega_3$. Низькоенергетична границя переходів між валентними зонами при $T = 0$ визначається величиною $\hbar\omega_4$. Таким чином, у виродженому напів-

провіднику p -типу навіть за підвищених температур пік поглинання, пов'язаний із верхньою парою валентних зон, зі зростанням концентрації дірок зсувається у бік більших енергій. Такий зсув піка не спостерігається у невироджених матеріалах, коли рівень Фермі розташований вище стелі валентної зони (на величину $> 2kT$), хоча коефіцієнт поглинання зростає за зростання концентрації дірок.

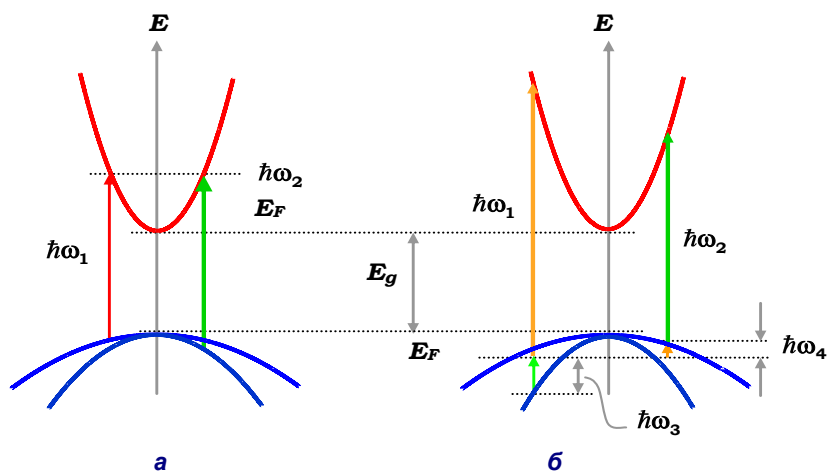


Рис. 15.35. Міжзонне поглинання в сильно легovanому напівпровіднику: a – n -типу, b – p -типу

У випадку непрямих переходів при збільшенні концентрації носіїв також відбувається зсув краю поглинання у бік більших енергій, хоча різкі пороги поглинання не спостерігаються навіть при температурах близьких до нуля, оскільки при кожній енергії фотона можливі процеси поглинання з переходами в деяку область станів зони провідності, як вище так і нижче рівня Фермі.

■ 15.8. Відбивання світла плоскою поверхнею напівпровідника

Раніше було розглянуте поглинання світла напівпровідниковою пластинкою обмеженої товщини. Як видно із (15.6), поглинання світла тонкою пластинкою суттєво визначається коефіцієнтом відбиття світла R поверхнею напівпровідника. Тепер розглянемо, як визначається коефіцієнт відбиття та обговоримо деякі аспекти відбивання світла поверхнею напівпровідника.

Нехай плоска хвиля, що розповсюджується в діелектричному середовищі із показником заломлення N , падає на плоску поверхню напів-

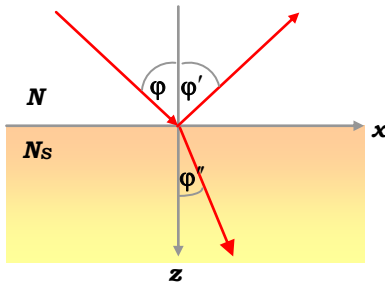


Рис. 15.36. Відбиття та заломлення світла на межі розподілу двох середовищ

провідника із комплексним показником заломлення N' (рис. 15.36). Виберемо систему координат так, щоб хвильовий вектор падаючого опромінення лежав в площині XOZ (площина падіння). Розкладемо амплітуду електричного вектору \mathbf{E}_0 цієї хвилі на компоненти: нормальну E_S та паралельну E_P площині падіння. Тобто, компонента E_S направлена вздовж осі OY , а E_P має компоненти $E_P \cos \varphi$ та $-E_P \sin \varphi$ відповідно в напрямках OX та OZ .

Таким чином, s - та p -компоненти можна записати для:

♦ **падаючої хвилі**

$$E_y = E_S \exp\{i\omega[t - N(x \sin \varphi + z \cos \varphi)/c]\}, \quad (15.162)$$

$$E_x \sec \varphi = -E_z \operatorname{cosec} \varphi = E_P \exp\{i\omega[t - N(x \sin \varphi + z \cos \varphi)/c]\}; \quad (15.163)$$

♦ **відбитої хвилі**

$$E'_y = E'_S \exp\{i\omega[t - N(x \sin \varphi' - z \cos \varphi')/c]\}, \quad (15.164)$$

$$E'_x \sec \varphi' = -E'_z \operatorname{cosec} \varphi' = E'_P \exp\{i\omega[t - N(x \sin \varphi' - z \cos \varphi')/c]\}; \quad (15.165)$$

♦ **заломленої хвилі**

$$E''_y = E''_S \exp\{i\omega[t - N_S(x \sin \varphi'' + z \cos \varphi'')/c]\}, \quad (15.166)$$

$$E''_x \sec \varphi'' = -E''_z \operatorname{cosec} \varphi'' = E''_P \exp\{i\omega[t - N_S(x \sin \varphi'' + z \cos \varphi'')/c]\}. \quad (15.167)$$

Компоненти магнітного вектора хвилі можна знайти за допомогою рівняння

$$-i\omega \mathbf{H} = \operatorname{rot} \mathbf{E}. \quad (15.168)$$

На межі розподілу, де $z = 0$, усі компоненти векторів електричного \mathbf{E} та магнітного \mathbf{H} полів мають однаково залежати від x . Це означає, що

$$N \sin \varphi = N \sin \varphi' = N_S \sin \varphi''. \quad (15.169)$$

Звідси, зокрема впливає закон Снелліуса $\varphi = \varphi'$ або рівність кутів падіння та відбиття та співвідношення між кутами падіння та заломлення

$$\frac{\sin \varphi}{\sin \varphi''} = \frac{N_S}{N}. \quad (15.170)$$

Використаємо граничні умови. Із неперервності тангенціальних компонент електричного поля при переході через межу розподілу отримаємо

$$E_S = E'_S = E''_S, \quad (15.171)$$

$$E_P \cos \varphi + E'_P \cos \varphi' = E''_P \cos \varphi'' = (E_P - E'_P) \cos \varphi. \quad (15.172)$$

Із умови неперервності магнітного поля при переході через межу розподілу аналогічно маємо

$$(E_S - E'_S)N \cos \varphi = E''_S N_S \cos \varphi'', \quad (15.173)$$

$$(E_P + E'_P)N = E''_P N_S. \quad (15.174)$$

Розв'язок рівнянь (15.171-15.174) дає зв'язок між амплітудами падаючого та відбитого й заломленого полів

$$E'_S = \frac{N \cos \varphi - N_S \cos \varphi''}{N \cos \varphi + N_S \cos \varphi''} E_S, \quad (15.175)$$

$$E'_P = \frac{N_S \cos \varphi - N \cos \varphi''}{N_S \cos \varphi + N \cos \varphi''} E_P, \quad (15.176)$$

$$E''_S = \frac{2N \cos \varphi}{N \cos \varphi + N_S \cos \varphi''} E_S, \quad (15.177)$$

$$E''_P = \frac{2N \cos \varphi}{N_S \cos \varphi + N \cos \varphi''} E_P. \quad (15.178)$$

Співвідношення (15.175–15.178) називаються формулами Френеля та відіграють значну роль у дослідження оптики напівпровідників.

Коефіцієнтом відбиття R називається відношення середнього за часом відбитого від поверхні потоку енергії електромагнітного поля до падаючого потоку. Потік енергії електромагнітної хвилі визначається вектором Пойтинга, що усереднюється за періодом

$$I = \frac{1}{2} \operatorname{Re}(\mathbf{E} \cdot \mathbf{H}^*), \quad (15.179)$$

або, використовуючи рівняння Максвелла,

$$I = \frac{c}{8\pi} \sqrt{\varepsilon} |E|^2. \quad (15.180)$$

Тоді для коефіцієнту відбиття отримаємо

$$R = \frac{|E'|^2}{|E|^2}. \quad (15.181)$$

Використовуючи (15.175–15.176), отримаємо коефіцієнт відбиття s - та p -поляризованої хвилі, відповідно

$$R_S = \frac{|N \cos \varphi - N_S \cos \varphi''|^2}{|N \cos \varphi + N_S \cos \varphi''|^2}, \quad (15.182)$$

$$R_P = \frac{|N_S \cos \varphi - N \cos \varphi''|^2}{|N_S \cos \varphi + N \cos \varphi''|^2}. \quad (15.183)$$

У простому випадку нормального падіння маємо

$$R_S = R_P = \frac{|N - N_S|^2}{|N + N_S|^2}. \quad (15.184)$$

Якщо частота світла знаходиться в області, де напівпровідник не поглинає, то коефіцієнт відбиття, як випливає із (15.184), визначається через показники заломлення напівпровідника n_2 і зовнішнього середовища n_1

$$R_S = R_P = \frac{(n_2 - n_1)^2}{(n_2 + n_1)^2}. \quad (15.185)$$

Інший спеціальний випадок відповідає таким кутам падіння та заломлення, коли $\varphi + \varphi'' = \pi/2$, оскільки тоді $R_P = 0$. При цьому $\operatorname{tg} \varphi = N_S / N$. У випадку непоглинаючого середовища кут φ дійсний і має спеціальну назву – кут Брюстера. Коли кут $\varphi + \varphi''$ проходить через $\pi/2$, фаза відбитої хвилі різко змінюється на протилежну, проходячи при цьому через нуль.

У важливому випадку відбиття світла межею розподілу між повітрям та провідним середовищем необхідно покласти $N = 1$ і $N_S = n - ik$ з n – коефіцієнтом заломлення та k – коефіцієнтом екстинції, що характеризує дисипативні процеси. Тоді, використовуючи змінні

$$\alpha^2 - \beta^2 = n^2 - k^2 - \sin^2 \varphi, \quad (15.186)$$

$$\alpha\beta = nk, \quad (15.187)$$

у термінах яких

$$N_S \cos \varphi'' = \alpha - i\beta, \quad (15.188)$$

для коефіцієнтів відбиття s - та p -поляризованих хвиль

$$R_S = \left| \frac{\cos \varphi - N_S \cos \varphi''}{\cos \varphi + N_S \cos \varphi''} \right|^2 = \frac{(\alpha - \cos \varphi)^2 + \beta^2}{(\alpha + \cos \varphi)^2 + \beta^2} \quad (15.189)$$

та

$$R_p = \frac{(\alpha - \sin \varphi \operatorname{tg} \varphi)^2 + \beta^2}{(\alpha + \sin \varphi \operatorname{tg} \varphi)^2 + \beta^2} R_s. \quad (15.190)$$

Загалом кажучи, плоскополяризована хвиля за відбиття поляризується еліптично. Для напівпровідників майже у всіх цікавих випадках величина n становить кілька одиниць, а k – дорівнює одиниці або є меншою. Тому $n^2 \gg k^2$ і $n^2 \gg \sin^2 \varphi$. Це означає, що майже завжди $\alpha \approx n$, $\beta \approx k$. Таким чином, R_s є функцією, що повільно змінюється за зростання кута падіння φ . R_s збільшується від типової для нормального падіння величини $\sim 0,3$ до 1 – за ковзного падіння. За нормального падіння коефіцієнт відбиття p -поляризованого світла має таке саме значення, як і R_s , але за збільшення кута падіння R_p демонструє немонотонну поведінку, і при $\alpha = \sin \varphi \operatorname{tg} \varphi$ проходить через мінімум, а за подальшого зростання кута φ швидко зростає та досягає одиниці за ковзного падіння $\varphi = 90^\circ$. Для значень коефіцієнту заломлення в інтервалі від 3 до 7 кут мінімального відбиття лежить у межах від 71° до 85° . За фіксованого кута падіння, згідно із (15.189–15.190) залишаються дві невідомі величини α та β . Виконуючи ще будь-які два вимірювання, можна визначити ці величини, що означає \square коефіцієнти заломлення та екстинції. Хоча вираз (15.129) є простішим для аналізу, дуже повільна зміна коефіцієнту R_s не дозволяє отримувати достатньо точних результатів для n та k . З іншого боку, швидка зміна коефіцієнту R_p в області кутів $\varphi = \arctg n$ дозволяє вимірювати n та k досить точно. Ще кращі результати можна отримати, вимірюючи відношення R_p / R_s .

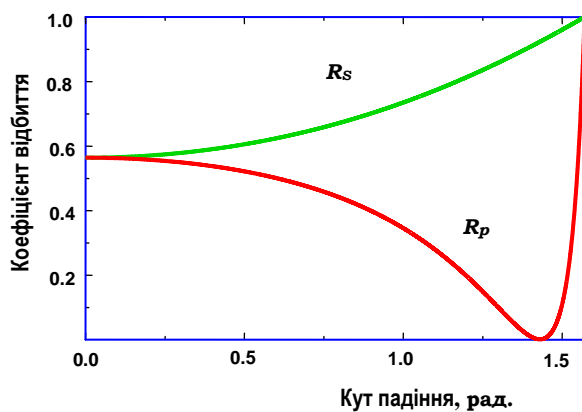


Рис. 15.37. Залежність коефіцієнтів відбиття від кута падіння

Дійсно, вимірювання цієї величини значно спрощує обчислення, а точність підвищується, оскільки величина сильніше залежить від кута падіння (порівн. рис. 15.37 та 15.38). Крім того, у цьому випадку не має необхідності проведення вимірювань абсолютних значень коефіцієнтів відбиття. Такий метод експериментального визначення коефіцієнтів заломлення та екстинції по-

плягає в аналізі сімейства кривих залежності R_p / R_s від n та k , що побудовані для трьох або двох кутів падіння. Ці кути мають вибиратись так, щоб один із них був трохи меншим, а інший – трохи більшим за кут мінімального відбиття.

Третій кут вибирають так, щоб він лежав між двома першими, тобто він мав дуже близьке значення до кута мінімального відбиття. Результати вимірювань, виконані за двох значень кута падіння, дозволяють визначити тільки по одному значенню величин n та k . Якщо ж вимірювання виконувались за трьох кутів падіння, то можна усереднити отримані значення n та k і тим самим різко підвищити точність.

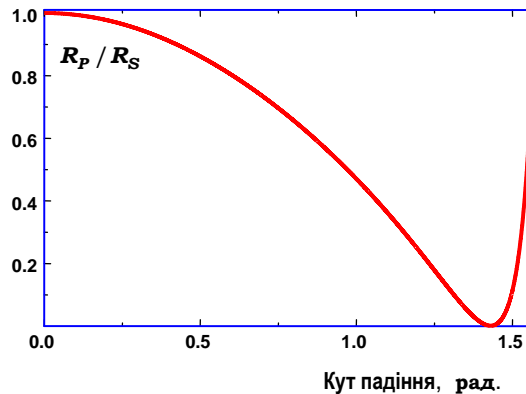


Рис. 15.38. Залежність R_p / R_s від кута падіння

■ 15.9. Дисперсія діелектричної функції. Співвідношення Крамерса-Кроніга

При розгляді дисперсії зручно розглядати внески до діелектричної функції (проникності), обумовлені зв'язаними та вільними електронами окремо. Якщо в кристалі внесок дають лише стани зв'язаних електронів, то кристал являє собою непровідний діелектрик. У метали ж домінуючу роль відіграють вільні електрони. У напівпровіднику істотний внесок дають обидва види електронів: перші зумовлюють інтенсивне поглинання з боку коротких хвиль від краю фундаментального поглинання, а в довгохвильовій області стає суттєвим внесок до поглинання вільними носіями.

Прийнятне феноменологічне пояснення залежності оптичних констант діелектрика від довжини хвилі або частоти можна отримати, виходячи із простої класичної теорії Лоренца, де тверде тіло розглядається як сукупність осциляторів, що здійснюють вимушені коливання під дією випромінювання. Розглянемо поляризацію, що виникає завдяки зміщенню зв'язаних електронів з їх положення рівноваги під дією електричного поля падаючої електромагнітної хвилі. Вважатимемо, що повертаюча сила пропорційна зміщенню x , а гальмівна сила –

пропорційна швидкості: повертаюча сила дорівнює $m\omega_0^2 x$, а гальмівна сила – $m\gamma(dx/dt)$. Тоді рівняння руху електрона можна записати у формі

$$m \frac{d^2 x}{dt^2} + m\gamma \frac{dx}{dt} + m\omega_0^2 x = -eE_x e^{i\omega t}, \quad (15.191)$$

де $E e^{i\omega t}$ – електрична компонента хвилі зовнішнього електромагнітного поля, що розповсюджується із частотою ω . Розв'язок цього рівняння зручно знайти після проведення попереднього Фур'є-перетворення. Як результат отримуємо, що координата електрона x змінюється за синусоїдальним законом із частотою прикладеного поля, а його комплексна амплітуда задається виразом

$$x = \frac{-eE_x/m}{-\omega^2 + \omega_0^2 + i\omega\gamma}. \quad (15.192)$$

Скористаємось рівнянням Максвелла

$$\varepsilon \mathbf{E} = \mathbf{E} + 4\pi \mathbf{P}. \quad (15.193)$$

Для одновимірної системи маємо

$$\varepsilon = 1 + 4\pi P_x / E_x. \quad (15.194)$$

Оскільки поляризацію одиниці об'єму можна представити у вигляді (N – концентрація електронів)

$$\mathbf{P} = Ne \langle \mathbf{x} \rangle, \quad (15.195)$$

то, вважаючи, що середні зміщення електронів зі стану рівноваги $\langle \mathbf{x} \rangle$ в одновимірному випадку задаються (15.192), отримаємо

$$\varepsilon(\omega) = 1 + \frac{4\pi Ne^2/m}{-\omega^2 + \omega_0^2 + i\omega\gamma}. \quad (15.196)$$

Запишемо діелектричну функцію через коефіцієнти заломлення n та екстинції k

$$\varepsilon = (n - ik)^2, \quad (15.197)$$

отримаємо

$$n^2 - k^2 - 1 = \left(\frac{4\pi Ne^2}{m} \right) \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \omega^2 \gamma^2}, \quad (15.198)$$

$$2nk = \left(\frac{4\pi Ne^2}{m} \right) \frac{\omega\gamma}{(\omega_0^2 - \omega^2)^2 + \omega^2 \gamma^2}. \quad (15.199)$$

Аналогічні формули можна отримати й за послідовного розгляду квантово-механічної задачі про реакцію електронної підсистеми на дію зовнішнього електромагнітного поля. Таким чином, ці формули досить до-

бре описують дисперсію оптичних параметрів напівпровідників поблизу резонансів. Наприклад, у сполук зі значною часткою іонного зв'язку завжди існують інтенсивні смуги поглинання у далекій інфрачервоній області. Значне поглинання викликає сильне відбиття, що часто досягає величини 0,9 у вузькому спектральному інтервалі (рис. 15.39), який часто

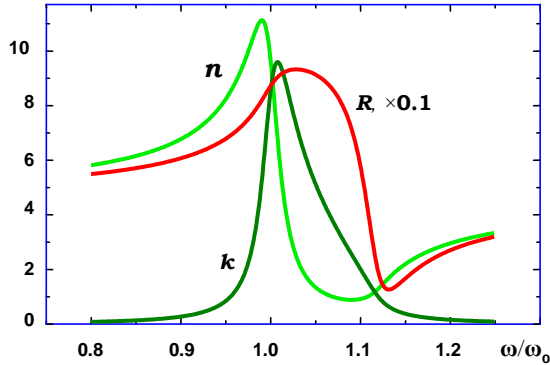


Рис. 15.39. Коефіцієнти поглинання R , заломлення n , та екстинції k у смугі залишкових променів для InSb

називають смугою залишкових променів. Отримані формули показують, що в околі частоти $\omega = \omega_0$ спостерігається максимум поглинання. У цій області частот n швидко збільшується зі зменшенням частоти ω . За подальшого зменшення частоти величина n проходить через максимум, а далі асимптотично зменшується за мірою віддалення частоти від $\omega = \omega_0$. При цьому резонансна частота ω_0 за сут-

тністю являє собою ту частоту, за якої провідність $2nk\omega$ досягає максимуму. Коефіцієнт затухання γ можна визначити з умови: коли провідність дорівнює половині своєї максимальної величини, зсув частоти від резонансу становить $1/2\gamma$. Таким чином γ є шириною резонансної кривої на напіввисоті. При $\omega \rightarrow 0$ із (15.199) видно, що поглинання є малим. З іншого боку, із (15.198) видно, що показник заломлення прямує до сталої (довгохвильової) величини, яка задається формулою

$$n_0^2 = 1 + \left(\frac{4\pi N e^2}{m \omega_0^2} \right) \quad (15.200)$$

Коли йдеться про **показник заломлення матеріалу**, найчастіше мають на увазі саме цю величину. Для довжин хвиль набагато більших, ніж ті, що відповідають смугі поглинання, величина k стає дуже малою порівняно з n . Це означає, що для показника заломлення можна записати

$$n^2 - 1 = \left(\frac{4\pi N e^2}{m} \right) \cdot \frac{1}{(\omega_0^2 - \omega^2)} \quad (15.201)$$

Криві дисперсії більшості матеріалів дуже добре слідують цьому закону, тому кращий метод визначення показника заломлення n_0 полягає в екстраполяції залежності $(n^2 - 1)^{-1}$ від λ^{-2} .

Оскільки дисперсія показника заломлення та коефіцієнту екстинції визначаються дисперсією діелектричної функції (15.197), визначення частотних властивостей діелектричної функції є важливою проблемою оптики напівпровідників. Реальна та уявна частини діелектричної функції пов'язані вельми універсальними співвідношеннями, що віддзеркалюють принцип причинності у фізичних явищах. Математично цей принцип виражається у загальному співвідношенні між індукцією та електричним полем

$$\mathbf{D}(t) = \mathbf{E}(t) + \int_0^{\infty} f(\tau)\mathbf{E}(t - \tau)d\tau. \quad (15.202)$$

Оскільки будь-яке змінне поле можна звести до сукупності монохроматичних компонентів, в яких залежність всіх величин від часу задається множителем $e^{-i\omega t}$, то для таких полів зв'язок між \mathbf{D} та \mathbf{E} має вигляд

$$\mathbf{D}(\omega) = \varepsilon(\omega)\mathbf{E}(\omega), \quad (15.203)$$

де діелектрична функція визначається як

$$\varepsilon(\omega) = 1 + \int_0^{\infty} f(\tau)e^{i\omega\tau}d\tau. \quad (15.204)$$

Функція $\varepsilon(\omega)$, взагалі-то є комплексною, тобто

$$\varepsilon(\omega) = \varepsilon'(\omega) + i\varepsilon''(\omega). \quad (15.205)$$

Із визначення (15.204) безпосередньо випливає, що

$$\varepsilon(-\omega) = \varepsilon^*(\omega). \quad (15.206)$$

Виділимо у цьому співвідношенні дійсну та уявну частини

$$\varepsilon'(-\omega) = \varepsilon'(\omega), \quad (15.207)$$

та

$$\varepsilon''(-\omega) = -\varepsilon''(\omega). \quad (15.208)$$

Отже $\varepsilon'(\omega)$ є парною, а $\varepsilon''(\omega)$ – непарною функцією частоти. Із визначення (15.204), зокрема випливає, що у всій верхній напівплощині комплексної площини ($\omega = \omega' + i\omega''$) діелектрична функція $\varepsilon(\omega)$ є однозначною функцією, що ніде не обертається на нескінченність: у верхній напівплощині ця функція не має особливих точок. Дійсно, при $\omega'' > 0$ підінтегральна функція у (15.204) містить експоненційний множник $e^{-\omega''\tau}$, що швидко зменшується за зростання τ , а оскільки функція $f(\tau)$ обмежена у всій області інтегрування, то інтеграл є збіжним. Функція $\varepsilon(\omega)$ не має особливих точок і на осі $\omega'' = 0$, за винятком, можливо, лише початку координат. Далі виведемо формулу, що пов'язує уявну та

дійсну частину діелектричної функції¹. Виберемо будь-яке позитивне дійсне значення $\omega = \omega_0$ і проінтегруємо вираз $\varepsilon(\omega)/(\omega - \omega_0)$ за контуром (рис. 15.40). Контур йде вздовж усієї дійсної осі, огинаючи точку $\omega = \omega_0$, а також точку $\omega = 0$, якщо вона є полюсом функції $\varepsilon(\omega)$ зверху. Контур

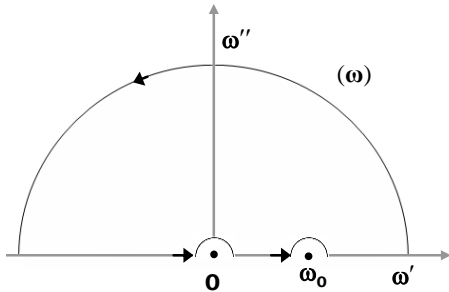


Рис. 15.40. Контур інтегрування при виведенні співвідношень Крамерса–Кроніга

замикається напівколом нескінченно великого радіусу. На нескінченності $\varepsilon \rightarrow 0$, тому функція $\varepsilon(\omega)/(\omega - \omega_0)$ прямує до нуля швидше за $1/\omega$. Це означає, що інтеграл

$$I = \int_C d\omega \frac{\varepsilon(\omega)}{\omega - \omega_0} \quad (15.209)$$

є збіжним. Оскільки функція $\varepsilon(\omega)$ не має особливостей у верхній напівплощині, крім точки $\omega = 0$, а точка $\omega = \omega_0$

як і $\omega = 0$ виключена з області інтегрування, то функція $\varepsilon(\omega)/(\omega - \omega_0)$ аналітична всередині контуру C , й інтеграл (15.209) дорівнює нулю. Вважатимемо спочатку, що $\varepsilon(\omega)$ не має особливостей у точці $\omega = 0$. Інтеграл за нескінченно віддаленим напівколом також обертається на нуль. Точку ж $\omega = \omega_0$ обійдемо за колом нескінченно малого радіусу $r \rightarrow 0$. Обхід відбувається за годинниковою стрілкою та дає до інтеграла внесок, що дорівнює $-i\pi\varepsilon(\omega_0)$. Якщо $\varepsilon(0)$ – обмежена величина, то обхід початку координат є зайвим, й інтегрування вздовж всієї дійсної осі дає

$$\lim_{r \rightarrow 0} \left\{ \int_{-\infty}^{\omega_0 - r} d\omega \frac{\varepsilon(\omega)}{\omega - \omega_0} + \int_{\omega_0 + r}^{\infty} d\omega \frac{\varepsilon(\omega)}{\omega - \omega_0} \right\} - i\pi\varepsilon(\omega_0) = 0. \quad (15.210)$$

Оскільки перший доданок у лівій частині рівняння є визначенням інтегралу у сенсі головного значення, то із (15.210) маємо

$$i\pi\varepsilon(\omega_0) = P \int_{-\infty}^{\infty} d\omega \frac{\varepsilon(\omega)}{\omega - \omega_0}, \quad (15.211)$$

де $P \int d\omega \dots$ – інтеграл у сенсі головного значення. Змінна інтегрування пробігає тут лише дійсні значення. Перепозначимо змінну інтегрування як ξ для заданого дійсного значення величини ω_0 (що позна-

¹ Читачам, хто не цікавиться математичними деталями, це виведення можна опустити.

чимо як ω) і запишемо функцію $\varepsilon(\omega)$ дійсної змінної $\varepsilon(\omega) = \varepsilon'(\omega) + i\varepsilon''(\omega)$. Виділяючи реальну та уявну частини у (15.211), отримуємо

$$\varepsilon'(\omega) = \frac{1}{\pi} P \int_{-\infty}^{\infty} d\xi \frac{\varepsilon''(\xi)}{\xi - \omega}, \quad (15.212)$$

$$\varepsilon''(\omega) = -\frac{1}{\pi} P \int_{-\infty}^{\infty} d\xi \frac{\varepsilon'(\xi)}{\xi - \omega}. \quad (15.213)$$

Якщо функція $\varepsilon(\omega)$ має полюс у точці $\omega = 0$, поблизу якої $\varepsilon(\omega) = iA/\omega$, то обхід цієї точки за напівколом дає в інтегралі додатковий дійсний внесок $-A/\omega$, що має бути доданий до лівої частини співвідношення (15.211). Відповідно такий доданок з'явиться у (15.213)

$$\varepsilon''(\omega) = -\frac{1}{\pi} P \int_{-\infty}^{\infty} d\xi \frac{\varepsilon'(\xi)}{\xi - \omega} + \frac{A}{\omega}. \quad (15.214)$$

Оскільки показник заломлення та коефіцієнт екстинції визначаються через діелектричну функцію, то із співвідношень Крамерса–Кроніга можна отримати дуже важливий висновок: показник заломлення визначається коефіцієнтом екстинції. Зокрема, у випадку власного напівпровідника, який на низьких частотах демонструє діелектричні властивості ($k \approx 0$), показник заломлення n_0 визначається виразом

$$n_0^2 - 1 = \frac{2}{\pi} \int_0^{\infty} \frac{2nk}{\xi} d\xi. \quad (15.215)$$

■ 15.10. Люмінесценція напівпровідників

Напівпровідники у збудженому стані можуть випромінювати електромагнітні хвилі. Такі процеси мають загальну назву – люмінесценція, і є оберненими до процесів поглинання світла. Взагалі кажучи, люмінесценція є процесом випромінювання світла, надлишковим над тепловим випромінюванням. При цьому це випромінювання має характеризуватися кінцевим часом, що істотно перевищує період коливань електромагнітного поля. Основною відмінністю люмінесценції від інших типів електромагнітного випромінювання є тривалість післясвітіння. У випадку люмінесценції акти поглинання та випромінювання світла розділені проміжними процесами, що спричиняє тривале світіння (із часом $\tau \gg 10^{-15}$ с) після припинення дії збудження.

Залежно від методів утворення збудженого стану, що спричиняють випромінювання фотонів, вирізняють кілька типів люмінесценції напівпровідників:

- ♦ випромінювання світла напівпровідником, що переведений у збуджений стан завдяки поглинанню енергії зовнішнього оптичного випромінювання;
- ♦ катодолюмінісценція, за якої випромінювання світла напівпровідником відбувається завдяки бомбардуванню зразка електронами;
- ♦ електролюмінісценція – випромінювання, що спостерігається під дією на напівпровідник електричного поля;
- ♦ хемілюмінесценція – випромінювання, що відбувається у результаті хімічних реакцій тощо.

У твердих тілах, зокрема у напівпровідниках, можна виділити два принципово різних типи люмінесценції – мономолекулярну та рекомбінаційну.

15.10.1. Мономолекулярне світіння твердих тіл

Мономолекулярним світінням твердих тіл називається ефект люмінесценції, за якого поглинання та випромінювання світла відбувається у домішковому центрі. Спектри люмінесценції твердих тіл відрізняються від відповідних спектрів ізольованих атомів. Ця різниця полягає в тому, що спектри випромінювання твердих тіл мають вигляд широкіх смуг і зазвичай зміщені у довгохвильовий бік від смуги поглинання: випромінювана енергія є меншою за енергію, що поглинається напівпровідником. Такі особливості оптичних властивостей твердих тіл обумовлені взаємодією центру світіння із кристалічним полем. Взаємодія центру з оточенням складається власне із поля кристалічної ґратки, що має кулонівську природу, обмінної взаємодії і взаємодії Ван-дер Ваальса, що мають квантово-механічну природу. У результаті атомні рівні домішок, що виступають у ролі центрів світіння, набувають вигляду казімолекулярних станів. Таким чином, поглинання світла домішковим центром може відбуватись за переходу з основного стану у збуджений, причому електрон попадає не в мінімум збудженого стану. Відбувається безвипромінювальна релаксація електрона, і він переходить у мінімальне незайняте положення поблизу мінімуму збудженого стану. І вже звідти може відбуватись перехід в основний стан із випромінюванням фотону. Схему мономолекулярної люмінесценції у напівпровіднику подана на рис. 15.41. Як видно, спектри поглинання (А) та випромінювання (В) рознесені за частотою та мають ненульову ширину. Зауважимо, оскільки центри випромінювання знаходяться, взагалі кажучи, у різних умовах (для кожного із центрів

кристалічний потенціал трохи відрізняється), то виникає додаткове розширення ліній – це так зване неоднорідне розширення.

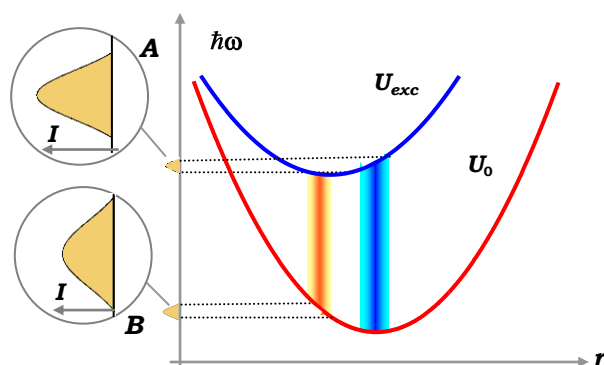


Рис. 15.41. Схема енергетичних рівнів домішкового центру, що служить центром світіння

15.10.2. Міжзонна випромінювальна рекомбінація

Рекомбінація електрон-діркових пар найчастіше супроводжується випромінюванням фотону, що виносить енергію, яка звільняється завдяки рекомбінації. Вважатимемо для визначеності, що збудження напівпровідника відбувається завдяки поглинанню зовнішнього випромінювання відповідної частоти. У напівпровіднику із прямою забороненою зоною поглинання світла супроводжується виникненням вільних електрона та дірки із майже однаковими хвильовими векторами. У непрямозонних напівпровідниках хвильові вектори електрона та дірки відрізняються принаймні на величину \mathbf{k}_0 (дефіцит імпульсу скомпенсований емітуванням або поглинанням фонону). Але загальною характерною рисою процесів люмінесценції є акт попереднього збудження напівпровідника з виникненням електрон-діркової пари (див процес 1 на рис. 15.41). Після акту поглинання світла вільна електрон-діркова пара бере участь у процесах розсіювання. У результаті за час релаксації електрон скочується на дно зони провідності, а дірка піднімається до стелі валентної зони (процес 2 на рисунку). Про такі процеси часто кажуть: електрон-діркова пара термалізується. Характерні часи встановлення такого стану – $\sim 10^{-10} - 10^{-11}$ с. Пряма рекомбінація вільних електрона та дірки у прямозонному напівпровіднику найбільш ймовірна, якщо їхні хвильові вектори будуть однаковими (процес 3 на рисунку). Зрозуміло, оскільки процес випромінювальної

рекомбінації є оберненим до процесу міжзонного поглинання, то спектр випромінювання можна описати формулою, подібною до (15.33)

$$I(\omega) = C\sqrt{\hbar\omega - E_g}. \quad (15.216)$$

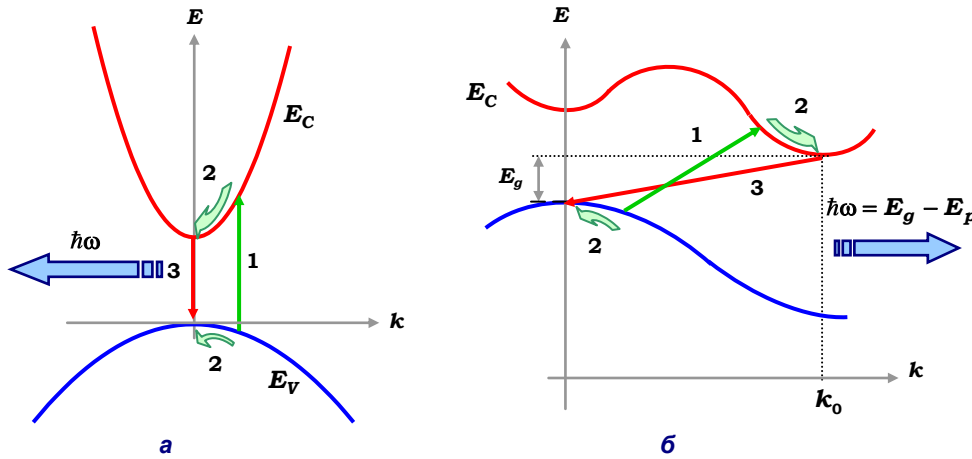


Рис. 15.42. Енергетична схема випромінювальних переходів у напівпровідниках: а – прямо зонному; б – непрямоzonному

Для точнішого обчислення форми спектра фотолюмінесценції за переходів **зона-зона** візьмемо до уваги, що спектр пропорційний добутку числа заповнення електронних станів f_C у зоні провідності та кількості вільних станів із заданою енергією у валентній зоні – $(f_h = 1 - f_V)$. Квазірівноважні функції розподілу електронів f_C і дірок f_h за невеликих інтенсивностей фотозбудження можна апроксимувати функціями розподілу Больцмана $f_C, f_h \sim \exp(-E/kT)$. Тоді спектр фотолюмінесценції можна описувати формулою

$$I_{PL}(\omega) = C\sqrt{\hbar\omega - E_g} \exp\{-(\hbar\omega - E_g)/kT\} \mathfrak{S}(\hbar\omega - E_g), \quad (15.217)$$

де $\mathfrak{S}(x)$ – функція Хевісайда ($\mathfrak{S}(x) = 1, x \geq 0, \mathfrak{S}(x) = 0, x < 0$).

У напівпровіднику з непрямою забороненою зоною поглинання світла відбувається із поглинанням або емітуванням фонону. У такому напівпровіднику за рекомбінації електрон-діркової пари також беруть участь фонони, які забезпечують виконання закону збереження імпульсу. Але найбільш ймовірним процесом при цьому буде емісія фонону, оскільки ймовірність поглинання фонону є в $N_p(N_p + 1)$ разів меншою за ймовірність процесу емісії. Таким чином, енергія фотону, що випромінюється за рекомбінації **зона-зона** у непрямоzonному на-

півпровіднику буде меншою від ширини забороненої зони на величину енергії емітованого фонону E_p , і спектр люмінесценції можна описати формулою

$$I(\omega) = C\sqrt{\hbar\omega - E_g + E_p}. \quad (15.218)$$

Іншими словами, випромінювання за рекомбінації вільних носіїв є пороговим процесом із боку низьких енергій $\hbar\omega \geq E_g$. Зростання температури та енергії збудження спричиняє заповнювання все більш високоенергетичних станів у зоні провідності. Це й стане причиною появи у спектрі люмінесценції короткохвильових "хвостів". Характерний вигляд спектру люмінесценції напівпровідника (n -InSb $n = 1,8 \cdot 10^{17} \text{ см}^{-3}$ при $T = 77 \text{ K}$) подано на рис. 15.43. Видно, що спектр в основному правильно описується моделями, що запропоновані для прямозонних напівпровідників: спектр характеризується порогом із боку низьких частот ($\sim 0,4 \text{ eV}$), має максимум і швидко спадає зі зростанням частоти. Якщо у напівпровіднику можливі як прямі, так і непрямі переходи, його спектр люмінесценції містить дві смуги – одна з яких (довгохвильова) визначається непрямими переходами, а інша (короткохвильова) – прямими.

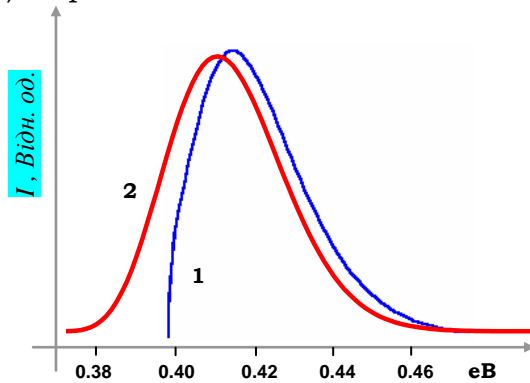


Рис. 15.43. Спектр випромінювання антимоніду індію n -типу за температури 77 К, $n = 1,8 \cdot 10^{17} \text{ см}^{-3}$: 1 – теорія, 2 – експеримент

15.10.3. Люмінесценція за екситонної рекомбінації

За низьких температур у чистих матеріалах при збудженні світлом, енергія якого майже дорівнює ширині забороненої зони, виникають зв'язані стани електронів і дірок – екситони з енергією зв'язку E_{exc} . За рекомбінації таких електрон-діркових пар, тобто при зникненні ек-

ситону випромінюватиметься світло у досить вузькому енергетичному інтервалі. У прямозонних напівпровідниках, таким чином, енергія екситонного випромінювання

$$\hbar\omega = E_g - E_{exc}. \quad (15.219)$$

Оскільки енергія екситону $E_n \sim n^{-2}$, то енергія іонізації збуджених станів екситону зменшується порівняно з енергією іонізації основного стану ($n=1$) як n^2 . Таким чином, випромінювання світла за анігіляції вільного екситону може містити серію вузьких ліній, відстань між якими зменшується як n^2 . Анігіляція екситону може відбуватись за участю ґратки: у процесі випромінювання світла за переходу екситону від одного збудженого стану до іншого можуть відбуватись процеси емісії фононів з енергіями E_p . Наприклад, у напівпровідниках із непрямыми зонами

$$\hbar\omega = E_g - E_{exc} - E_p. \quad (15.220)$$

У прямозонних напівпровідниках випромінювання світла за анігіляції екситону також може відбуватись з емісією одного або кількох фононів. Але зі зростанням кількості частинок, що беруть участь в процесі, ймовірність процесу сильно зменшується. Це означає, що лінії, які відповідають багатофононним процесам, будуть слабкими. Такі лінії у спектроскопії часто називають фононними повтореннями. Дуже часто екситон буває локалізований у деякій потенціальній ямі поблизу дефекту. У такій ситуації йдеться про існування зв'язаного екситону. Дійсно, екситон можна трактувати як збуджені електрон і дірку, що притягуються один до одного завдяки кулонівській взаємодії. Якщо зразок містить невелику кількість домішок у нейтральному стані, що реалізується в умовах низьких температур, то через Ван-дер-Ваальсову взаємодію екситони притягуватимуться до області локалізації домішок.

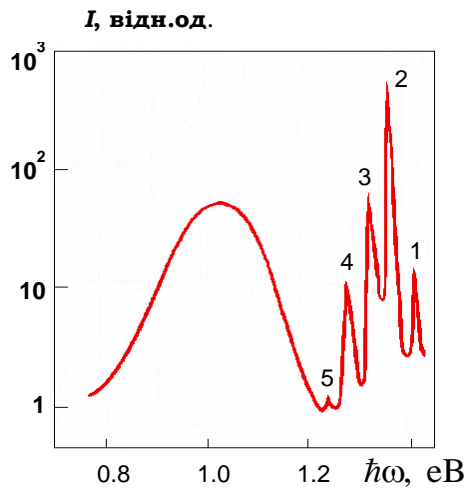


Рис. 15.44. Спектр фотолюмінесценції InP при $T = 6$ К

При цьому найбільш короткохвильова лінія відповідає люмінесценції за рекомбінації вільного екситону. Набір яскраво виражених піків, що зсунуті у довгохвильовий бік, відповідає анігіляції зв'язаного екситону, що супроводжується емісією повздовжніх оптичних фононів з енергією 43 меВ. Смути 3, 4, 5 є фононними повтореннями, що відповідають емісії одного, двох, та трьох фононів, відповідно. Видно, що зі зростанням кількості емітованих фононів амплітуда піків зменшується.

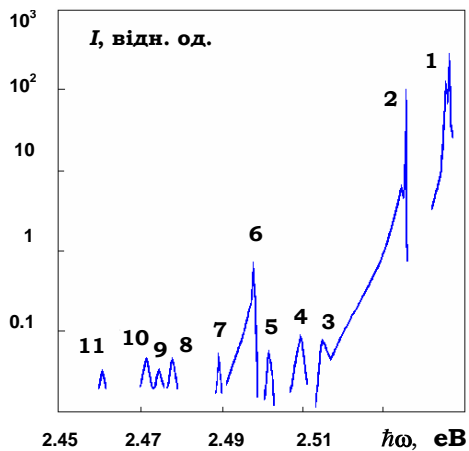


Рис. 15.45. Спектр екситонної люмінесценції чистого CdS при $T = 1.6$ К.

Піки 1 та 2 відповідають безфононій випромінювальній рекомбінації, піки 3-11 – фононним повторенням

Оскільки у результаті для екситону виникатиме самоузгоджена потенціальна яма, то притягання спричинить зменшення енергії екситону: нейтральні домішки будуть захоплювати екситони, і з'являться зв'язані екситони. Спектр випромінювання зв'язаних екситонів містить вузькі лінії і відповідає енергіям, меншим за ті, що спостерігаються за анігіляції вільного екситону. Характерний вигляд спектру екситонної фотолюмінесценції подано на рис. 15.44 для чистого фосфіду індію, де в короткохвильовій області розташовані характерні вузькі екситонні лінії.

Аналогічну поведінку демонструє спектр екситонної люмінесценції чистого кристалу CdS, що вимірювався за температури 1,6 К (рис. 15.45). Тут також спостерігаються безфононні піки зв'язаних екситонів (1 та 2), а також їхні фононні повторення (3-11). При цьому, наприклад піку 3 відповідає процес анігіляції зв'язаного екситону 2 із випромінюванням поперечного акустичного фонону (ТА з енергією 20,6 меВ), а піку 7 – анігіляції екситону 1 із випромінюванням фононів та повздовжнього оптичного фонону (ЛО) з енергією 37,7 меВ.

15.10.4. Люмінесценція за переходів між зоною та домішковим рівнем

У напівпровідниках важливу роль у процесах люмінесценції відіграють різного роду дефекти, енергетичні рівні яких попадають всередину забороненої зони напівпровідника. Такими дефектами можуть бути як домішки, так і різного роду порушення регулярності кристалічної ґратки. Ці дефекти, незалежно від їхньої природи, переважно відіграють роль центрів рекомбінації. І, таким чином, люмінесценція в основному відбувається через такі центри рекомбінації. Кінетику таких переходів ми розглядали раніше. Тепер зупинимось на аналізі спектрів люмінесценції за таких механізмів рекомбінації.

У напівпровіднику можливими є випромінювальні переходи електронів із донорного рівня до валентної зони або із зони провідності на акцепторний рівень. Наприклад, після процесу переходу електрона із

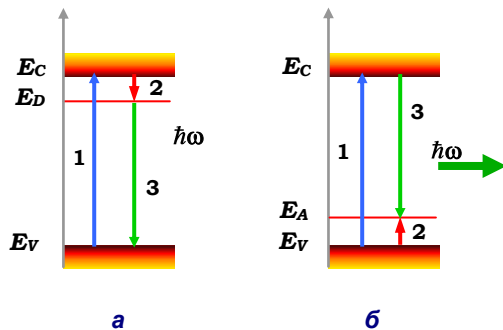


Рис. 15.46. Випромінювальні переходи між зоною та домішковими рівнями

та дірки із валентної зони (процес 3). При цьому необхідною є умова: захоплення дірки із валентної зони хоч і менш ймовірно порівняно із захопленням електрона із зони провідності, але його переріз є досить помітною величиною. У випадку акцепторного напівпровідника, або коли енергетичний рівень рекомбінаційного центру розташований на невеликій відстані від стелі валентної зони, після переходу електрона із валентної зони до зони провідності люмінесценція виникає, коли рекомбінують електрон із зони провідності та дірка, що попередньо була захоплена на рекомбінаційний рівень (процес 3 на рис. 15.46 б). У випадку прямозонних напівпровідників енергія фотону визначається співвідношенням

$$\hbar\omega = E_g - E_i. \quad (15.221)$$

Якщо люмінесценція відбувається у непрямозонному напівпровіднику, то енергія випроміненого фотону визначається співвідношенням

$$\hbar\omega = E_g - E_i - E_p. \quad (15.222)$$

У цих формулах через E_i позначено енергію іонізації домішкового атому або енергію електрона у потенціалі структурного дефекту. Іншими словами, у непрямозонних напівпровідниках за малих концентрацій домішок у переходах типу **зона провідності-акцептор** або **донор-зона провідності** для виконання закону збереження імпульсу бере участь фоновий. Наприклад, у процесах люмінесценції кремнію, легovanому бором, беруть участь поперечний акустичний фоновий ТА з енергією 0.016 еВ і поперечний оптичний фоновий ТО з енергією 0,055 еВ. На рис. 15.47 видно смугу люмінесценції із центральною енергією 1,05 еВ, що утворюється переходом із випромінюванням оптичного фоновий ТО $\hbar\omega = E_g - E_A - E_p(TO)$ та смугу енергії 1,1 еВ, що утворюється пере ходом **зона-акцептор** із випромінюванням поперечного акустичного фоновий $\hbar\omega = E_g - E_A - E_p(TA)$.

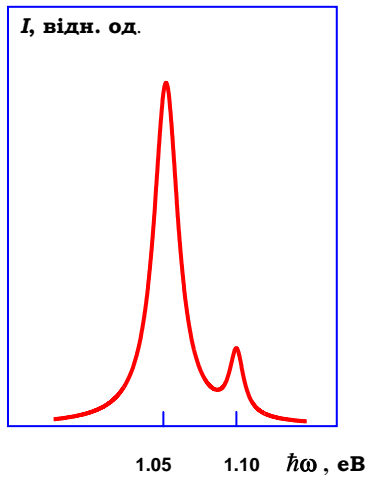


Рис. 15.47. Спектр домішкової випромінювальної рекомбінації у кремнії, легovanому бором при 20 К

За збільшення концентрації домішок у спектрах люмінесценції виникають специфічні зміни. Спочатку завдяки змінам у щільності станів домішкових рівнів смуги люмінесценції розширюються. Зі зменшенням відстані між домішками за досить великих їх концентрацій хвильові функції домішкових рівнів починають перекриватися. Завдяки взаємодії між такими станами та принципу Паулі ці енергетичні рівні розмиваються у домішкові зони. Коли домішкова зона стає настільки широкою, що починає перекриватись із валентною зоною, електрони (якщо домішка донорного типу) або дірки (якщо домішка – акцептор) перестають бути локалізованими на домішках і починають поводитись як вільні квазічастинки. У такому разі кажуть, що відбува-

ється перехід від локалізованого стану у делокалізований, або перехід Мотта.² Крім збільшення щільності домішкових станів, функція розподілу носіїв за збільшення їхньої концентрації стає виродженою: форма смуг люмінесценції відступає від експоненціальної, що відповідає розподілу Больцмана, а все більш стає схожою на розподіл Фермі-Дірака. До того ж, збільшення концентрації домішок спричиняє зсув смуг спектру випромінювання, що пов'язано із додатковою взаємодією, обумовленою багатоелектронними ефектами.

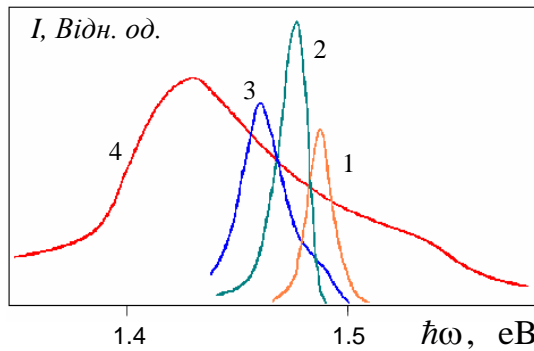


Рис. 15.48. Електролюмінесценція GaAs при $T = 4,2$ К, легованого Zn до різних концентрацій:

1 – $N_A = 3.7 \cdot 10^{17} \text{ см}^{-3}$; 2 – $N_A = 1.9 \cdot 10^{18} \text{ см}^{-3}$;
 3 – $N_A = 1.3 \cdot 10^{19} \text{ см}^{-3}$; 4 – $N_A = 1.25 \cdot 10^{20} \text{ см}^{-3}$

15.10.5. Випромінювальні переходи у донорно-акцепторних парах

У компенсованих напівпровідниках із досить високою концентрацією донорів та акцепторів можуть відбуватись специфічні випромінювальні переходи. Дійсно, якщо енергіями іонізації донорів та акцепторів у кристалі є E_D та E_A , відповідно, то за достатніх їх концентрацій можуть існувати пари донор-акцептор, розділені невеликою відстанню R . Таку пару можна розглядати як нерухомі молекули, занурені в кристал. Кулонівська взаємодія між іонізованими атомами зменшуватиме енергію комплексу на величину

$$\Delta E_{Coul} = \frac{e^2}{\epsilon R}, \quad (15.223)$$

де ϵ – діелектрична проникність кристалу. Таким чином, енергія між енергетичними рівнями у донорно-акцепторній парі

² Зауважимо, що часто перехід Мотта відбувається за концентрацій домішок, менших від тих, що необхідні для перекривання домішкової зони із валентною або зоною провідності, оскільки домішкові рівні розширюються, утворюючи наполовину заповнену зону через виродження за спіном. Тоді, якщо розширення рівня стає більшим за додаткову енергію, необхідну для розміщення двох електронів на одному домішковому рівні даного атому, замість розміщенні їх на рівні атома, домішкова зона набуває ознак зони провідності металу.

$$\hbar\omega = E_g - E_D - E_A + \Delta E_{Coul}. \quad (15.224)$$

Оскільки домішки зазвичай можуть займати дискретні позиції у ґратці (найчастіше це домішки заміщення, що розташовуються у вузлах ґратки), відстані в донорно-акцепторних парах змінюються дискретними порціями. Існує суттєва різниця між кулонівською взаємодією між електроном і діркою, що утворюють екситон, і в донорно-акцепторних парах. Якщо відстань між електроном і діркою в екситонах визначається деталями квантово-механічної взаємодії між носіями в кристалі, то відстань між донором та акцептором визначається в основному геометричними факторами та концентрацією домішок. Оскільки величина енергетичного зазору (15.224) змінюється дискретно, це спричиняє наявність тонкої структури у спектрі люмінесценції. За великих відстаней між донорами та акцепторами (що відбувається за низьких концентрацій домішок) лінії випромінювання перекриваються, утворюючи широкі смуги. Збільшення концентрації домішок спричиняє виникнення у спектрі люмінесценції набору вузьких ліній. Випромінювання донорно-акцепторних пар детально вивчалось у фосфіді галію, легovanому парами Si-S, Se-Si, Si-Zn. Оскільки зазвичай випромінювання донорно-акцепторних пар спостерігається на бінарних сполуках, то можливі два типи геометрії донорно-акцепторних пар, з якими пов'язуються два типи спектрів випромінювання. У системах першого типу як донори, так й акцептори локалізовані у вузлах однієї підґратки. Наприклад, пари S-Si та Se-Si локалізуються у вузлах, де розташовується фосфор, а пара Si-Zn – локалізується у вузлах галію. У системах другого типу донори та акцептори розташовуються у вузлах різних підґраток. Наприклад, у парі S-Zn сірка розташована у позиціях фосфору, а цинк – у вузлах галію. На рис. 15.49–15.50 подано характерний вигляд спектрів випромінювання за рекомбінації донорно-акцепторних пар у системах першого та другого типу, відповідно. Оскільки системи другого типу характеризуються більшою кількістю можливостей взаємного розташування донорів та акцепторів у парі, спектри другого типу мають більшу кількість різких піків, що розділені невеликою енергією.

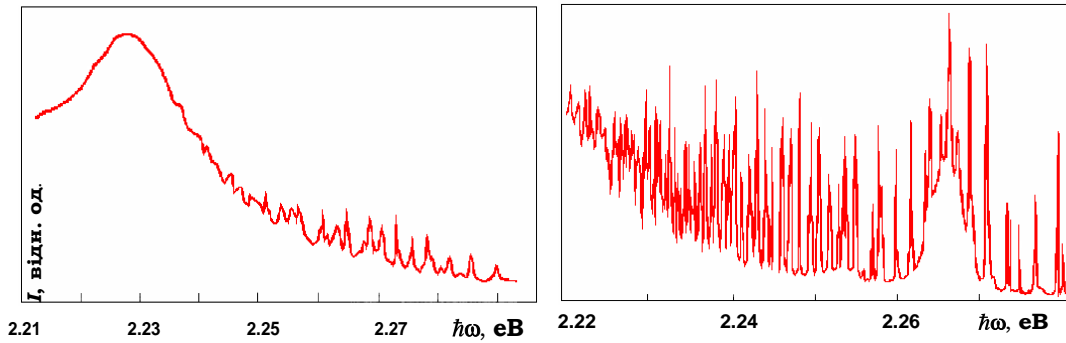


Рис. 15.49. Спектр фотолюмінесценції GaP при $T=1.6$ К за механізмом донорно-акцепторної рекомбінації у парі Si-Te

Рис. 15.50. Спектр фотолюмінесценції GaP при $T=1.6$ К за механізмом донорно-акцепторної рекомбінації у парі S-Mg (S – у позиціях P, Mg – у позиціях Ga)

15.10.6. Еволюція люмінесценції у часі

Випромінювання світла у напівпровідниках пов'язується із наявністю нерівноважних носіїв і, таким чином, актуальною стає проблема релаксації люмінесценції. Проблеми релаксації концентрації надлишкових нерівноважних носіїв заряду розглядалися раніше (див. розд. 10), зокрема обговорювались процеси генерації і рекомбінації носіїв у напівпровідниках. Оскільки інтенсивність випромінювання світла за рекомбінації носіїв пропорційна концентрації надлишкових нерівноважних носіїв, то всі моделі, що використовувались при обговоренні релаксації надлишкової концентрації нерівноважних носіїв можуть бути застосовані й при вивченні еволюції у часі люмінесценції. Тепер розглянемо процеси релаксації люмінесценції, не торкаючись деталей механізмів процесів рекомбінації, що спричиняють ті чи інші залежності часів релаксації.

Завдяки ненульовому часу життя нерівноважного стану напівпровідника, при вимкненні збудження (що викликає нерівноважний стан, (напр., при вимкненні зовнішнього підсвічування) випромінювання із зразку спостерігається деякий час. Оскільки інтенсивність випромінювання пропорційна числу актів рекомбінації за одиницю часу

$$I = -C \frac{dn}{dt}, \quad (15.225)$$

то зміна інтенсивності випромінювання світла із плином часу сильно залежатиме від механізму рекомбінації у напівпровіднику. Наприклад, за люмінесценції за мономолекулярним механізмом, коли процеси по-

глинання і випромінювання світла відбувається на одному і тому самому центрі, зменшення концентрації нерівноважних електронів визначається формулою

$$n = n_0 e^{-t/\tau} \quad (15.226)$$

із початковою концентрацією носіїв n_0 і часом життя збудженого на центрі електрона τ . Таким чином, інтенсивність мономолекулярної люмінесценції релаксує за законом

$$I = I_0 e^{-t/\tau} \quad (15.227)$$

із часом релаксації τ .

У випадку рекомбінаційного випромінювання релаксація люмінесценції визначається як механізмами збудження, так і величиною інжекції. У випадку слабкого оптичного збудження, коли відбувається біполярна оптична генерація носіїв заряду, а концентрація надлишкових нерівноважних носіїв є набагато меншою за концентрацію рівноважних носіїв ($\delta n \ll n_0, p_0$), швидкість рекомбінації пропорційна концентрації надлишкових носіїв, тобто $d(\delta n)/dt = -(1/\tau)\delta n$ (див. (10.14)).

Тоді $\delta n = \delta n_0 e^{-t/\tau}$ і релаксація інтенсивності люмінесценції відбувається за законом, аналогічним (15.227): відбувається мономолекулярне рекомбінаційне світіння. Час релаксації при цьому визначається механізмом рекомбінаційного процесу. Наприклад, при випромінюванні донорно-акцепторних пар час релаксації визначається формулою

$$(\tau_{DA}) \sim \exp[-2(R/a_D)], \quad (15.228)$$

де a_D – борівський радіус донору та вважається, що борівські радіуси донорів та акцепторів приблизно однакові.

У випадку сильної інжекції носіїв, коли $\delta n \gg n_0, p_0$, швидкість рекомбінації, отже й інтенсивність люмінесценції, пропорційна квадрату надлишкових носіїв заряду (див (10.17)), і концентрація надлишкових нерівноважних носіїв зменшується із часом за гіперболічним законом (10.31) $\delta n(t) = \delta n_0 (1 + \gamma \delta n_0 t)^{-1}$ (δn_0 – концентрація надлишкових нерівноважних носіїв у початковий момент часу $t = 0$). Інтенсивність люмінесценції можна описати співвідношенням

$$I = C \frac{\delta n_0^2}{(1 + \gamma \delta n_0 t)^2}. \quad (15.229)$$

Такий тип люмінесценції часто називають бімолекулярним рекомбінаційним світінням.

15.10.7. Генерація світла в напівпровідниках і напівпровідникових структурах

Для збудження люмінесценції у напівпровідниках необхідно створити умови для генерації великої кількості електронно-діркових пар. Як зазначалось вище, одним із методів генерації електрон-діркових пар може бути дія зовнішнього електричного поля. Електричне збудження напівпровідника є найбільш прямим методом, оскільки при цьому відбувається безпосереднє перетворення енергії електричного поля на енергію світлового випромінювання. Але при цьому виникає проблема утримання високої концентрації електронно-діркових пар, яка найбільш ефективно розв'язується при випромінюванні світла із *p-n*-переходу.

15.10.7.1. Випромінювання світла із *p-n*-переходу за прямого зміщення

Пряме зміщення дозволяє забезпечити дуже високу щільність електронно-діркових пар у вузькій області поблизу *p-n*-переходу. При цьому струм зростає експоненційно зі зростанням напруги, що прикладена до переходу. Якщо внутрішній опір діоду невеликий, то вся напруга падає на *p-n*-переході. Тобто в області переходу утворюється сильне поле, що через механізм тунелювання спричиняє ефективну генерацію електрон-діркових пар. У напівпровідникових світловопромінювальних структурах зазвичай обидві області *p*- та *n*-типів провідності сильно легуються, і, таким чином, внутрішній опір переходу зводиться до мінімуму. Для розуміння механізму виникнення люмінесценції із області *p-n*-переходу розглянемо енергетичну діаграму *p-n*-переходу, до якого прикладене пряме зміщення (рис. 15.51)

$$V \geq (E_g + \xi_p) / e. \tag{15.230}$$

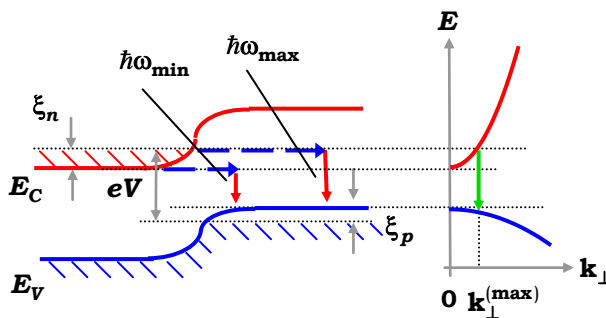


Рис. 15.51. Інжекційна люмінесценція в *p-n*-переході за прямого зміщення *V*

Величина цього зміщення достатня для інжекції електронів до *p*-області. За однакового рівня легування в обох областях, завдяки тому, що ефективна маса електронів є меншою від ефективної маси дірок квазірівень Фермі в *n*-області заходить глибше до зони провідності, ніж квазірівень

Фермі в p -області – углиб валентної зони. Тобто виконується нерівність $\xi_n > \xi_p$. В умовах низьких температур і за зміщення $V = (E_g + \xi_p)/e$ виникає потік електронів до p -області, у той час як дірки блоковані бар'єром, що дорівнює $\xi_n - \xi_p$. Таким чином, у вузькому шарі p -області виникає велика концентрація електрон-діркових пар і можуть ефективно відбуватись процеси рекомбінації із випроміненням фотону³. У моделі прямих параболічних зон мінімальна енергія фотону, що випромінюється за такого процесу,

$$\hbar\omega_{\min} = eV - \xi_n - \xi_p, \quad (15.231)$$

що відповідає переходу між краями зон. Варто зазначити: якщо краї зон розширені експоненційними хвостами (за рахунок ненульової температури), то низькоенергетичний спад спектру випромінювання не матиме різкого обриву. Максимальна енергія фотону, що випромінюються за дуже низьких температур, визначається формулою

$$\hbar\omega_{\max} = eV - \xi_n + \frac{m_e^*}{m_p^*} \xi_p, \quad (15.232)$$

де доданок $-\xi_n + \frac{m_e^*}{m_p^*} \xi_p$ враховує збереження поперечного (до площини переходу) імпульсу. Зрозуміло, що коли збереження імпульсу за сталої енергії досягається за рахунок розсіювання електронів, то максимальна енергія випроміненого фотону

$$\hbar\omega_{\max} = eV. \quad (15.233)$$

У непрямозонних напівпровідниках, наприклад GaAs, для збереження імпульсу випромінюється фонон з енергією E_{ph} , і максимальна енергія фотону визначається величиною

$$\hbar\omega_{\max} = eV - E_{ph}. \quad (15.234)$$

Таким чином, електролюмінесценція з p - n -переходу матиме спектр випромінювання, що зсувається у синій бік за збільшення зміщення V . Крім того, за збільшення зміщення ширина переходу зменшується, що викликає зростання ймовірності тунелювання. Цей факт є причиною різкого зростання інтенсивності випромінювання

$$I \sim \exp\left[-\frac{\text{const}}{E(V)}\right], \quad (15.235)$$

³ Такий механізм випромінювання фотона можна розглядати як процес перенесення електрона, що відбувається у два етапи: на першому – тунелювання електрона із n - до p -області; на другому – рекомбінація електрон-діркової пари з випромінюванням фотону. З огляду на це, в аспекті явищ перенесення такі процеси часто називають тунелюванням за участі фотона.

де електричне поле в p - n -переході $E = (E_g - V)/x(V)$ пропорційне $(E_g - V)^{1/2}$ у випадку різкого переходу та $(E_g - V)^{3/2}$ – у випадку плавного.

15.10.7.2. Світлодіоди

На основі ефектів електролюмінісценції виготовляють широкий спектр електролюмінісцентних приладів, найпоширенішими з яких є світлодіоди (в англійській літературі часто застосовується аббревіатура LED – light emitting diode). Інжекція неосновних носіїв в p - n -переході, що включений у пряму напругу, спричиняє високоефективне перетворення енергії зовнішнього електричного поля на випромінювання світла. Процеси спонтанного випромінювання світла можуть відбуватись у напівпровідниках із прямими та із непрямыми переходами завдяки процесам випромінювальної рекомбінації. Випромінювальна рекомбінація має місце, якщо у напівпровідник інжектують надлишкові носії заряду, внаслідок чого виникає нерівноважна ситуація. Електрони та дірки можуть або спонтанно рекомбінувати з випромінюванням фотонів, або рекомбінувати через домішково-дефектні стани із випромінюванням фотонів, або брати участь у безвипромінювальних процесах рекомбінації (процеси Оже). Ефективність генерації випромінювання залежить від співвідношення між часами життя надлишкових носіїв за випромінювальних процесів рекомбінації – τ_r і безвипромінювальних процесів рекомбінації – τ_0 . Для характеристики ефективності процесів випромінювання світла вводять внутрішню квантову ефективність як відношення числа фотонів, що генерується всередині напівпровідника, у розрахунку на один інжектований електрон: ця величина визначається відношенням оберненого часу життя надлишкових носіїв за випромінювальних рекомбінаційних процесів до повного оберненого часу життя надлишкових носіїв у напівпровіднику

$$\eta_i = \frac{1/\tau_r}{1/\tau_r + 1/\tau_0} = \frac{\tau_0}{\tau_0 + \tau_r}. \quad (15.236)$$

Ясно, що для достатньої ефективності випромінювання в електролюмінісцентних приладах необхідно, щоб якість кристалу була достатньо високою, тобто виконання умови

$$\tau_0 \gg \tau_r. \quad (15.237)$$

При цьому, незважаючи на те, що внутрішня квантова ефективність процесів випромінювання може досягати одиниці, зовнішня ефективність взагалі є значно меншою, оскільки через великі значення показника заломлення більша частина випромінюваного світла відби-

ватиметься від поверхні всередину напівпровідника і там поглинається. Через поверхню проходить невелика частка світла, який падає на поверхню під кутами, меншими за $\arcsin(1/n_r)$, де n_r – коефіцієнт заломлення матеріалу на частоті випромінювання. Зазвичай цей кут невеликий і може

досягати $\approx 16^\circ$. Величина втрат за рахунок повного внутрішнього відбиття у світловипромінювальних структурах загалом визначається коефіцієнтом заломлення напівпровідника на частоті випромінювання. Ці втрати можуть бути досить суттєвими і для їхньої мінімізації використовують конструкційні особливості приладів (рис. 15.52). При цьому необхідно брати до уваги також особливості

природи напівпровідника, що є основою для світловипромінювальних діодів. Наприклад, червоне світло, що випромінюється світлодіодом на основі **Zn-O** в кристалі GaP, зазнає лише невеликого ослаблення, і в середньому кожен фотон може кілька разів відбитися від поверхонь структури. У результаті, якщо число відбиттів є великим, то ймовірність виходу фотону із діоду зростає. Це буде вірним лише у випадку, якщо після всіх перевідбиттів світло потрапить на випромінювальну поверхню під кутом, меншим за критичний, або у випадку, коли в точці падіння наявний поверхневий дефект. На відміну від непрямоzonної системи **Zn-O в GaP**, світло, що випромінюється прямоzonним напівпровідником зазвичай сильніше поглинається. Але, з іншого боку, у таких системах легше досягти великих значень внутрішньої ефективності, ніж у непрямоzonних структурах. Загалом, коефіцієнти поглинання в p - та n - областях переходів можуть сильно відрізнятись, і, окрім того, частина світла втрачається при поглинанні на контактах. Наприклад, у світлодіодах на основі **GaAs** випромінювання зазвичай виводиться через область n -типу, оскільки поглинання в p -області значно сильніше, і в результаті до випромінювальної поверхні доходить не більше половини генерованих в середині структури світла. У матері-

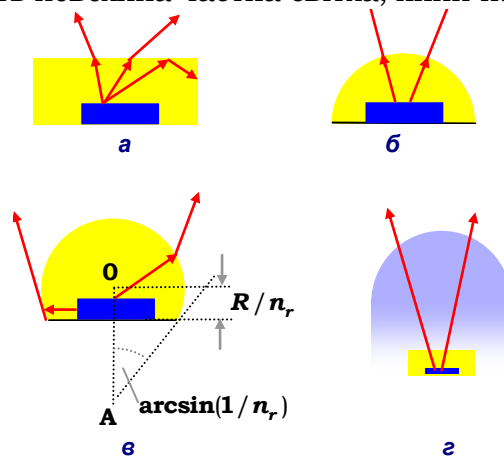


Рис. 15.52. Конструктивні типи світлодіодів: а – плоска структура; б – напівсфера; в – сфера Вейерштраса; г – структура зі скляним або полімерним покриттям

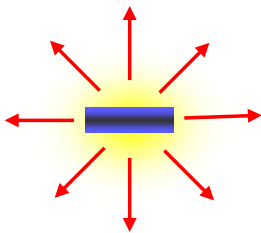


Рис. 15.53
Випромінення світла з кристалу кубічної симетрії

алі кубічної симетрії спонтанне випромінювання емітується ізотропно (рис. 15.53), тому для простої плоскої структури, що показана на рис. 15.52, частку світла, що проходить через верхню поверхню можна оцінити як

$$\eta_e \approx \frac{1}{2} \left(1 - \sqrt{1 - \frac{1}{n_r^2}} \right), \quad (15.238)$$

де фактор $\frac{1}{2}$ виникає через припущення: із прила-

ду може вийти світло, емітоване у напрямку верхньої поверхні. Для структур із прозорими p - та n -областями або із дзеркальною нижньою поверхнею цей фактор дорівнює одиниці. Для напівпровідників із $n_r \approx 3 - 3.5$ із такої **найпростішої** структури можна вивести не більше 1,5 % згенерованого світла. Однак, використання напівсфери або сфери Вейерштраса (див. рис. 15.52) дозволяє суттєво поліпшити вихід світла. В обох структурах, якщо випромінювальний перехід (позначено синім кольором) достатньо малий, усе світло, що емітує у верхню напівсферу, падає на зовнішню поверхню кристалу під кутом, меншим за критичний. У таких структурах ефективність може дорівнювати $\eta_e \approx 0.35$. Крім того, використання структури у вигляді сфери Вейерштраса дозволяє отримувати випромінюване світло, колімоване у конічний пучок із напівкутом $\arcsin(1/n_r)$. Однак, таке зменшення втрат на відбиття досягається за рахунок значного збільшення довжини оптичного шляху у зразку, що спричиняє збільшення втрат і може призвести до зникнення всіх переваг від поліпшеної геометрії структури. Інший, найбільш поширений метод, що дозволяє значно поліпшити випромінювальні властивості світлодіодів порівняно із плоскою структурою, вільний від недоліків (через самопоглинання), полягає у використанні покриття зі скла або полімеру із високими показниками заломлення. Це дозволяє збільшити кут повного внутрішнього відбиття, що автоматично збільшує ефективність структури.

15.10.7.3. Напівпровідникові лазери

Емісія світла у напівпровідникових структурах може **служити** основним механізмом для створення генераторів когерентного випромінювання. Суттєва відмінність напівпровідникових інжекційних лазерів від газових або кристалічних полягає в тому, що переходи, які зумовлюють лазерне випромінювання, відбуваються не між дискретними рівнями, а між парою широких енергетичних зон. Із цієї причини спонтанне випромінюванням займає відносно широкий спектральний

інтервал. В GaAs цей інтервал може досягати 30 нм. Ширина лінії когерентного випромінювання є значно меншою. У напівпровідникових лазерах вдається отримати великі значення оптичного підсилення, що пов'язано: по-перше, з ефективним способом накачування через інжекцію носіїв у вузьку зону p - n -переходу за прямого зміщення; по-друге, – із високою щільністю станів в енергетичних зонах. У прямозонних напівпровідниках весь струм використовується для заповнення енергетичних зон, зі станів в яких відбуваються процеси випромінювальної рекомбінації. У газовому лазері, накачування якого здійснюється за допомогою електричного розряду, більша частина збуджених атомів виявляється у збуджених станах, за яких не відбувається лазерних переходів. У результаті ефективність накачування виявляється низькою. Накачування в кристалічних лазерах також виявляється неефективним, оскільки відбувається за допомогою оптичного випромінювання, що охоплює широкий спектр енергетичних станів, лише незначна частина яких задовольняє умовам накачування лазерного середовища.

Розглянемо детальніше умови генерації лазерного випромінювання в p - n -переході. При виведенні (15.18) для коефіцієнту поглинання світла напівпровідником ми брали до уваги як процеси поглинання фотонів системою, так і процеси емісії фотонів. Таким чином, формальною умовою генерації світла, як випливає із (15.18), є виконання нерівності

$$F(E_j) > F(E_i), \quad (15.239)$$

що означає: за умов, коли число заповнення збуджених станів є більшим за число заповнення незбуджених станів, коефіцієнт поглинання α від'ємний, тобто (15.239) означає наявність інверсії заселеності рівнів у системі, що є основною умовою появи лазерної генерації. Як ми вже зазначали, аналізуючи процеси емісії фотонів із p - n -переходу за інжекції неосновних носіїв, щільність енергетичних станів у зоні провідності перевищує щільність станів дірок у валентній зоні, що означає виконання умови (15.239). Іншою умовою появи лазерної генерації є наявність достатньо великого коефіцієнту оптичного підсилення G , який визначається співвідношенням $G = -\alpha$. З іншого боку, цей коефіцієнт залежить від співвідношення швидкостей спонтанного та вимушеного випромінювання. Розглянемо фотони, що належать до визначеної моди і взаємодіють з електроном, який зв'язаний парою енергетичних станів, розділених енергією $\hbar\omega$. Відношення ймовірностей випромінювання світла до ймовірності поглинання дорівнює відношенню суми числа заповнення фотонів у моді n_ω і числа заповнення спонтанно емітованих фотонів ($=1$) до n_ω . При цьому вважа-

ється, що один зі станів, між якими можливий перехід, заповнений, а інший – вільний. В умовах, коли нижній рівень заповнений, а верхній вільний, коефіцієнт поглинання α_0 можна визначити так, щоб ймовірність поглинання фотону в одиницю часу дорівнювала $\alpha_0 v_g$, де v_g – групова швидкість хвильового пакету. Якщо до системи надійшло n_ω фотонів, то швидкість, з якою система їх поглине, дорівнюватиме $\alpha_0 v_g n_\omega$. Якщо, навпаки, верхній стан заповнений, а нижній вільний, то швидкість спонтанної емісії фотонів, що належать даній моді, дорівнює $\alpha_0 v_g$. Зазначимо, що таке визначення коефіцієнту поглинання має сенс, тільки якщо конкретизована спектральна ширина лінії поглинання $\Delta\omega_s$, у межах якої коефіцієнт α_0 можна вважати сталим. Таким чином, щільність електромагнітних мод, що можуть взаємодіяти із визначеною парою електронних рівнів, дорівнює $N_\omega \Delta\omega_s$, де N_ω є числом мод в одиничному частотному інтервалі⁴. Щільність фотонів, що беруть участь у процесі випромінювання – $P_\omega = n_\omega N_\omega \Delta\omega_s$. Звідси випливає, що число переходів в одиницю часу дорівнює $\alpha_0 v_g P_\omega$. Взагалі, коли ймовірності заповнення рівнів електронами набуває будь-яких значень між нулем та одиницею, можна ввести функції розподілу f_e – ймовірність заповнення верхнього рівня електронами, та f_h – ймовірність заповнення нижнього рівня діркою. Тоді, беручи до уваги принцип Паулі, швидкість спонтанного випромінювання фотонів, що належать до інтервалу мод $N_\omega \Delta\omega_s$, одиницею об'єму напівпровідника, можна записати у вигляді

$$r_\omega \Delta\omega_s = f_e f_h \alpha_0 v_g N_\omega \Delta\omega_s. \quad (15.240)$$

Аналогічно швидкість вимушеної емісії світла можна записати як $f_e f_h \alpha_0 v_g P_\omega$, а швидкість поглинання фотонів як $(1 - f_e)(1 - f_h) \alpha_0 v_g P_\omega$. Тоді сумарна швидкість поглинання за вимушеної емісії запишеться як

$$r_\omega^T = (1 - f_e - f_h) \alpha_0 v_g P_\omega. \quad (15.241)$$

⁴ В одиничному об'ємі матеріалу, що характеризується показником заломлення n_r , число електромагнітних мод, що приходяться на частотний інтервал $d\omega$, визначається як $N_\omega d\omega = 2 \cdot 4\pi q^2 dq$, де $q = n_r \omega / c$ – хвильовий вектор світла, що випромінюється, а фактор 2 враховує дві перпендикулярні поляризації. Оскільки групова швидкість пакету визначається як $v_g = d\omega / dq$, то для числа мод в одиничному частотному інтервалі знайдемо $N_\omega = 8\pi n_r^2 \omega^2 / c^2 v_g$.

Звідси отримуємо модифікований коефіцієнт поглинання (що зі знаком мінус дорівнює коефіцієнту підсилення)

$$\alpha = (1 - f_e - f_h)\alpha_0. \quad (15.242)$$

Виключимо звідси за допомогою (15.240) α_0 , і знайдемо коефіцієнт підсилення як функцію швидкості спонтанного випромінювання

$$G = \frac{1 - f_e - f_h}{f_e f_h} \frac{\pi}{2} \left(\frac{c}{n_r \omega} \right)^2 r_\omega. \quad (15.243)$$

У напівпровідниках верхній і нижній рівні, між якими відбуваються переходи, які формують спектр поглинання або генерують світло, що емітується зі зразку, можна ідентифікувати, відповідно до зони провідності та валентної зони. Крім того, можна вважати, що функції розподілу електронів і дірок за енергетичними станами є функціями Фермі, тобто

$$f_e = \frac{1}{1 + \exp\left[\frac{(\tilde{E}_C - \xi_e)}{kT}\right]}, \quad (15.244)$$

$$f_h = \frac{1}{1 + \exp\left[\frac{(\xi_h - \tilde{E}_V)}{kT}\right]}, \quad (15.245)$$

де ξ_e та ξ_h – квазірівні Фермі для електронів і дірок, а \tilde{E}_C та \tilde{E}_V деякі енергетичні стани у зоні провідності та валентній зоні, відповідно. Введення цих величин відповідає умові: у кожній зоні встановлюється рівноважний розподіл носіїв, хоча електрони та дірки не знаходяться в рівновазі один з одним. Використання функцій розподілу Фермі означає, що часи вільного пробігу носіїв вважаються малими порівняно із часом їхньої рекомбінації. За цієї умови відхилення від рівноваги швидко релаксують, і, якщо застосовується метод інжекції, що створює носії із великою енергією, то рівноважний розподіл носіїв встановлюється за дуже короткі часи так, що рекомбінують майже рівноважні носії. У такому разі (15.243) можна, використавши вирази (15.244-15.245), записати у вигляді

$$G = \frac{\pi r_\omega}{2} \left(\frac{c}{n_r \omega} \right)^2 \left(1 - \exp\left[\frac{\hbar\omega - \Delta\xi}{kT}\right] \right), \quad (15.246)$$

де через $\hbar\omega = \tilde{E}_C - \tilde{E}_V$ позначено енергію фотона, що випромінюється системою, а через $\Delta\xi = \xi_e - \xi_h$ – відстань між квазірівнями Фермі електронів і дірок.

Із (15.243) видно, що умовою від'ємного поглинання (генерації світла) є

$$f_e + f_h > 1, \quad (15.247)$$

звідки, використовуючи (15.244-15.245), отримуємо нерівність

$$\hbar\omega = \tilde{E}_C - \tilde{E}_V < \xi_e - \xi_h. \quad (15.248)$$

Звідси, зокрема впливає, що існує внесок у підсилення від пар енергетичних рівнів, які задовольняють умові (15.248), при цьому рівні не обов'язково розташовані між квазірівням Фермі. Хоча для отримання великих значень підсилення необхідно, щоб обидві величини f_e і f_h були великими, чого можна досягнути лише у випадку, коли обидва стани лежать у проміжку між ξ_e і ξ_h . Оскільки ефект оптичного підсилення реалізується за великих значень чисел заповнення, це означає: концентрація електрон-діркових пар має бути великою, отже оптичне підсилення обмежено у просторі областю, що знаходиться безпосередньо поблизу переходу. Таким чином, величини підсилення, що можуть бути досягнуті на практиці, отримуються лише для випромінювання, що розповсюджується у площині переходу. Критичною є вимога, щоб p - n -перехід був плоским. Зворотній зв'язок, необхідний для виникнення лазерної генерації, можна створити, обмеживши активну область двома оптично плоскими відбиваючими поверхнями, тобто утворивши інтерферометр Фабрі-Перо. Поверхні мають бути паралельними одна до одної і перпендикулярними до площини переходу. Схе-

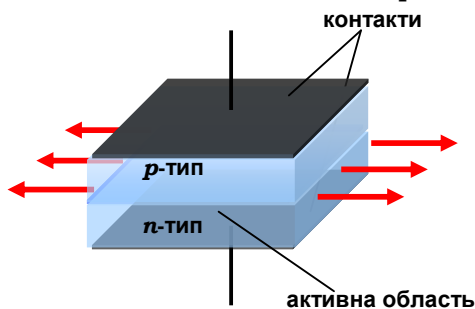


Рис. 15.54. Структурна схема напівпровідникового лазера типу Фабрі-Перо

матично такий лазер подано на рис. 15.54. Оскільки спектральний діапазон, в якому відбувається оптичне підсилення, є достатньо широким, а на довжині резонатора d вміщується велика кількість хвиль ($\sim 10^3 - 10^4$), то умова Фабрі-Перо для позитивного зворотного зв'язку

$$M\lambda = 2n_r d \quad (15.249)$$

виконується для великої кількості мод M . Лазерний ефект виникає тільки у

випадку, коли оптичне підсилення для хвилі, що проходить через резонатор, перевищує втрати при проходженні світла всередині резонатора.

В інжекційних лазерах існує порогова величина густини струму, починаючи з якої ця умова виконується. За струмів, нижчих за цей поріг, відбувається лише спонтанне випромінювання. За перевищення порогового значення струму у системі відбувається оптичне підсилення, і діод починає працювати як лазер. Для визначення порогового значення струму необхідно знати величину втрат та оптичне підсилення, якого можна досягти за заданої густини струму. Оскільки для

виникнення лазерної генерації необхідно створити високу щільність електрон-діркових пар, одним із домінуючих процесів втрат є поглинання вільними носіями. Оптичні втрати можуть також виникати завдяки домішкам та іншим дефектам ґратки, що дають додаткові внески до поглинання та розсіювання світла. Для хвилі, що експоненційно наростає із відстанню, як у разі оптичного підсилення, пропускання на площинах, що утворюють оптичний резонатор, являє собою оптичні втрати, еквівалентні поглинанню з коефіцієнтом поглинання

$$\alpha_R = -\frac{1}{2d} \ln(R_1 R_2), \quad (15.250)$$

де R_1 і R_2 – коефіцієнти відбиття від двох граней резонатора. Іншим каналом втрат є так звані дифракційні втрати. Вони виникають завдяки тому, що електромагнітна хвиля, яка розповсюджується у вузькій (шириною кілька довжин хвиль) активній області дифрагує в суміжні поглинаючі області, що оточують активний шар.

Втрати, пов'язані із поглинанням вільними носіями, і дифракційні втрати зазвичай є порівняними за величиною та сумарно із α_R дають коефіцієнт поглинання

$$\alpha_t = \alpha_{FC} + \alpha_d + \alpha_R, \quad (15.251)$$

де α_{FC} та α_d – коефіцієнти поглинання вільними носіями та дифракційних втрат, відповідно. Зазвичай втрати щодо коефіцієнта підсилення повільно змінюються залежно від енергії фотону. Це означає, що генерація відбуватиметься при енергій, близьких до максимуму кривої підсилення.

Для характеристики величини підсилення в лазерних структурах зручно користуватись безрозмірним параметром

$$Z(T) = \frac{r_\omega \Delta\omega}{R_S} \left(1 - \exp \left[\frac{\hbar\omega_{\max} - \Delta\xi}{kT} \right] \right), \quad (15.252)$$

що являє собою відношення швидкості спонтанної емісії фотонів з енергією, що відповідає максимуму оптичного підсилення – $\hbar\omega_{\max}$, до повної швидкості спонтанної рекомбінації R_S . Іншими словами, цей параметр показує міру відносних енергій фотону, за яких криві спонтанного випромінювання та підсилення досягають максимуму, тобто визначає границю величини інверсної заселеності, якої можна досягти у системі.

В інжекційних світловипромінюючих структурах максимум підсилення завжди досягається за енергії фотону, меншій за ту, що відповідає максимуму спектру спонтанного випромінювання (різниця буде більшою, якщо поперечний імпульс електрона не зберігається, тобто, у

випадку, коли імпульс електрона до випромінювального переходу і після нього відрізняються⁵). Якби енергії обох піків були однакові, виконувалась би рівність $R_S = r_\omega \Delta\omega$. А оскільки ми припускаємо, що в системі існує позитивне підсилення, тобто $\Delta\xi > \hbar\omega$, то маємо, що вираз у дужках у (15.252) зменшується за збільшення температури. Таким чином, за зниженні температури, в умовах, коли енергії максимумів підсилення та спонтанного випромінювання майже збігаються, параметр $Z(T)$ наближається до максимального значення $Z(T) = 1$.

Використовуючи (15.252), із формули (15.246) знайдемо вираз для коефіцієнта підсилення в околі максимального значення

$$G_{\max} = \frac{\pi\lambda^2 R_S}{2n_r^2 \Delta\omega} Z(T). \quad (15.253)$$

З іншого боку, швидкість спонтанного випромінювання можна записати через густину струму як

$$R_S = \frac{\eta I}{ed}, \quad (15.254)$$

де η – квантова ефективність спонтанної генерації фотонів у кристалі, d – товщина активного шару. При цьому вважатимемо, що весь струм дає внесок до інжекції носіїв до активної області. Використовуючи (15.253–15.254), можна знайти вираз для оцінки критичного струму, за якого система досягає порогу генерації. Дійсно, перехід через поріг генерації означає, що підсилення G_{\max} починає перевищувати сумарний коефіцієнт поглинання α_T . Таким чином, для порогової густини струму можна записати

$$I_{th} = \frac{2n_r^2 \Delta\omega ed \alpha_T}{\pi\lambda^2 Z(T)\eta}. \quad (15.255)$$

Типові значення порогової густини струму для інжекційних лазерів на основі p - n -переходу в арсеніді галію за температури рідкого азоту дорівнюють $10^{-1} \text{ А}\cdot\text{см}^{-2}$. За зниженні температури величина I_{th} змінюється дуже повільно, але за збільшення температури густина порогового струму швидко ($\sim T^3$) зменшується. Таким чином, лазерний режим у простому лазері із p - n -переходом за кімнатних температур можна реалізувати лише використовуючи короткі імпульси струму великої густини для того, щоб породжуваний ним нагрів переходу був невеликий.

⁵ Початковий та кінцевий стани електронів відповідають знаходяться різним значенням \mathbf{k} , але сумарний імпульс зберігається за рахунок пружних розсіювань зарядженими домішками.

Швидке зростання порогового струму за підвищення температури від азотної робить використання лазерів на простих $p-n$ -переходах непридатними для практичного застосування. Досягнути значного зменшення порогового струму та підвищення ефективності лазерів за кімнатної температури можна використовуючи гетероструктури, наприклад, на основі $\text{GaAs} - \text{Ga}_x\text{Al}_{1-x}\text{As}$. Принцип дії таких гетероструктур, як генератори лазерного випромінювання з обмеженням, полягає в наявності сходинки у зоні провідності (рис. 15.55), що дає можливість утримувати область інверсної заселеності у вузькій обмеженій області. Для більшого розуміння переваг лазерів на гетероструктурах розглянемо рисунок детальніше та проаналізуємо особливості роботи такої структури як лазерної.

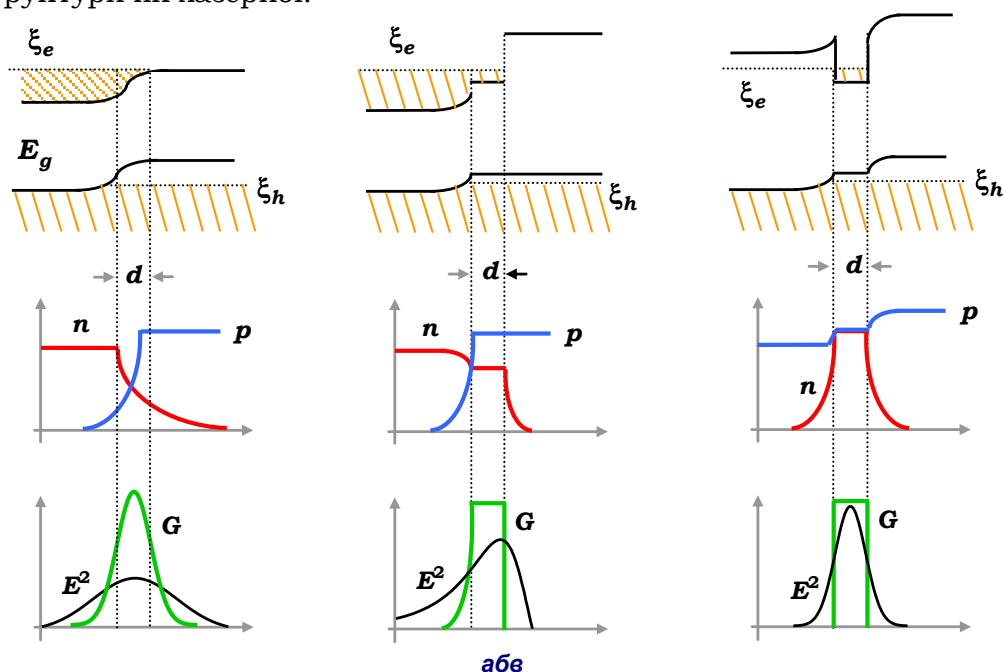


Рис. 15.55. Енергетичні діаграми напівпровідникових лазерних структур (зверху); схематична поведінка щільності електронів (червоні криві) і дірок (сині криві); амплітуди світлового поля E^2 (чорні криві); коефіцієнта оптичного підсилення G (зелені криві) для:
а – $p-n$ -перехід у напівпровіднику; **б** – лазер з одним гетеропереходом;
в – лазер із подвійною гетероструктурою

На рис. 15.55 *а* подано зонну діаграму напівпровідникового лазера із $p-n$ -переходом, розподіл концентрації носіїв, коефіцієнт підсилення і амплітуду світлового поля в околі активного шару (товщиною d) за прямого зміщення. Видно, що розподіл концентрації електронів і дірок і

розподіл коефіцієнту оптичного підсилення а також, особливо дуже широкий розподіл інтенсивності світлового поля у простій p - n -структурі призводять до того, що світлове поле генерації проникає далеко за межі активного шару в області з високим коефіцієнтом поглинання, що спричиняє суттєві втрати. Крім того, товщина активного шару у таких структурах є меншою за область рекомбінації нерівноважних носіїв заряду, що інжектуються в область p - n -переходу. Це означає, що область оптичного підсилення "розмазана" у широкому okolí активного шару. Усі ці фактори визначають великі енергетичні втрати, високі значення порогового струму та низький ККД за кімнатних температур. Навпаки, у гетеролазерах (на рис. 15.55 б подано зонну діаграму напівпровідникової гетероструктури, розподіл концентрації носіїв, світлового поля і коефіцієнту оптичного підсилення за прямого зміщення. У цьому гетеролазері з однією структурою за рахунок гетеропереходу на відстані d від інжектуючого p - n -переходу утворюється потенціальний бар'єр за рахунок гетеропереходу із більш широкозонним напівпровідником. Якщо швидкість рекомбінації на гетерограниці недостатньо велика (це спостерігається на всіх гетеропереходах, коли збігаються параметри кристалічної ґратки напівпровідників, що утворюють гетероперехід), то нерівноважні носії, що не встигли рекомбінувати, відбиваються від бар'єру та повертаються до активної області, тим самим підвищуючи концентрацію електрон-діркових пар в області оптичного підсилення. Таким чином, інверсна заселеність рівнів в активному шарі виникає за невеликих значень густини струму, тим самим ефективно знижуючи його порогове значення. Стрибок показника заломлення на гетерограниці одночасно спричиняє зменшення проникнення світлового поля до сильно поглинаючої p -області (нижній графік на рис. 15.55 б). Усі ці фактори сприяють суттєвому покращенню параметрів гетеролазерів і підвищують можливість виникнення лазерної генерації за низьких значень струму і кімнатних температур. Ще кращі результати отримують при використанні подвійні гетероструктури з активним шаром із вузькозонного напівпровідника, що розташований між двох шарів широкозонних напівпровідників (рис. 15.55 в). Така структура завдяки наявності двостороннього бар'єру характеризується ефектом двостороннього обмеження та збігом області інверсної заселеності та великих значень

світлового поля. На рис. 15.56 подано структуру гетеролазера GaAs-AlGaAs із подвійною гетероструктурою, де резонатор Фабрі-Перо утворений сколотими гранями напівпровідникового кристалу (площини відколу {110}), а активний шар утворений кристалічними площинами з орієнтацією {100}.

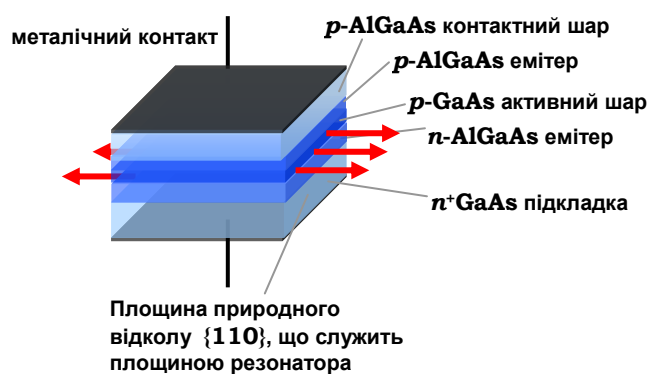


Рис. 15.56. Схема лазера на подвійній гетероструктурі

■ 15.11. Фотоелектричні ефекти у напівпровідниках

Під дією зовнішнього випромінювання електронам у напівпровідникових кристалах передається енергія фотонів, у результаті можуть генеруватися додаткові нерівноважні носії, що спричиняє підвищення їхньої концентрації, або може спостерігатись емісія електронів із напівпровідника. У такому разі кажуть, що в напівпровіднику відбувається внутрішній або зовнішній фотоелектричний ефект.

15.11.1. Зовнішній фотоелектричний ефект

При поглинанні напівпровідником фотона достатньої енергії може відбутись емісія електрона із кристалу. При цьому електрон проходить через межу розподілу напівпровідник-вакуум так, щоб виявитись на вакуумному рівні E_{vac} . Таким чином, для виникнення зовнішнього фотоелектричного ефекту необхідно, щоб енергія фотона перевищувала порогову величину, яка для власного або невиродженого напівпровідника визначається сумою енергії електронної спорідненості та ширини забороненої зони (рис. 15.57), оскільки найвищий енергетичний рівень, що зайнятий електронами, є степою валентної зони

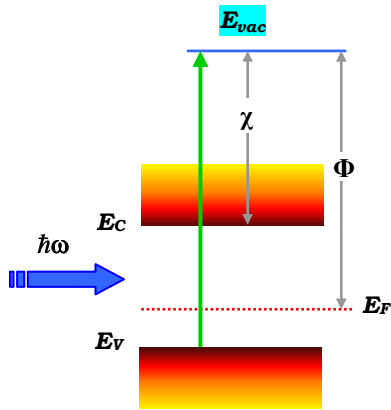


Рис. 15.57. Зовнішній фотоелектричний ефект у власному напівпровіднику

$$E_T = \chi + E_g \quad (15.256)$$

У випадку сильно легovanого напівпровідника *p*-типу, коли рівень Фермі розташований на величину ξ_p нижче стелі валентної зони (рис. 15.58), мінімальне значення енергії фотона, що призводить до емісії електрона на вакуумний рівень визначається порогом

$$E_T = \chi + E_g + \xi_p, \quad (15.257)$$

а для сильно легovanого напівпровідника *n*-типу, коли рівень Фермі лежить на величину ξ_n вище дна зони провідності, порогова енергія фотона визначається формулою

$$E_T = \chi - \xi_n, \quad (15.258)$$

що зрозуміло із рис. 15.59. Як видно із рис. 15.58–15.59, поріг зовнішнього фотоелектричного ефекту у сильно легovanих напівпровідниках *n*-типу істотно (на величину $\sim E_g$) менший, ніж у напівпровідниках *p*-типу.

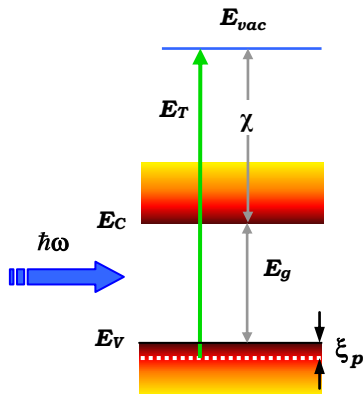


Рис. 15.58. Зовнішній фотоелектричний ефект у сильно легovanому напівпровіднику *p*-типу

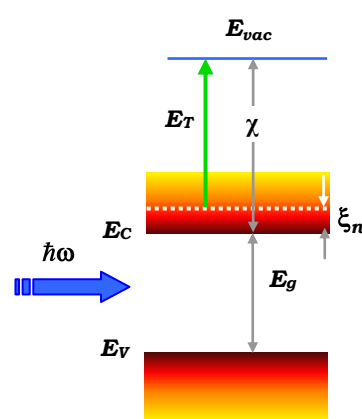


Рис. 15.59. Зовнішній фотоелектричний ефект у сильно легovanому напівпровіднику *n*-типу

Зрозуміло, що поріг фотоелектричного ефекту, що визначається співвідношеннями (15.256–15.258), являє собою мінімальні можливі значення E_T . У реальних ситуаціях на формування порогової енергії впливають як процеси поглинання і розсіяння електронів на дефектах, так і стан поверхні. Крім того, необхідно мати на увазі, що ці процеси проходять за законами збереження енергії та імпульсу. Наприклад, для зовніш-

нього фотоефекту у власних непрямозонних напівпровідниках збереження імпульсу забезпечується емісією фонона, і порогове значення енергії визначатиметься формулою

$$E_T = \chi + E_g + E_{ph}, \quad (15.259)$$

де E_{ph} – енергія емітованого фонона.

За мірою збільшення енергії фотонів, при переході через порогове значення, число емітованих електронів різко зростає. Далі зростання уповільнюється та за енергій, що перевищують поріг приблизно на 1eV, переходить у плато. Структура плато відображає властивості зонної структури напівпровідника. Початкова ділянка різкого зростання квантового виходу (кількість емітованих електронів, що приходить на один фотон збудження) описується такою залежністю

$$Y(\hbar\omega) = A \cdot (\hbar\omega - E_T)^a, \quad (15.260)$$

де A та a – константи. Константа a визначається механізмом збудження електрона зовнішнім випромінюванням. Наприклад, за емісії електрона шляхом прямого переходу із валентної зони із врахуванням розсіювання на дефектах $a = 2$. Якщо електрон емітує із дискретного поверхневого стану нижче рівня Фермі, коли $E_T = \chi + E_C - E_F$, величина $a = 1$. Взагалі різним областям енергетичного спектру фотонів відповідають різні механізми збудження електронів, що викликають їхню емісію із кристалу і, як результат, – різні залежності квантового виходу від $(\hbar\omega - E_T)$. Наприклад, для енергій фотонів поблизу порогу характерна кубічна залежність виходу від $(\hbar\omega - E_T)$, а для збудження високоенергетичними фотонами – лінійна.

15.11.2. Внутрішній фотоефект. Фотопровідність

За внутрішнього фотоефекту поглинання світла спричиняє збудження електрона та переведення його до зони провідності. За рахунок цього змінюватиметься електричний опір напівпровідника. Додаткова провідність, обумовлена носіями заряду, що створені оптичною генерацією, називається фотопровідністю. За міжзонних переходів спостерігається явище власної фотопровідності. Для прямозонних напівпровідників порогове значення енергії фотона, за якої виникає фотопровідність, дорівнює ширині забороненої зони. У випадку непрямих переходів, коли закон збереження імпульсу забезпечується за рахунок випромінювання фонона з енергією E_{ph} , довгохвильовий край спектру фотопровідності

$$\hbar\omega = E_g + E_{ph} . \quad (15.261)$$

У сильно легovanого напівпровідника n -типу рівень Фермі лежить вище дна зони провідності на величину ξ_n , тому довгохвильова границя спектру фотоструму за внутрішнього фотоефекту визначатиметься співвідношенням (рис. 15.60 а)

$$\hbar\omega = E_g + \xi_n . \quad (15.262)$$

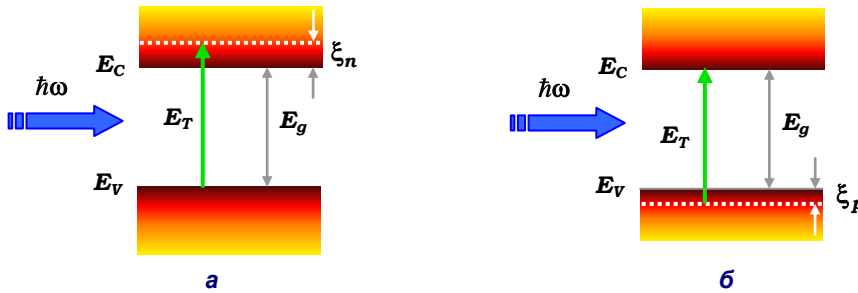


Рис. 15.60. Внутрішній фотоефект у сильно легovanих напівпровідниках: а – n -типу; б – p -типу

У сильно легovanого напівпровідника p -типу рівень Фермі лежить нижче стелі валентної зони на величину ξ_p . Це є причиною того, що порогове значення енергії фотону, за якого можна спостерігати внутрішній фотоефект, виражається формулою (рис. 15.60 б)

$$\hbar\omega = E_g + \xi_p . \quad (15.263)$$

Смуга власного поглинання у напівпровідниках завжди характеризується чітко вираженою довгохвильовою границею. У принципі напівпровідник може мати й короткохвильову границю, якщо його зона провідності є вузькою. Однак у багатьох напівпровідниках зона провідності перекривається із розташованими вище за енергією дозволенними зонами. Таким чином, часто у напівпровідниках зона провідності являє собою дуже широку суцільну енергетичну смугу дозволених станів. Цей факт свідчить, що спектральний розподіл фотоструму може розповсюджуватись далеко в короткохвильову область спектру. Але зі збільшенням енергії фотонів збільшується й коефіцієнт поглинання, тобто майже все світло поглинається вузьким приповерхневим шаром, тобто область генерації фотоносіїв розташована поблизу поверхні, де час життя носіїв заряду є меншим, ніж в об'ємі. Така зміна часу життя носіїв заряду спричинятиме зменшення фотопровідності в області коротких довжин хвиль. Якщо у забороненій зоні напівпровідника існують локальні рівні, оптичне поглинання може викликати

переходи між рівнями домішок та енергетичними зонами (рис. 15.61), що викличе виникнення домішкової фотопровідності. А оскільки енергія іонізації домішки менша за ширину забороненої зони, то спектр домішкового фотоструму буде зсунутий у довгохвильовий бік щодо спектру власної фотопровідності.

Поглинання світла вільними носіями та коливаннями ґратки не можуть безпосередньо спричинити зміни концентрації носіїв, отже й виникнення фотоефекту. Однак збільшення концентрації носіїв заряду у цьому випадку може відбуватись опосередковано, коли поглинання світла суттєво підвищує кінетичну енергію вільних носіїв заряду, які, віддаючи енергію на генерацію додаткових носіїв заряду, збільшують провідність напівпровідника.

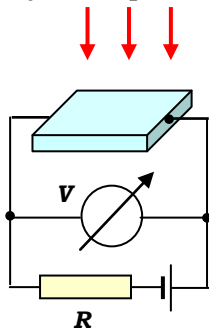


Рис. 15.62. Схема експерименту з вивчення ефектів фотопровідності

а фотопровідність напівпровідника, обумовлена безпосередньою дією світла. Зазначимо, що у разі, коли енергія фотону є меншою за ширину забороненої зони, фотон генерує носій одного знаку (напр., у результаті домішкового поглинання), а за енергії фотону, більшої за ширину забороненої зони, генеруються пари носіїв – електрон + дірка. Чутливість вимірювань можна підвищити, освітлюючи зразок імпульсним джерелом і використовуючи фазочутливий детектор. Таким чином, відносна зміна електропровідності визначається формулою

$$\frac{\Delta\sigma}{\sigma_0} = \frac{\Delta n\mu_n + \Delta p\mu_p}{n_0\mu_n + p_0\mu_p} = \frac{b\Delta n + \Delta p}{bn_0 + p_0}, \quad (15.267)$$

де $b = \mu_n / \mu_p$ – відношення рухливостей електронів і дірок. Концентрації надлишкових носіїв заряду Δp і Δn залежать від довжини хвилі та інтенсивності світла. Розглянемо шар зразка Δx , на який падає

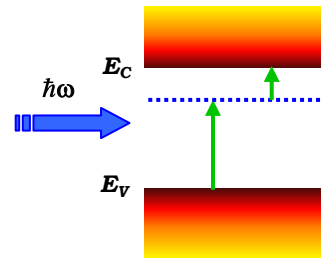


Рис. 15.61. Схема виникнення домішкової фотопровідності

світло інтенсивністю I . Якщо напівпровідник характеризується коефіцієнтом поглинання α , то кількістю енергії світла, що поглинається за одиницю часу одиницею площі матеріалу на перерізі Δx , є

$$\frac{dI}{dx} = \alpha I. \quad (15.268)$$

Таким чином, при поглинанні світла в області власного поглинання із довжиною хвилі λ одиницею об'єму напівпровідника за одиницю часу утворюється надлишкові електрони та дірки. Швидкість генерації надлишкових носіїв при цьому пов'язана з інтенсивністю та довжиною хвилі світла формулою

$$G = \beta \frac{\alpha I}{2\pi\hbar c / \lambda} = \beta \frac{\alpha I}{\hbar\omega}, \quad (15.269)$$

де $\hbar\omega$ – енергія квантів світла, а β – квантовий вихід фотоіонізації, що визначає число електрон-діркових пар⁶, які утворюються при поглинанні одним фотоном, якщо інтенсивність світла вимірюється числом фотонів, що проходять за одну секунду через поверхню одиничної площі. Зрозуміло, що завдяки процесам рекомбінації, які підсилюються зі зростанням концентрації нерівноважних носіїв заряду, фотопровідність збільшуватиметься після початку освітлення до максимального значення зі сталою часу τ_n для електронів і τ_p – для дірок. Дійсно, оскільки швидкість генерації нерівноважних носіїв залишається сталою за незмінної інтенсивності світла, то через проміжок часу τ_n або τ_p швидкість рекомбінації зрівняється зі швидкістю рекомбінації, і в напівпровіднику встановиться стаціонарний стан із постійними концентраціями надлишкових носіїв заряду Δn_{st} і Δp_{st} , що визначаються з умов

$$\Delta n_{st} = G\tau_n, \quad \Delta p_{st} = G\tau_p. \quad (15.270)$$

Підставимо ці вирази до (15.267). Із врахуванням (15.269) отримаємо, що відносна стаціонарна фотопровідність

$$\frac{\Delta\sigma_{ph.st}}{\sigma_0} = e\beta\alpha \frac{I}{\hbar\omega} \cdot \frac{b\tau_n + \tau_p}{bn_0 + p_0}. \quad (15.271)$$

Часто використовують таку характеристику напівпровідника, як питома фоточутливість, що визначається відношенням фотопровідності до інтенсивності опромінюючого світла. Якщо рухливість або час життя носіїв одного типу істотно перевершують ці параметри у носіїв іншого типу, то йдеться про монополярну фотопровідність. Наприклад,

⁶ У випадку домішкової фотопровідності коефіцієнт β визначає число надлишкових носіїв заряду.

якщо рухливість і час життя електронів є набагато більшими за рухливість і час життя дірок, у напівпровіднику реалізується електронна фотопровідність

$$\Delta\sigma_{ph}^n = e\beta\alpha\mu_n\tau_n \frac{I}{\hbar\omega}. \quad (15.272)$$

У протилежному випадку – у напівпровіднику спостерігається фотопровідність діркового типу

$$\Delta\sigma_{ph}^p = e\beta\alpha\mu_p\tau_p \frac{I}{\hbar\omega}. \quad (15.273)$$

За відомої фотопровідності неважко визначити стаціонарний фотострум

$$J_{phot} = e\beta\alpha(\mu_n\tau_n + \mu_p\tau_p) \frac{IE}{\hbar\omega}. \quad (15.274)$$

Оскільки добуток рухливості на величину електричного поля дорівнює дрейфовій швидкості носія, то фотострум у зразку, довжина якого в напрямку поля дорівнює l , визначатиметься відношеннями часів релаксації до часів дрейфу носіїв

$$J_{phot} = e\beta\alpha \left(\frac{\tau_n}{t_n} + \frac{\tau_p}{t_p} \right) \frac{I}{\hbar\omega} l, \quad (15.275)$$

де часи дрейфу електронів і дірок, відповідно

$$t_n = \frac{l}{v_{dr}^n} = \frac{l}{\mu_n E}, \quad t_p = \frac{l}{v_{dr}^p} = \frac{l}{\mu_p E}. \quad (15.276)$$

Таким чином, вимірюючи фотострум у зразку заданої геометрії за заданої довжини хвилі опромінюючого світла, можна визначити квантовий вихід фотоіонізації як функцію енергії фотону. Наприклад, для германію за температури 300 К квантовий вихід для енергій фотонів аж до 2.7 еВ дорівнює одиниці, при подальшого зростання енергії фотонів досить різко збільшується, досягаючи значення 1.8 при $\hbar\omega = 5\text{eV}$. Аналогічно для кремнію при $T = 300\text{K}$, $\beta = 1$ за енергій фотонів, менших від 3 еВ. Збільшення енергії фотонів спричиняє різке збільшення квантового виходу так, що при $\hbar\omega = 5\text{eV}$ квантовий вихід дорівнює 3.3. Така поведінка квантового виходу фотоіонізації пояснюється тим, що поглинання високоенергетичного фотону напівпровідником спричиняє появу високоенергетичних носіїв заряду, енергія яких достатня для утворення вторинних електрон-діркових пар через механізм ударної іонізації.

15.11.3. Релаксація фотопровідності

Уявимо, що зразок напівпровідника освітлюється імпульсом світла прямокутної форми (рис. 15.63). Зрозуміло, що через ненульові часи релаксації стаціонарне значення фотопровідності досягається не миттєво, а за деякий час після початку дії світла на зразок. Після вимкнення освітлення нерівноважні носії, згенеровані світлом, ще деякий час даватимуть внесок до провідності, оскільки їхня концентрація після припинення процесів генерації зменшуватиметься та за час релаксації зменшиться в e разів. Швидкість зміни концентрації надлишкових нерівноважних носіїв заряду у напівпровіднику визначається різницею між швидкістю їхньої генерації і рекомбінації. Нехай процеси рекомбінації визначаються швидкістю R . Тоді, беручи до уваги, що швидкість генерації носіїв за фотоіонізації визначається величиною G (див. (15.269)), можна записати

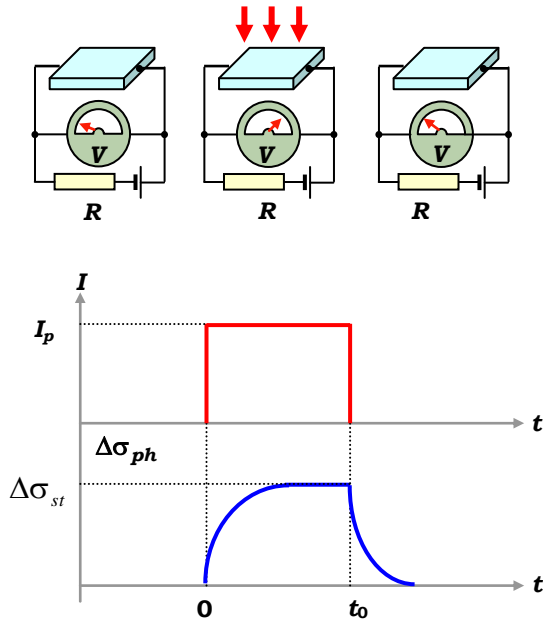


Рис. 15.63. Релаксація фотопровідності за освітлювання зразка прямокутним імпульсом світла

$$\frac{d\Delta n}{dt} = \beta \frac{\alpha I}{\hbar\omega} - R. \tag{15.277}$$

Це рівняння має досить універсальний характер за припущенням, що поглинання світла та рекомбінація вздовж товщини зразка є однорідними. Його легко розв'язати у граничних випадках слабого та сильного рівнів накачування.

Якщо інтенсивність випромінювання є невеликою, тобто основний внесок до електронних процесів у напівпровіднику здійснюють рівноважні носії, тобто $\Delta n \ll (n_0 + p_0)$, то можна вважати, що швидкість рекомбінації пропорційна концентрації надлишкових носіїв

$$R = \frac{\Delta n}{\tau}. \tag{15.278}$$

Тоді (15.277) запишеться у вигляді простого диференціального рівняння зі змінними, що розділяються

$$\frac{d\Delta n}{dt} = \beta \frac{\alpha I}{\hbar\omega} - \frac{\Delta n}{\tau}. \quad (15.279)$$

Беручи до уваги, що у початковий момент, коли було ввімкнене освітлення, концентрація надлишкових носіїв дорівнювала нулю ($\Delta n = 0$ при $t = 0$), отримаємо закон зміни концентрації надлишкових носіїв після ввімкнення освітлення

$$\Delta n = \Delta n_{st} (1 - e^{-t/\tau}), \quad (15.280)$$

де стаціонарне значення надлишкової концентрації носіїв визначається (15.279). Таким чином, наростання фотопровідності після ввімкнення освітлення невеликої інтенсивності має експоненційний характер і визначається часом релаксації, що є часом життя нерівноважних носіїв. Після вимкнення освітлення релаксація фотопровідності визначається також із (15.279), але з іншими початковими умовами: $\Delta n = \Delta n_{st}$ при $t = t_0$. Розділюючи змінні та виконуючи інтегрування отриманого рівняння, дістаємо

$$\Delta n = \Delta n_{st} \cdot e^{-t/\tau}, \quad (15.281)$$

тобто релаксація фотопровідності за малих рівнів освітлення визначається експоненційним законом зі сталою часу, що є часом життя нерівноважних носіїв заряду.

За великих рівнів освітлювання основний внесок до електронних процесів у напівпровіднику дають нерівноважні носії. Це означає, що $\Delta n \gg (n_0 + p_0)$ і можна вважати, що швидкість рекомбінації пропорційна квадрату концентрації надлишкових носіїв, тобто

$$R = \gamma(\Delta n)^2. \quad (15.282)$$

Підставляючи цей вираз до (15.277), отримаємо рівняння, що описує динаміку зміни концентрації надлишкових носіїв заряду за великих рівнів збудження

$$\frac{d\Delta n}{dt} = \beta \frac{\alpha I}{\hbar\omega} - \gamma(\Delta n)^2. \quad (15.283)$$

Це рівняння також є рівнянням зі змінними, що розділяються

$$\frac{d\Delta n}{\beta \frac{\alpha I}{\hbar\omega} - \gamma(\Delta n)^2} = dt \quad (15.284)$$

і може бути проінтегровано безпосередньо. Використовуючи початкові умови – при $t = 0$ $\Delta n = 0$, отримуємо що при ввімкненні опромінювання концентрація нерівноважних носіїв наростатиме згідно із законом

$$\Delta n = \Delta n_{st}^{(I)} \cdot \text{th}(\Delta n_{st}^{(I)} \cdot t), \quad (15.285)$$

де стаціонарне значення надлишкової концентрації носіїв, що досягається при $t \rightarrow \infty$, має вигляд

$$\Delta n_{st}^{(I)} = \left(\beta \frac{\alpha I}{\gamma \hbar \omega} \right)^{1/2}. \quad (15.286)$$

Якщо після досить тривалої дії на зразок інтенсивного опромінювання освітлення вимкнути, концентрація надлишкових носіїв релаксуватиме згідно із (10.31) від величини $\Delta n_{st}^{(I)}$, оскільки за квадратичної рекомбінації велика надлишкова концентрація носіїв зменшується за гіперболічним законом. Таким чином, вимкнення освітлення спричинить спад концентрації надлишкових носіїв за законом

$$\Delta n = \frac{\Delta n_{st}^{(I)}}{1 + \Delta n_{st}^{(I)} t}. \quad (15.287)$$

Ми отримали досить передбачений результат: фотопровідність залежить від інтенсивності збуджуючого світла. За слабких рівнів освітлювання фотопровідність визначають процеси генерації носіїв, і її залежність від інтенсивності світла є лінійною

$$\sigma_{ph} \sim \frac{I}{\hbar \omega}. \quad (15.288)$$

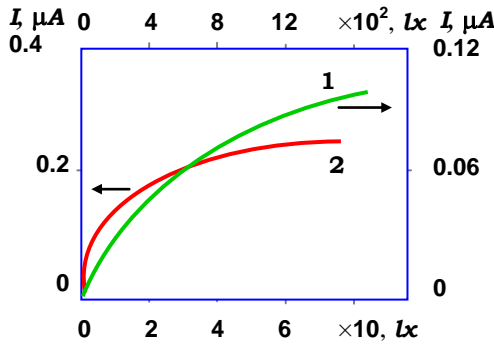


Рис. 15.64. Залежність стаціонарного фотоструму від інтенсивності освітлювання у ВіС за слабого (1) та сильного (2) освітлювання

За збільшення інтенсивності світла все більшу роль починають відігравати рекомбінаційні процеси, і зростання фотопровідності уповільнюється (рис. 15.64). За великих інтенсивностей освітлювання поведінка фотопровідності визначається коренем квадратним від інтенсивності освітлювання

$$\sigma_{ph} \sim \sqrt{\frac{I}{\hbar \omega}}. \quad (15.289)$$

Крім випромінювальної рекомбінації із випромінюванням світла з енергією $\hbar \omega = E_g$ у напівпровіднику

може відбуватись рекомбінація через рекомбінаційні центри без випромінюванням фотонів і випромінюванням світла із частотою, меншою за ширину забороненої зони. Такими центрами рекомбінації можуть виступати домішки та недосконалість ґратки. Наявність таких центрів впливатиме на кінетику фотопровідності, оскільки електрони не тільки рекомбінуватимуть із дірками, а також захоплюватимуться

на центри захоплення. Центри, будучи зайнятими електронами, не можуть захопити дірку, тому рекомбінація на таких центрах не відбувається. Такі пастки мають спеціальну назву центрів прилипання, і для опису впливу наявності таких центрів на кінетику фотопровідності вводять додаткові часи захоплення носіїв заряду на центри прилипання. Таким чином, наявність центрів прилипання спричинить зменшення швидкості наростання концентрації надлишкових носіїв заряду за освітлення зразку. У разі вимкнення освітлення наявність центрів прилипання спричинить спустошення таких пасток і спад фотопровідності затягнеться у часі.

■ 15.12. Задачі

1. Визначити вплив центрів прилипання на кінетику фотопровідності напівпровідника.

Розв'язок. Рівнями рекомбінації зазвичай є глибокі рівні. Неглибокі рівні при цьому відіграють роль центрів прилипання. Нехай τ_1 – час захоплювання дірки на акцепторі, а τ_2 – час життя дірки на акцепторі, тобто $1/\tau_1$ є ймовірністю, з якою дірка може бути захоплена на акцепторі протягом часу τ_2 утримуватись на такому центрі. Якщо концентрація нерівноважних надлишкових носіїв Δp та Δn , то концентрація дірок, що захоплені на центрах прилипання, визначається як

$$\Delta p_t = \Delta n - \Delta p. \quad (15.290)$$

Нехай $1/\tau_p$ – ймовірність рекомбінації електрон-діркової пари. Тоді зміна концентрації надлишкових носіїв відбуватиметься згідно із рівнянням

$$\frac{d\Delta p}{dt} = G - \frac{\Delta p}{\tau_p} - \frac{\Delta p}{\tau_1} + \frac{\Delta p_t}{\tau_2}. \quad (15.291)$$

В умовах рівноваги швидкість захоплення на центри прилипання має дорівнювати швидкості теплової генерації дірок із акцепторів d зони, тобто

$$\frac{\Delta p_0}{\tau_1} = \frac{\Delta p_t^{(0)}}{\tau_2}. \quad (15.292)$$

У стаціонарному випадку швидкість генерації діро можна виразити через стаціонарну надлишкову концентрацію

$$G = \frac{\Delta p_0}{\tau_p}, \quad (15.293)$$

а (15.291) із урахуванням (15.290) записати як

$$0 = G - \frac{\Delta p_0}{\tau_p} - \frac{\Delta p}{\tau_1} + \frac{\Delta n_0 - \Delta p_0}{\tau_2}. \quad (15.294)$$

Звідси, використовуючи (15.293), отримуємо надлишкову концентрацію електронів у стаціонарному випадку

$$\Delta n_0 = G\tau_p \left(1 + \frac{\tau_2}{\tau_1} \right). \quad (15.295)$$

Вважаючи напівпровідник слабо легованим, тобто $n_0 = p_0$, з (15.267) отримуємо

$$\frac{\Delta\sigma}{\sigma_0} = \frac{G}{n_0} \cdot \tau_p \frac{1 + b(1 + \tau_2 / \tau_1)}{1 + b}. \quad (15.296)$$

Ця формула описує відносну зміну фотопровідності напівпровідника з ефективним часом релаксації

$$\tau_p^{eff} = \tau_p \frac{1 + b(1 + \tau_2 / \tau_1)}{1 + b}. \quad (15.297)$$

Звідси видно, що у випадку, коли час утримання носія на центрі прилипання істотно перевищує час захоплення дірки на акцептор, ефективний час релаксації значно перевищує час релаксації за відсутності центрів прилипання τ_p . Це означає, що наявність центрів прилипання істотно збільшує відносну фотопровідність, тобто центри прилипання збільшують фоточутливість напівпровідника. Цей результат дуже просто пояснити. Дійсно, чим більше дірок захоплюватимуться на центрах прилипання, тим більше електронів залишається вільними у зоні провідності, що забезпечить значне зростання провідності. Кількість дірок на центрах прилипання є тим більшою, чим більшим є час перебування дірки на такому центрі τ_2 .

2. Знайти порогову енергію поглинання фотону для прямих переходів у сильно легованому арсеніді галію n -типу. Вважати, що хвости зон за такого легування не утворюються та $N_e = 2 \cdot 10^{18} \text{ cm}^{-3}$.

3. Визначити світловий потік, що проходить через кристал германію завтовшки 1 мм, що освітлюється світлом з енергією фотонів $\hbar\omega_{phot} = 0.82 \text{ eV}$. Потік світла становить $10^{18} \text{ фот / см}^2\text{с}$, коефіцієнт поглинання – $\alpha = 10 \text{ см}^{-1}$. Вважати, що ймовірність переходу не залежить від енергії, а процес випромінювання визначається непрямыми переходами за участю фононів з енергією $\hbar\omega_{phon} = 0.03 \text{ eV}$. Використати такі дані: час випромінювальної рекомбінації $\tau_r = 0.1 \text{ s}$, ширина забороненої зони $E_g = 0.72 \text{ eV}$, ефективні маси $m_e^* = 0.2m$ (4 долини), $m_h^* = 0.3m$, коефіцієнт заломлення $n = 4$.

4. Визначити форму спектру випромінювання у стаціонарному режимі для попередньої задачі

Список літератури

1. Смит Р. Полупроводники. – М.: Мир, 1982.
2. Панков Ж. Оптические процессы в полупроводниках,. – М.: Мир, 1973.

3. Зеегер К. Физика полупроводников. – М.: Мир, 1977.
4. Богданкевич О.В., Дарзбек С.А, Елисеєв П.Г.. Полупроводниковые лазеры.– М.: Наука, 1976
6. Kasap S.O., Principles of Electronic Materials and Devices, McGraw-Hill, 2005.

Розділ 16

НИЗЬКОВИМІРНІ НАПІВПРОВІДНИКОВІ СИСТЕМИ

Зменшення розмірів твердого тіла спричиняє очевидне підсилення впливу поверхні на його фізичні властивості – оптичні, електронні, електрофізичні, зокрема за розмірів частинок істотно менших за довжину хвилі зовнішнього світла, але набагато більших за сталу ґратки спектр поглинання світла стає залежним від форми частинки. Виникають так звані розмірні ефекти. За подальшого зменшення розмірів частинки виникають суттєво мікроскопічні ефекти – так звані квантово-розмірні ефекти.

Таким чином, виникає питання – а що таке низьковимірні системи? Очевидно, в ідеальному випадку, низьковимірні системи можна визначити як такі, розмірність яких менша за три. До низьковимірних систем належали би дво-, одно- та нульвимірні системи. Але в реальному житті всі системи, з яким доводиться стикатись, є тривимірними. І все ж існують ефекти, пов'язані із розмірністю, і таким чином можна визначити низьку розмірність реальних об'єктів. Ми визначатимемо низьковимірний об'єкт як такий, що має принаймні вздовж однієї з осей системи координат характерні розміри, набагато менші від характерних параметрів розмірності довжини зовнішніх "впливів". При цьому для існування розмірних ефектів у низьковимірних системах необхідно, щоб розмір системи був порівняним із деякими характерними довжинами, що відіграють вирішальну роль при виникненні тих чи інших ефектів. Наприклад, при розгляді квантово-розмірних ефектів роль внутрішніх параметрів для тривимірної системи може відігравати довжина вільного пробігу носіїв, довжина хвилі де Бройля тощо. У випадку класичних розмірних ефектів таким параметром розмірності довжини може бути довжина хвилі зовнішнього випромінювання при вивченні оптичних властивостей, або довжина вільного пробігу носіїв при розгляді впливу товщини плівки на її провідність.

Таким чином, існує можливість спостереження як класичних, так і квантових розмірних ефектів. Розглянемо обидва типи ефектів, маючи на увазі, що вони можуть використовуватись у сучасних електронних і мікроелектронних пристроях та є яскравими прикладами нових галузей фізики напівпровідників.

■ 16.1. Квантові та класичні розмірні ефекти. Короткий огляд

Визначатимемо *квантові розмірні ефекти* як такі, що спричиняють зміни термодинамічних, кінетичних та оптичних властивостей системи, якщо хоча б один із геометричних розмірів зразку стає сумірним із довжиною хвилі де Бройля λ_B електронів об'ємного зразку того самого матеріалу. Ці ефекти зумовлені квантуванням руху електронів у напрямку, в якому розмір системи є порівняним із λ_B . У масивному ідеальному кристалі стан електронів визначається чотирма квантовими числами – трьома компонентами квазіімпульсу та номером енергетичної зони. У кристалічній плівці товщиною L (рис. 16.1) рух електронів у площині плівки залишається вільним. Компоненти квазіімпульсу k_x і k_y можуть, як раніше, набувати будь-яких значень, дозволених зонною теорією.

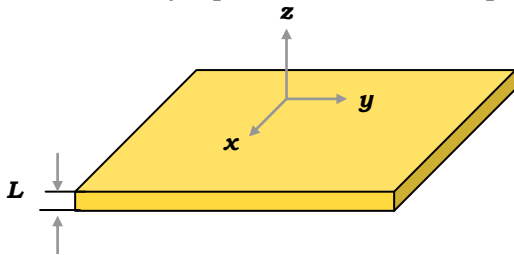


Рис. 16.1. Схема квазідвовимірної системи

Вздовж вісі OZ квазіімпульс може мати тільки дискретний ряд значень, оскільки хвильова функція електрона на поверхні плівки має дорівнювати нулю. Таким чином, на товщині L плівки має вклатись ціле число напівхвиль де Бройля $\lambda_B / 2 = \pi \hbar / |p_z|$. Звідси маємо закон квантування проекції квазіімпульсу p_z

$$|p_z| = \frac{\pi \hbar}{L} n. \quad (16.1)$$

У наближенні ефективної маси закон квантування енергії поперечного руху визначається формулою

$$E_n(\mathbf{k} = 0) = \frac{\pi^2 \hbar^2}{2m_z^* L^2} n^2. \quad (16.2)$$

Тут m_z^* – ефективна маса електрона у напрямку OZ. Рівні енергії (16.2) називаються рівнями розмірного квантування. Графічно енергія носіїв $E_n(\mathbf{k})$ для різних значень z-компоненти квазіімпульсу (що в дійсності означає – для різних значень n) являє собою набір параболоїдів, вставлених один в один так, щоб дно кожного наступного було розташоване вище попереднього. Дно параболоїда $E_n(\mathbf{k} = 0)$ відповідає енергії руху у поперечному щодо площини плівки напрямку, тобто спектр носіїв складається з набору двовимірних розмірних підзон, кожна з

яких містить значення енергії для всіх можливих значень компонентів k_x та k_y за заданого p_z . За зменшення товщини L енергія рівнів $E_n(\mathbf{k} = 0)$ зростає, і як наслідок – збільшується відстань між підзонами. Таким чином, енергетичний спектр носіїв у тонкій плівці має вигляд

$$E_n(\mathbf{k}) = E_n(\mathbf{k} = 0) + \frac{\hbar^2 k_x^2}{2m_x^*} + \frac{\hbar^2 k_y^2}{2m_y^*}. \quad (16.3)$$

Отже електрони у плівці утворюють двовимірний електронний газ, якщо вони заповнюють одну або кілька двовимірних підзон.

Важливою характеристикою електронних станів є їхня щільність, що має сенс концентрації електронів, яка приходить на одиничний інтервал енергії. У подальшому розглянемо детально, як обчислюються щільності електронних станів для низьковимірних систем, а поки що – деякі якісні висновки щодо поведінки щільності електронних станів. Якщо, наприклад у масивному зразку напівпровідника, щільність станів

$$N(E) \sim \sqrt{E}, \quad (16.4)$$

то в тонкій плівці розмірна підзона дає сталий (що для квадратичного закону дисперсії не залежить від енергії) внесок до щільності станів, який на одиницю площі плівки дорівнює

$$\Delta N(E) = s \frac{m^*}{2\pi\hbar^2}, \quad (16.5)$$

де s – кратність спінового та міждолинного виродження. А повна щільність станів, як буде далі, є східчастою функцією енергії E . При цьому n -й стрибок відбувається за енергії $E = E_n(\mathbf{k} = 0)$ і відбиває появу або зникнення внеску n -ї підзони.

Для сталих значень енергії щільність станів на одиницю об'єму плівки як функція її товщини L зазнає стрибків за значень товщини

$$L = n \frac{\lambda_B}{2}. \quad (16.6)$$

Між цими стрибками щільність станів змінюється обернено пропорційно товщині. Період осциляцій за товщиною для щільності станів

$$\Delta L = \frac{\pi\hbar}{\sqrt{2m^* E}}. \quad (16.7)$$

Оскільки електронні властивості металів, напівметалів і вироджених напівпровідників визначаються поведінкою електронів з енергіями у вузькому інтервалі (завширшки $\sim kT$) навколо поверхні Фермі (тобто із центром E_F), то термодинамічні й кінетичні характеристики кванто-

во-розмірних систем залежать від щільності станів на рівні Фермі $N(E_F)$. Таким чином, для квантових ям стрибкоподібна поведінка $N(E_F)$ за зміни товщини ями L приводить до осцилюючої залежності від L питомого опору, константи Холла та магнетопровідності. Крім того, виникають особливості в тунельних характеристиках тонких металічних плівок (напр. такі, що спостерігались у плівках Pb, Mg, Au, Ag). Дійсно, тунелювання електронів наводить безпосередні докази існування рівнів розмірного квантування у шаруватих системах із надтонкими шарами. Відомо, що ймовірність тунелювання електронів через бар'єр визначається параметрами, що характеризують бар'єр, і щільністю початкових і кінцевих станів. Таким чином, у тонкоплівковій системі **метал-діелектрик-метал** особливості щільності станів визначатимуть особливості тунельного струму I_T від напруги U . Найбільш цікавими з погляду практичних застосувань є квантово-розмірні ефекти у напівпровідникових гетероструктурах. Наприклад, у гетероструктурах із квантовими ямами $\text{Al}_x\text{Ga}_{1-x}\text{As-GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ рух електронів обмежено шарами GaAs, а шари $\text{Al}_x\text{Ga}_{1-x}\text{As}$ відіграють роль потенціальних бар'єрів. Якщо товщина бар'єрних шарів не дуже мала, то гетероструктуру можна розглядати як набір не пов'язаних між собою плівок GaAs. При цьому, якщо товщина плівок становить величину $\sim 10^{-7}$ - 10^{-6} см, то розмірне квантування спричинить ефективну модифікацію енергетичних станів структури, що істотно змінює, наприклад їхні оптичні властивості. Розмірне квантування зумовлює зсув дна зони провідності E_C в арсеніді галію на величину $E_1(\mathbf{k} = 0)$. Це приводить до ефективного збільшення ширини забороненої зони, що викликає зсув червоної межі смуги поглинання у жовтий бік залежно від товщини квантової ями. Квантово-розмірні явища спостерігаються для тонких шарів GaAs – завтовшки порядку десятків нанометрів. Піки у спектрі поглинання зумовлені переходами електронів із заповненої n -ї підзони у валентній зоні до порожньої n -ї підзони у зоні провідності. Аналогічні особливості спостерігаються й в спектрах люмінесценції.

Залежність оптичних властивостей систем із квантовими ямами від їхньої товщини використовується для створення лазерів із вдосконаленими характеристиками. Зокрема, у таких лазерах спостерігається короткохвильовий зсув частот випромінювання, зниження порогової потужності накачування тощо.

Слід зазначити, що квантово-розмірні ефекти спостерігаються лише у досконалих та однорідних за товщиною плівках. Кількісно це означає, що поширення ліній рівнів δE розмірного квантування через розсію-

вання носіїв заряду на домішках, фононах і шорсткостях поверхні півки має бути малим щодо відстані між енергетичними рівнями розмірного квантування ΔE . При цьому флуктуації товщини півки мають бути меншими за довжину хвилі електрона на рівні Фермі λ_F . Осциляції, обумовлені квантово-розмірними ефектами, найбільш яскраво виявляються у тонких півках, де $L \sim \lambda_F$ за низьких температур коли температурне розмиття щільності станів $N(E)$ є меншим за ΔE ($kT \ll \Delta E$). Таким вимогам до систем краще за все відповідають епітаксіальні шари (типу шарів GaAs у багатошарових гетероструктурах), а також півки напівметалів (Bi, Sn, Sb та їхні сплави) і вироджених напівпровідників із вузькою забороненою зоною (InSb, PbTe) в інтервалах товщин $L \sim 10^3 - 10^4$ нм. У металічних півках через малі значення λ_F вимогу однорідності півки за товщиною досить важко виконати.

Крім півок і гетероструктур, квазідвовимірний електронний газ можна отримати й поблизу меж розподілу на контактах між двома напівпровідниками в інверсних і збагачених шарах МДН-структур і модуляційно-легованих гетероструктурах. Важливою перевагою цих систем над півковими є можливість керованої зміни концентрації носіїв у широких межах. Модуляційно-леговані гетероструктури, які складаються із переходів GaAs-AlGaAs (рис. 16.2), вирощуються зазвичай методами молекулярно-променевої епітаксії. При

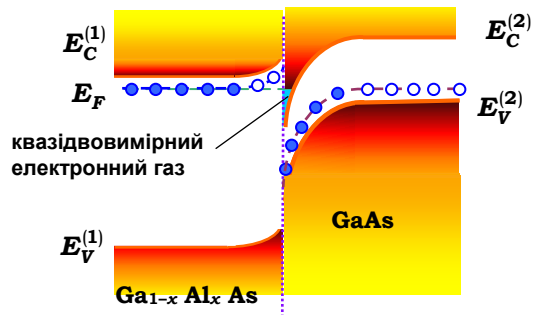


Рис. 16.2. Утворення каналу квазідвовимірних електронів на межі розподілу модуляційно-легованої гетероструктури.

Синіми кульками позначено домішкові рівні, заповнені електронами, білими – дірки. Рівень Фермі збігається із рівнем донорної домішки в AlGaAs

цьому шар або частина шару AlGaAs легується (концентрація донорів $\sim 10^{18} \text{ см}^{-3}$), а шар GaAs не легується або слабо легується акцепторами. Таке легування спричиняє викривлення зон, оскільки електрони, намагаючись встановити єдине значення рівня Фермі в системі, переходять із донорів до потенціальної ями, що утворюється вигином зон, з одного боку, і розривом дна зони провідності на гетеропереході – з іншого. Такі електрони можуть вільно рухатись тільки вздовж границі гетеропереходу. Квантування поперечного руху в ямі (аналог розмір-

ного квантування у плівці) приводить до виникнення квазідвовимірного електронного газу із поверхневою концентрацією 10^{11} – 10^{12} см^{-2} . У такій системі поверхневі стани та дефекти (завдяки відповідності ґраток GaAs та AlGaAs) не впливають суттєво на формування фізичних властивостей. З іншого боку, розсіювання електронів на домішках, що їх породили, також незначне через їхнє просторове розділення. За рахунок цього досягаються високі значення рухливості електронів. Наприклад, за температури 4К рухливість електронів може досягати величин $\mu = 10^6$ $\text{см}^2/\text{Вс}$, за температури 77К – $\mu = 10^5$ $\text{см}^2/\text{Вс}$, і за 300К – $\mu = 10^4$ $\text{см}^2/\text{Вс}$. Завдяки таким високим значенням рухливостей польові транзистори, створені на модуляційно-легованих гетероструктурах із квазідвовимірним електронним газом, характеризуються великою швидкістю ≥ 10 пс в інтервалі температур 77–300 К. Останнім часом все більшу увагу як дослідники, так і практики приділяють системам із квантовими точками (системи, в яких квантово-розмірні об'єкти обмежені у просторі вздовж всіх трьох осей). У таких системах електронні стани квантуються, подібно станам у тривимірному потенціальному ящику.

Специфіка фізичних ефектів у низьковимірних системах, що не пов'язана із просторовим квантуванням, полягає, як зазначалось вище, у наявності меж розподілу. Такі (класичні розмірні ефекти) можуть спостерігатись у тонких плівках і малих частинках. Історично першим класичним ефектом, що пов'язаний із пониженою розмірністю системи, було визначення залежності провідності тонких плівок металів і сильно легованих напівпровідників від товщини плівки. Виявилось, що додаткове розсіювання носіїв границями плівки спричиняє зменшення провідності тонких плівок. Вплив розмірів зразка (товщини плівки) на його електропровідність стає помітним, коли ці розміри стають порівняними із довжиною вільного пробігу електрона. Таким чином, розмірні ефекти різкіше виражені у чистих монокристалічних плівках, оскільки довжина вільного пробігу в них істотно перевищує таку довжину у розупорядкованих із великою кількістю домішок зразках. До класичних розмірних ефектів можна зарахувати також можливість збудження поверхневої хвилі (поверхневого поляритону або плазмон-поляритону) *p*- та *s*-поляризації, якщо поверхня, де збуджується хвиля, вкрита дуже тонкою плівкою іншого матеріалу. Оскільки ці ефекти походять від наявності поверхні, де локальні поля дуже неоднорідні, то зрозуміло, що головним чином вони пов'язані з нелокальними електродинамічними взаємодіями. А оскільки такі ефекти розігруються на відстанях, що порівняні із товщиною перехідного шару поблизу поверхні, які в свою чергу є набагато меншими за довжину

хвилі зовнішнього випромінювання (якщо йдеться про оптичні властивості), то доводиться мати справу зі специфічними розподілами електромагнітного поля, коли не можна говорити про довжину хвилі, оскільки на таких відстанях фронт хвилі ще не може сформуватись. У такому випадку кажуть, що фізичні ефекти розігруються у ближньому полі. Розвиток сучасних експериментальних методів вивчення властивостей мезооб'єктів (таких, лінійні розміри яких досить малі для того, щоб ефекти нелокальності та неоднорідності поля відігравали суттєву роль, а з іншого боку ці розміри досить великі для того, щоб об'єкти ще можна було описувати в класичному наближенні) спричинив виникнення цілої нової галузі фізики – фізики ближнього поля.

■ 16.2. Електронні фундаментальні довжини

Як ми обговорювали у попередніх розділах, електрон у напівпровіднику характеризується ефективною масою m^* , яка зазвичай менша за масу вільного електрона. Таким чином довжина хвилі де Бройля в напівпровіднику є більшою за довжину хвилі де Бройля вільного електрона

$$\lambda_B = \frac{h}{p} = \frac{h}{\sqrt{2m^*E}} = \lambda_B^0 \sqrt{\frac{m}{m^*}} \quad (16.9)$$

Відомо, що відношення ефективних мас електронів для InSb $m^*/m = 0.014$, для GaAs – $m^*/m = 0.067$, для GaN – $m^*/m = 0.172$, для SiC – $m^*/m = 0.41$. Тоді із (16.9) маємо абсолютні значення довжин хвиль де Бройля для цих напівпровідників (рис. 16.3). Зясуємо, за яких розмірів системи можна очікувати ефекти розмірного (просторового) квантування. Нехай геометричні розміри напівпровідникового зразка

$$L_x \times L_y \times L_z.$$

Припустимо, що $L_x < L_y < L_z$. Якщо система, що розглядається, є однорідною, а процеси розсіювання електронів є слабкими, то можна вважати рух електронів у такому зразку квазібалістичним, і тільки лінійні розміри системи необхідно порівнювати із довжиною хвилі де Бройля. Оскільки

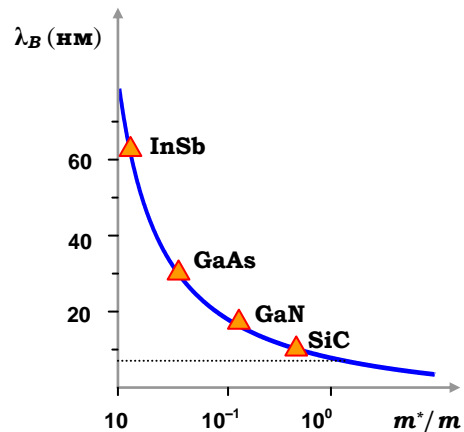


Рис. 16.3. Довжини хвиль електронів у напівпровідниках при $T=300$ К залежно від ефективної маси

тільки ціле число напівхвиль де Бройля може вкластися вздовж будь-якої обмеженої системи, дістанемо, що тільки дискретна кількість електронних станів може існувати у такій системі. Таким чином, рух електрона в обмеженому зразку може бути квантований. При цьому умови квантування можна визначити ієрархією довжин:

♦ **тривимірний рух електронів** можливий у тривимірному об'ємному зразку, де

$$\lambda_B \ll L_x, L_y, L_z. \tag{16.9}$$

У цьому випадку електрон поводитья як вільна частинка з ефективною масою m^* .

♦ **двовимірна система** або квантова яма – якщо квантування відбувається вздовж однієї з осей i (наприклад, Ox), а вздовж інших двох напрямків зберігається квазивільний рух

$$\lambda_B \cong L_x \ll L_y, L_z; \tag{16.10}$$

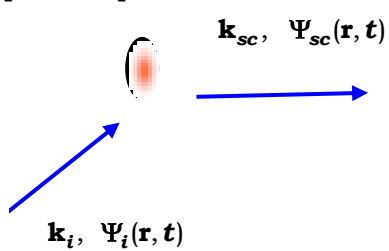
♦ **одновимірна система** або квантовий дріт¹ (квантова нитка – quantum wire) – якщо квантування відбувається вздовж двох напрямків так, що електрон рухається вільно лише у напрямку розповсюдження дроту

$$L_x \cong L_y \cong \lambda_B \ll L_z; \tag{16.11}$$

♦ **нульвимірна система** або квантова точка – якщо квантування відбувається вздовж усіх трьох осей системи

$$\lambda_B \cong L_x \cong L_y \cong L_z. \tag{16.12}$$

Тепер візьмемо до уваги, що у реальних системах електрони піддаються процесам розсіювання, що спричиняє стохастизацію їхнього руху. При цьому вони перестають поводитися як квантові об'єкти. Процеси розсіювання поділяються на дві групи – пружні та непружні.



За пружного розсіюванні з погляду класичної фізики відбувається тільки передача імпульсу, а непружні процеси супроводжуються передачею як імпульсу, так й енергії. Розглянемо вплив різних типів процесів розсіювання на квантову поведінку електронів. Нехай, наприклад електрон пружно розсіюється домішкою (рис. 16.4). При цьому до розсіювання він характеризувався хвильовою функцією із хвильовим век-

Рис. 16.4. Зміна хвильового вектора та хвильової функції за пружного розсіювання електрона

¹ Такі системи ще називають квантовими голками (quantum needle)

тором $\mathbf{k}_i = \mathbf{p}_i / \hbar$. Після розсіювання його хвильовий вектор зміниться – $\mathbf{k}_{sc} = \mathbf{p} / \hbar$, а енергія не зміниться – $E(p) = E(p_i) = \hbar\Omega$. Тоді маємо, що хвильова функція електрона, яка до розсіювання мала вигляд

$$\Psi_i(\mathbf{r}, t) = e^{-i\Omega t} e^{i\frac{\mathbf{p}_i \mathbf{r}}{\hbar}}, \quad (16.13)$$

перейде у хвильову функцію

$$\Psi_{sc}(\mathbf{r}, t) = e^{-i\Omega t} \sum_{\mathbf{p}, |\mathbf{p}|=p_i} A_{\mathbf{p}} e^{i\frac{\mathbf{p} \mathbf{r}}{\hbar}} = e^{-i\Omega t} \psi(\mathbf{r}). \quad (16.14)$$

При цьому як завжди $|A_{\mathbf{p}}|^2$ визначає ймовірність знаходження частинки з деяким моментом \mathbf{p} після розсіювання. Падаюча та розсіяна хвилі утворюють комплексну хвильову картину, і за пружного розсіювання важливою властивістю цієї картини стане збереження фази електрона. У результаті квадрат модуля хвильової функції (16.14) залишається не залежним від часу

$$|\Psi_{sc}(\mathbf{r}, t)|^2 = |\psi(\mathbf{r})|^2. \quad (16.15)$$

Іншими словами, пружне розсіювання не порушує когерентність електронного руху.

Якщо τ_e – середній час між двома актами пружного розсіювання, то можна визначити середню довжину пробігу електрона між цими актами як

$$l_e = v\tau_e, \quad (16.16)$$

де v – середня швидкість електронів. Таким чином, на відстанях, що перевершують величину l_e , електрони можна розглядати як когерентні

хвильові утворення. У разі непружного розсіювання електронів (рис. 16.5) замість (16.14) запишемо

$$\Psi_{sc}(\mathbf{r}, t) = \sum_{\mathbf{p}, |\mathbf{p}|=p_i} A_{\mathbf{p}} e^{i\frac{\mathbf{p} \mathbf{r}}{\hbar}} e^{-i\Omega(\mathbf{p})t}. \quad (16.17)$$

Тепер квадрат модуля хвильової функції залежить не тільки від координати, а також від часу, тобто когерентність втрачається. Нехай τ_E – середній час між двома актами непружного розсіювання, а середня довжина пробігу електрона за непружного розсіювання – L_E . Електрон зберігає квантову когерентність на відстанях, менших за L_E . Зазвичай L_E набагато перевищує довжину l_e , тобто

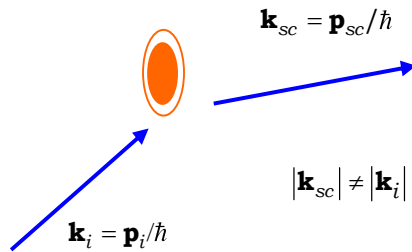


Рис. 16.5. Непружне розсіювання електрона

менших за L_E . Зазвичай L_E набагато перевищує довжину l_e , тобто

електрон багато разів розсіюється перед тим, як його енергія помітно зменшується (рис. 16.6). У результаті його рух є хаотичним, і середнє переміщення його за час τ_E визначається коефіцієнтом дифузії D

$$L_E = \sqrt{D\tau_E}, \quad (\tau_E \gg \tau_e) \quad (16.18)$$

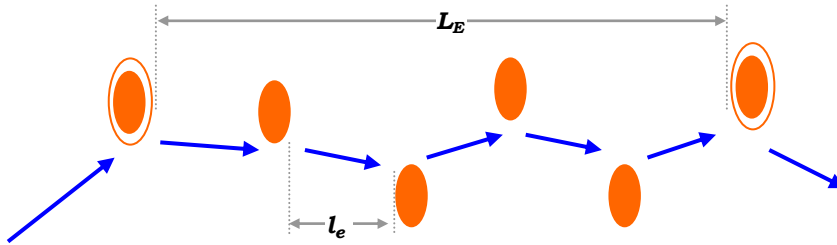


Рис. 16.6. Стохастизація руху електрона за пружного та непружного розсіювання

Коефіцієнт дифузії можна подати формулою

$$D = \frac{v^2 \tau_e}{\alpha}, \quad (16.19)$$

де параметр $\alpha = 3$ для тривимірного електронного газу, $\alpha = 2$ – для двовимірного і $\alpha = 1$ – для одновимірного.

За збільшення температури число електронів, що можуть брати участь у транспорті, збільшується та все більша кількість електронів із різною енергією можуть брати участь у кінетичних процесах. Когерентність у системі при цьому порушується через широкий розподіл фаз хвильових функцій. Дійсно, нехай $k_B T \ll E_F$. Тоді кожен електронний стан має майже однакову амплітуду, але трохи різні фази. Поки ця різниця є малою, усереднення за електронними станами всередині температурно-розмитої функції розподілу електронів не порушує їхню фазу і, тим самим, зберігає когерентність електронів. Але за збільшення температури, коли електрони з істотно різними енергіями беруть участь у кінетичних процесах, когерентність у системі порушується, оскільки усереднення за станами відбувається із значно розмитою функцією розподілу електронів. Нехай L_T – характерна довжина, на якій порушується когерентність електронів завдяки ненульовій температурі. Температурне розмиття функції розподілу пропорційне $k_B T$ і призводить до того, що розмиття фаз змінюється з часом як $t \cdot (k_B T / \hbar)$. Таким чином, для оцінки часу розмиття фаз можна використати величину

$$\tau_T = \frac{\hbar}{k_B T}. \quad (16.20)$$

Якщо процеси розсіювання є пружними, то електрон дифундує в просторі на відстань $\sim\sqrt{Dt}$ за час t , що перевищує середній час вільного пробігу τ_e . Таким чином, протягом часового інтервалу τ_T електрон продифундує завдяки термічному руху на довжину

$$L_T = \sqrt{D\tau_T} = \sqrt{D\hbar/k_B T}. \quad (16.21)$$

Отже когерентність електронів порушується на відстанях, що перевершують довжину L_T .

Зрозуміло, що ефекти дефазування, викликані непружним розсіюванням і температурним розмиванням функції розподілу, діють одночасно. Тоді просторовий масштаб, пов'язаний із втратою квантово-механічної когерентності, можна визначити як меншу з двох довжин L_E та L_T

$$l_\phi \cong \min\{L_E, L_T\}. \quad (16.22)$$

Оскільки електронний транспорт визначається суперпозицією електронних хвильових функцій, із вищенаведеного розгляду можна зробити висновок, що довжина когерентності l_ϕ визначає просторовий розмір, нижче якого електронний транспорт має квантовий характер.

Можна також розглянути нерівноважну біполярну систему, що містить як електрони, так і дірки. Тоді, якщо τ_R – час життя носіїв у системі, то довжина $L_R = \sqrt{D\tau_R}$ визначатиме середню відстань, на якій нерівноважні носії дифундують, перед тим як рекомбінувати.

Таким чином, прилади, геометричні розміри яких становлять порядок довжини когерентності, не можна описувати макроскопічними матеріальними параметрами – провідністю, дрейфовою швидкістю тощо. Такі прилади необхідно описувати в рамках квантової теорії. З іншого боку, для приладів, в яких транспортні відстані L_{tr} є більшими порівняно із довжиною когерентності l_ϕ можна описувати в рамках класичної фізики.

Розглянемо тепер умови реалізації квантового та класичного режимів електронного транспорту. Для цього припустимо, що транспорт відбувається вздовж однієї вісі, тобто повний струм уздовж інших двох осей, є нульовим. Але поперечні розміри приладу (вздовж осей) можуть відігравати важливу роль. Можна визначити кілька принципово різних типів транспорту електронів у просторово обмежених системах. Почнемо із *квантового та мезорежимів транспорту*. Якщо довжина хвилі де Бройля перевищує лінійні розміри приладу L_z

$$\lambda_B \geq L_z, \quad (16.23)$$

і крім цього, виконується нерівність

$$l_e \gg \lambda_B, \quad (16.24)$$

то маємо квантово-балістичний режим електронного транспорту. Випадає, коли

$$l_\phi > L_z, \lambda_B \quad (16.25)$$

відповідає мезорежиму електронного транспорту. Таким чином, можна стверджувати, що поперечні (до вісі OZ) розміри приладу визначають одно-, дво- та тривимірний характер квантово-балістичного та мезо- режимів електронного транспорту.

Класичний режим електронного транспорту реалізується у випадку, якщо виконується нерівність

$$L_z > l_\phi. \quad (16.26)$$

Якщо лінійний розмір приладу вздовж вісі OZ при цьому є меншим за довжину вільного пробігу електрона, то відбувається класичний балістичний режим електронного транспорту. Іншими словами, у цьому випадку електрони можуть рухатись уздовж приладу за класичними траєкторіями без зіткнень. Якщо

$$L_z \gg l_\phi, \quad (16.27)$$

то електронний транспорт має дифузійну природу. У випадку, коли виконуються співвідношення

$$L_z \sim l_E \gg l_e, \quad (16.28)$$

електрони не втрачають енергію, рухаючись уздовж приладу. Тоді кажуть про реалізацію квазібалістичного режиму електронного транспорту. Зрозуміло, що можуть реалізуватись три класичних режими електронного транспорту для одно-, дво- та тривимірних рухів електронів. Якщо обидва поперечні розміри приладу L_x та L_y набагато більші за довжину хвилі де Бройля, але порівняні із характеристичними класичними довжинами, спостерігатимуться поперечні класичні розмірні ефекти в електронному транспорті. У цьому випадку зіткнення електронів із поперечними поверхнями приладу визначатимуть транспорт вздовж вісі OZ. Наприклад, якщо один із двох поперечних розмірів приладу має порядок довжини вільного пробігу електронів $L_x, L_y \sim l_e$, опір приладу вздовж вісі OZ залежатиме від його бічних меж. Якщо поперечні розміри приладу стають порівняними з однією із дифузійних довжин, ми отримуємо інший тип класичних розмірних ефектів. Наприклад, якщо L_x або L_y мають порядок довжини релаксації енергії L_E , границі приладу забезпечують додатковий канал релаксації

енергії. Такий дифузійно-розмірний ефект визначає середню енергію нерівноважних електронів.

На останок зведемо ієрархію довжин електронного транспорту до таблиці

Таблиця 16.1

Квантовий режим	→	Міжконтактна відстань L_z є порівняною із довжиною хвилі електрона $L_z \sim \lambda_B$		
Мезорежим	→	Міжконтактна відстань L_z є меншою за довжину, на якій відбувається дефазування хвильової функції електронів $L_z \leq l_\phi$		
Класичний режим (одно-, дво- або тривимірний електронний транспорт)	→	Міжконтактна відстань L_z перевищує довжину, на якій відбувається дефазування хвильової функції електронів $L_z > l_\phi$		
		класичний балістичний режим: $l_e \geq L_z$	квазібалістичний режим (енергія зберігається): $L_E \geq L_z \geq l_e, l_\phi$	поперечні розмірні ефекти
			середньої довжини вільного пробігу: $L_x, L_y \sim l_e$	дифузійні: $L_x, L_y \sim L_E, L_R$

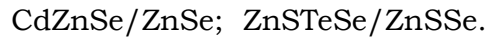
■ 16.3. Гетероструктури. Квантові ями, квантові голки, квантові точки

Гетероструктура формується, коли два різних матеріали з'єднуються в одну структуру із межею розподілу між ними. Умови формування гетероструктур і різні типи гетерограниць ми розглядали у розд. 13. Скористаємось цими відомостями та розглядатимемо гетероструктури, в яких може спостерігатись просторове кантування. Найпростішими та найпоширенішими подібними гетероструктурами є квантові ями. Властивостям квантових ям, а також квазіодновимірних – квантових голок і квазінульвимірних – квантових точок – об'єктів й буде присвячено обговорення у цьому розділі.

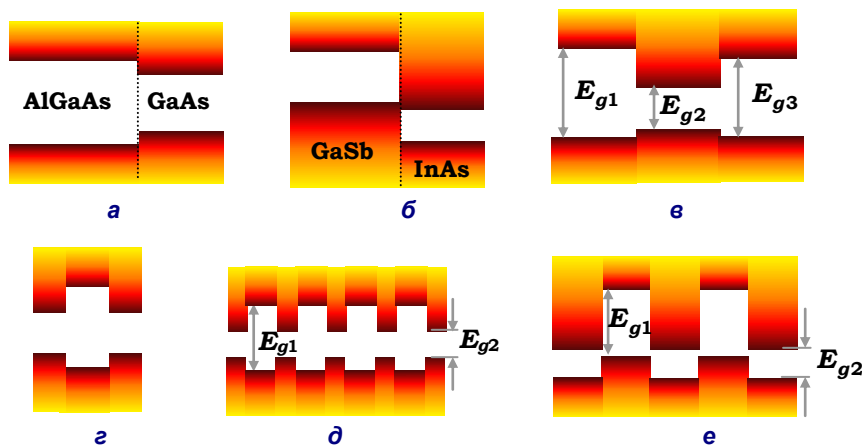
Сучасні технології дозволяють виготовляти ці межі розподілу майже ідеальними. Наприклад, перехідний шар, що утворюється на межі розподілу завдяки досягненням сучасних технологій (базуються на

методі молекулярно-променевої епітаксії) може бути доведений до моноатомного. Найпростішим прикладом гетероструктури є два напівпровідникових матеріали з однією межею розподілу (гетероконтактом). Відомими прикладами гетероструктур можна назвати SiO_2/Si – гетероз'єднання з дуже малою щільністю дефектів на інтерфейсі. Можна вказати, наприклад на такі широкоуживані пари III-V-напівпровідників, що утворюють промислові гетероструктури $\text{GaAs}/\text{AlGaAs}$; GaInAs/InP ; $\text{GaInAs}/\text{AlInAs}$; GaSb/AlSb ; GaN/AlN .

Іншим прикладом гетероструктур є структури сполуки II-VI



Головною метою при отриманні гетероструктур є контрольована модифікація енергетичних зон носіїв у них. Інакше кажучи, гетероструктури можна розглядати як штучні речовини із наперед заданими властивостями. На рис. 16.7 подані зонні діаграми для деяких напівпровідникових гетероструктур. Рис. 16.7 а-б демонструють два типи гетероструктур з однією межею розподілу, рис. 16.7 в-г – різні типи гетероструктур із двома межами розподілу, а рис. 16.7 д-е – багатоінтерфейсні гетероструктури. Далі розглянемо ідеалізовану картину трьох типів квантових структур – квантові ями, дроти та точки.



16.7. Енергетичні діаграми різних гетероструктур

■ 16.4. Квантові ями

Квантова яма може утворитись, коли тонкий шар з одного матеріалу вирощується між двома іншими шарами із різних матеріалів, що характеризуються ширшою заборненою зоною, ніж матеріал центра-

льного шару (рис. 16.8). Розривність зон при цьому приводить до обмеженості руху (конфайнменту) носіїв усередині ями. Рух електрона тепер відбуватиметься у потенціалі. Для більшості випадків квантових ям такий потенціал можна вважати прямокутним. Розглянемо ідеалізований потенціал (рис. 16.8).

$$V(z) = \begin{cases} 0 & \text{при } |z| \leq L/2, \\ V_b & \text{при } |z| \geq L/2, \end{cases} \quad (16.29)$$

де V_b і L – глибина та ширина ями, відповідно. Оскільки потенціал залежить тільки від однієї змінної, можна застосувати метод змінних, що розділяються: рух електронів у площині $ХОУ$ буде вільний, тоді хвильову функцію можна представити у вигляді

$$\Psi(x, y, z) = e^{ik_x x + ik_y y} \chi(z). \quad (16.30)$$

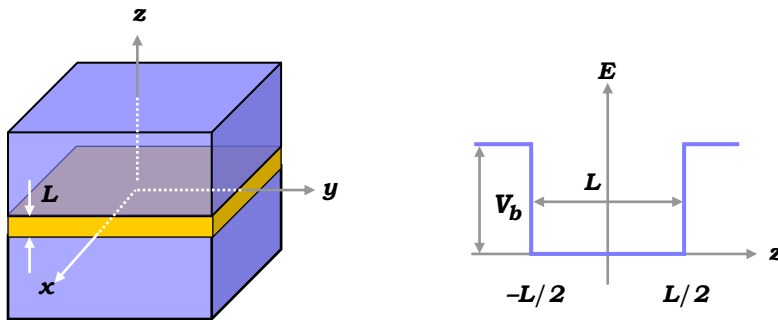


Рис. 16.8. Квантова яма як тонка плівка напівпровідника між товстих шарів матеріалу із широкою забороненою зоною та ідеалізована енергетична діаграма такої структури

Після підстановки цієї хвильової функції до рівняння Шредингера

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E \psi(\mathbf{r}), \quad (16.31)$$

отримаємо рівняння для невідомої функції $\chi(x)$

$$\left[-\frac{\hbar^2}{2m^*} \frac{\partial^2}{\partial z^2} + V(z) \right] \chi(z) = \left(E - \frac{\hbar^2 k_{\square}^2}{2m^*} \right) \chi(z). \quad (16.32)$$

Тут $k_{\square} = (k_x, k_y)$, а $(\hbar^2 k_{\square}^2)/(2m^*)$ – кінетична енергія руху електрона з ефективною масою m^* у площині $ХОУ$. Введемо позначення

$$\epsilon = E - \frac{\hbar^2 k_{\square}^2}{2m^*}. \quad (16.33)$$

Оскільки E – повна енергія частинки, то ϵ буде енергією руху частинки у напрямку, нормальному до площини XOY , тобто поперечного руху. Таким чином, ми розділили змінні: поперечну (z) від площинних (x, y) і тим самим звели проблему до одновимірної

$$\left[-\frac{\hbar^2}{2m^*} \frac{d^2}{dz^2} + V(z) \right] \chi(z) = \epsilon \chi(z). \quad (16.34)$$

16.4.1. Зв'язані стани електронів

Згідно із загальними властивостями рівняння Шредингера можна очікувати наявності двох типів його розв'язків: при $\epsilon < V_b$ електрони будуть зв'язані в ямі, а при $\epsilon > V_b$ рух електрона буде необмежений, а стани незв'язані. Існування ненульового розв'язку (16.34) за умови $\epsilon < V_b$ можливе лише, якщо енергія електрона набуватиме дискретних значень, тобто виконуються дисперсійні співвідношення. Оскільки рухи вздовж осі OZ та у протилежному напрямку є еквівалентними, то хвильова функція може мати вигляд як парних, так і непарних розв'язків. При цьому умови існування парного розв'язку мають вигляд (див. додаток Н)

$$\operatorname{tg} k_w L / 2 = k_b / k_w, \quad \operatorname{tg} k_w L / 2 \geq 0. \quad (16.35)$$

Аналогічно, для непарного розв'язку

$$\operatorname{ctg} k_w L / 2 = -k_b / k_w, \quad \operatorname{ctg} k_w L / 2 \leq 0. \quad (16.36)$$

Легко показати (див. додаток Н), що у будь-якій ямі завжди існує хоча б один енергетичний рівень

$$\epsilon_1 = \frac{\hbar^2 k_{w,1}^2}{2m^*}. \quad (16.37)$$

Якщо висота потенціального бар'єру достатня, то в ямі може формуватися другий енергетичний рівень з енергією, що майже дорівнює V_b

$$\epsilon_2 \approx V_b. \quad (16.38)$$

Загалом при $\epsilon < V_b$ у потенціальній ямі може формуватися кількість рівнів, що визначається формулою

$$N = 1 + \operatorname{int} \left[\sqrt{\frac{2m^* V_b L^2}{\pi^2 \hbar^2}} \right]. \quad (16.39)$$

Таким чином, варіюючи параметри квантової ями – підбираючи компоненти структури із певною шириною забороненої зони і задаючи

товщину шару напівпровідника – можна отримати систему, що характеризується наперед заданою структурою енергетичних рівнів, тобто отримувати штучні матеріали. На рис. 16.9 подано кількість зв'язаних у квантовій ямі станів як функцію ширини ями для двох значень її глибини у системі AlGaAs/GaAs. Видно, що за однієї і тієї самої ширини ями кількість рівнів більша в ямі, що характеризується більшою ефективною масою носіїв. З іншого боку, у квантових ямах із невеликими значеннями ефективної маси кількість рівнів фіксується в досить широкому інтервалі ширин ями. Цей факт є дуже корисним для технологів, оскільки незначне варіювання ширини ями при виготов-

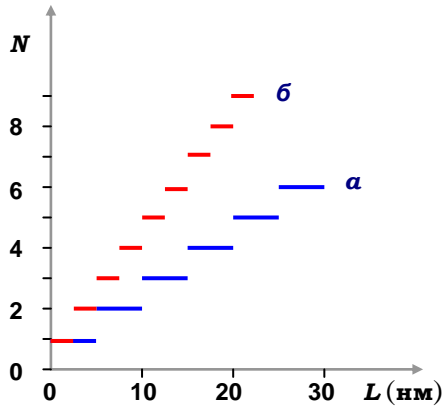


Рис. 16.9. Число енергетичних станів в ямі із параметрами: а – $V_b = 224$ меВ, $m^* = 0.067 m$, б – $V_b = 150$ меВ, $m^* = 0.4 m$

ленні структур з одновимірним конфайнментом не спричинить змін кількості енергетичних рівнів у випадку, якщо ефективна маса носіїв є малою. Точне значення для енергій зв'язаних у ямі рівнів можна знайти для двох крайніх випадків. Перший відповідає дуже неглибокій ямі, коли існує тільки один рівень у ямі

$$\epsilon_1 \approx \frac{m^* L^2 V_b^2}{2\hbar^2}. \quad (16.40)$$

Інший випадок – коли яма є нескінченно глибокою

$$\epsilon_n = \frac{\hbar^2 \pi^2 n^2}{2m^* L^2}, \quad n = 1, 2, 3, \dots \quad (16.41)$$

В останньому випадку відстань між рівнями зростає за збільшення номера рівня n . Наприклад, для тонкої півки $L = 12,5$ нм з арсеніду галію, де $m^* = 0,067m$, перші два рівні відповідають енергіям $\epsilon_1 \approx 35$ меВ, $\epsilon_2 \approx 140$ меВ. Отже, для квантових ям завширшки ~ 10 нм характерні значення енергій рівнів зв'язаних на ямі становлять порядок сотень міліелектронвольт, що порівняні з енергіями Фермі вродженого двовимірного електронного газу.

Таким чином, на противагу руху електронів в об'ємному зразку напівпровідника, у системах з одновимірним конфайнментом (де рух електронів обмежений уздовж однієї вісі – OZ), енергія руху електронів буде квантованою та утворюватиме підзони. Така ситуація описується формулою

$$E_{n,k_{\square}} = \epsilon_n + \frac{\hbar^2}{2m^*} (k_x^2 + k_y^2), \quad (16.42)$$

де ϵ_n – квантована енергія, пов'язана з поперечним обмеженням руху електрона. Таким чином, електронні стани у цьому випадку описуються двома квантовими числами – дискретним номером підзони n і неперервним значенням хвильового числа k_{\square} , що відповідає вільному руху у площині ями. Графічно закон дисперсії електрона у квантовій ямі матиме вигляд вкладених один в один параболоїдів обертання (рис. 16.10). При цьому мінімальне значення енергії відповідає найнижчому енергетичному рівню локалізованого в ямі стану при $\mathbf{k}_{\square} = 0$. Отже навіть при $n = 1$ енергія електрона у квантовій ямі більша за енергію руху електрона в об'ємному зразку (у площині $k_z = 0$).

Систему з електронним конфайнментом можна отримати на межі розподілу при утворенні гетерограниці за методом так званого модуляційного легування. Ще раз пояснимо коротко ідею методу на прикладі гетерограниці AlGaAs/GaAs. Якщо при утворенні гетерограниці між AlGaAs, легованого кремнієм, і GaAs, слабо легованого акцепторною домішкою, області легування не зачіпають сусідній напівпровідник, то електрони, що перейшли із домішок до зони провідності AlGaAs, тепер перейдуть праворуч від межі розподілу та опиняться у "кишені" нижче рівня Фермі поблизу границі з боку арсеніду галію (рис. 16.2). Електрони будуть просторово розділені із домішками і сильно локалізованими у тонкому шарі поблизу гетерограниці, матимуть високу рухливість і характеризуватимуться майже двовимірним характером їхнього руху. Таким чином, ці електрони можуть розглядатися як квазідвовимірна система із всіма притаманними таким системам властивостями. Для неї можуть бути застосовані методи обчислення енергетичного спектру, при які йшлося вище. Зрозуміло, що тепер для побудови моделі необхідно використовувати інший, відмінний від прямокутного, потенціал. Як показує досвід, зручною моделлю потенціальної ями у цьому випадку служитиме трикутна потенціальна яма.

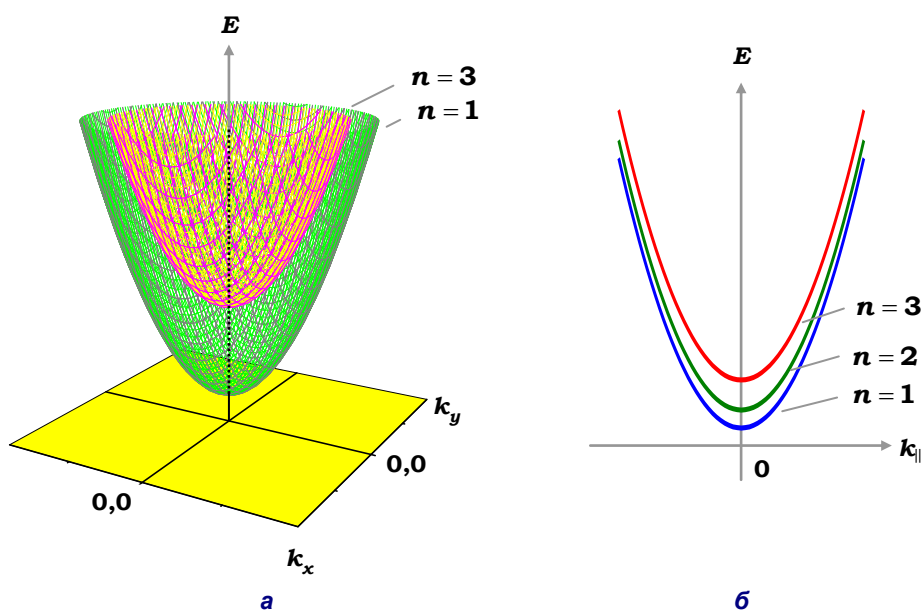


Рис. 16.10. Закон дисперсії електрона у квантовій ямі:
 а – тривимірний вигляд вкладених параболоїдів обертання;
 б – переріз дисперсійних поверхонь для фіксованого напрямку квазіімпульсу

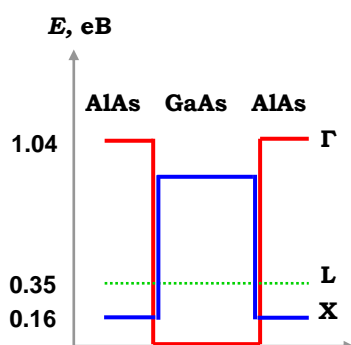


Рис. 16.11. Формування енергетичних бар'єрів у різних кристалографічних напрямках гетероструктури AlAs/GaAs/AlAs

Розглядаючи властивості квантоворозмірних гетероструктур, варто пам'ятати, що вихідні матеріали, на базі яких утворюються квантові ями, є кристалічними речовинами. Це означає, що вирощуючи ту чи іншу гетероструктуру, необхідно враховувати, що за тих самих геометричних параметрів тонких епітаксціальних плівок можна отримувати принципово різні системи щодо їхньої енергетичної структури. На рис. 16.11 подано просторові профілі енергетичної структури системи AlAs/GaAs/AlAs для різних груп носіїв. Видно, що квантова яма утворюється тільки для носіїв із Γ долини; для носіїв з X долини область GaAs являє собою бар'єр, а носії з L долини взагалі не відчувають наявності бар'єрних структур.

16.4.2. Квантові ефекти у неперервному спектрі. Поперечний транспорт

У попередньому підрозділі ми вивчали ефекти конфайнменту електронів в умовах, коли енергії не перевищували глибину потенційної ями. Розглянемо тепер поведінку квантової частинки, якщо її енергія перевищує висоту потенціального бар'єру й вона характеризується неперервним енергетичним спектром. Як і раніше, дослідимо прямокутну квантову яму. Аналізуючи рівняння Шредингера при $\epsilon > V_b$, можна зробити висновок, що для областей, далеких від квантової ями, розв'язком рівняння будуть дві плоскі хвилі, що розповсюджуються в протилежних напрямках. Перша хвиля відповідає розповсюдженню електрона зліва направо, тобто при $z \rightarrow -\infty$ маємо падаючу на потенціал та відбиту хвилю, а при $z \rightarrow \infty$ – тільки хвилю, що пройшла потенціальний бар'єр. Інша ситуація відповідає протилежному напрямку руху електрона. Обидва рухи відповідають одному й тому самому значенню енергії. Оскільки рухи є еквівалентними, розглянемо тільки один, за якого електрон розповсюджується зліва направо. Введемо хвильовий вектор в області бар'єра як

$$k_b = \sqrt{\frac{2m}{\hbar^2} (\epsilon - V_b)}. \quad (16.43)$$

При цьому зауважимо, що для значень енергії, більших за висоту бар'єру ($\epsilon > V_b$), хвильовий вектор є реальним. Хвильова функція матиме вигляд

$$\chi(z) = \begin{cases} e^{ik_b(z+L/2)} + re^{-ik_b(z+L/2)}, & z \leq -L/2 \\ ae^{ik_w z} + be^{-ik_w z}, & -L/2 \leq z \leq L/2 \\ te^{ik_b(z-L/2)}, & z \geq L/2 \end{cases}. \quad (16.44)$$

Константи a , b , r , t знайдемо з умови неперервності хвильової функції та її похідних на бар'єрі. Ця процедура майже така сама, як у випадку $\epsilon < V_b$. Але існує суттєва розбіжність. У цьому випадку система із чотирьох алгебраїчних рівнянь є неоднорідною, і для існування нетривіальних розв'язків не потребує виконання умови занулення детермінанту, що спричинило б обмеження припустимих значень енергії. У результаті дістанемо, що константи, які характеризують хвильову функцію руху із неперервною енергією, можна виразити у термінах амплітуди падаючої хвилі, яку ми вибрали рівною 1. Прирівнюючи значення хвильових функцій та їхніх похідних при $z = -L/2$ і $z = L/2$, отримаємо систему рівнянь

$$\begin{cases} 1 + r = ae^{-ik_w L/2} + be^{ik_w L/2} , \\ t = ae^{ik_w L/2} + be^{-ik_w L/2} , \\ \frac{\kappa_b}{\kappa_w} (1 - r) = ae^{-ik_w L/2} - be^{ik_w L/2} , \\ \frac{\kappa_b}{\kappa_w} t = ae^{ik_w L/2} - be^{-ik_w L/2} . \end{cases} \quad (16.45)$$

Додаючи до першого й другого рівнянь (16.45) третє та четверте, і віднімаючи від першого й другого третє й четверте, отримаємо систему, з якої виключені невідомі a та b

$$\begin{cases} \frac{1}{t} + \frac{r}{t} \frac{\kappa_w - \kappa_b}{\kappa_w + \kappa_b} = \cos k_w L - i \sin k_w L , \\ \frac{1}{t} + \frac{r}{t} \frac{\kappa_w + \kappa_b}{\kappa_w - \kappa_b} = \cos k_w L + i \sin k_w L . \end{cases} \quad (16.46)$$

Віднімемо від другого рівняння цієї системи перше та отримаємо зв'язок між коефіцієнтами r і t

$$r(\epsilon) = \frac{i}{2} \left(\frac{\kappa_w}{\kappa_b} - \frac{\kappa_b}{\kappa_w} \right) t(\epsilon) \sin k_w L . \quad (16.47)$$

Підставимо співвідношення (16.47) до (16.46), дістанемо

$$\begin{cases} \frac{1}{t} + \frac{i}{2} \frac{(\kappa_w - \kappa_b)^2}{\kappa_w \kappa_b} \sin k_w L = \cos k_w L - i \sin k_w L , \\ \frac{1}{t} + \frac{i}{2} \frac{(\kappa_w + \kappa_b)^2}{\kappa_w \kappa_b} \sin k_w L = \cos k_w L + i \sin k_w L . \end{cases} \quad (16.48)$$

Додамо до першого рівняння отриманої системи друге. У результаті знайдемо

$$t(\epsilon) = [\cos k_w L - (i/2)(\kappa_w / \kappa_b + \kappa_b / \kappa_w) \sin k_w L]^{-1} , \quad (16.49)$$

і разом із (16.47) отримуємо набір коефіцієнтів, що визначають хвильову функцію електрона за межами потенціальної ями.

Для обчислення коефіцієнтів відбиття та проходження використаємо формулою потоку частинок

$$\mathbf{j} = -\frac{i\hbar}{2m} (\psi^* \nabla \psi - \psi \cdot \nabla \psi^*). \quad (16.50)$$

Для обчислення щільності потоку частинок, що падають на область ями за руху зліва направо, використовуємо хвильову функцію

$$\psi_{in} = e^{ik_b(z+L/2)}. \quad (16.51)$$

Тоді

$$j_{in} = \frac{\hbar k_b}{m}. \quad (16.52)$$

Хвильова функція електронів, що відбиті потенціалом і рухаються із правої до лівої частини системи,

$$\psi_r = r e^{-ik_b(z+L/2)}. \quad (16.53)$$

Підставляючи цей вираз до (16.50), отримуємо щільність потоку частинок, відбитих областю неоднорідності потенціалу

$$j_r = \frac{\hbar}{m} k_b |r|^2. \quad (16.54)$$

Аналогічно, застосовуючи хвильову функцію частинок, що пройшли бар'єрну область,

$$\psi_{tr} = t e^{ik_b(z-L/2)}, \quad (16.55)$$

отримаємо щільність потоку частинок, що пройшли над бар'єрною областю

$$j_{tr} = \frac{\hbar}{m} k_b |t|^2. \quad (16.56)$$

Коефіцієнти проходження та відбиття, що визначені формулами

$$T(\epsilon) = \frac{j_{tr}}{j_{in}}, \quad R(\epsilon) = \frac{j_r}{j_{in}}, \quad (16.57)$$

таким чином можна обчислити згідно із наступними виразами

$$T(\epsilon) = |t|^2 = \left[1 + \frac{1}{4} \left(\frac{k_w}{k_b} - \frac{k_b}{k_w} \right)^2 \sin^2 k_w L \right]^{-1}, \quad (16.58)$$

$$R(\epsilon) = |r|^2 = 1 - T(\epsilon). \quad (16.59)$$

Ці формули ілюструють суто квантовий ефект при проходженні частинкою бар'єрної області з енергією, що перевищує висоту бар'єра. По-перше, ми бачимо, що при $\epsilon > V_b$ коефіцієнт відбиття R є ненульовим. Це типовий квантовий ефект відбиття частинки потенційною ямою, що зникає у класичному наближенні. По-друге, знаменник у

(16.58) осцилює зі зміною енергії, тобто коефіцієнт проходження частинкою бар'єрної області є осцилюючою функцією енергії. За енергій, що задовольняють умові

$$k_w L = \sqrt{\frac{2m\epsilon}{\hbar^2}} L = n\pi, \quad (16.60)$$

коефіцієнт проходження дорівнює одиниці. При цьому коефіцієнт відбиття занулюється (!). За малих значень різниці $\Delta = \epsilon - V_b$ маємо, що $T(\epsilon) \sim \Delta$, і коефіцієнт проходження стає зникаюче малим при $\Delta \rightarrow 0$. Інший цікавий квантовомеханічний ефект пов'язаний із поведінкою хвильової функції: ймовірність знаходження частинки в області квантової ями є осцилюючою функцією енергії. Така поведінка може визначати технологічно важливі процеси проникнення та покидання електронами квантової ями. На рис. 16.12 подано поведінку коефіцієнта проходження як функції енергії частинок (а – дірок, б – електронів) у квантовій ямі, що утворена структурою AlGaAs/GaAs із шаром GaAs товщиною 25 нм, що характеризується глибиною потенціальної ями 150 меВ для дірок і 224 меВ – для електронів. Видно, що згідно із (16.60) відстань між двома одиничними значеннями коефіцієнту проходження збільшується зі збільшенням ефективної маси носія, а також, що коефіцієнт проходження за зростання енергії частинки асимптотично наближується до одиниці. Цей результат свідчить, що частинка із дуже великою енергією, пролітаючи над областю потенціальної ями, не "помічає" її. Але навіть за дуже великих значень енергії частинки її коефіцієнт відбиття не тотожний нулем, що є проявом квантової поведінки частинки.

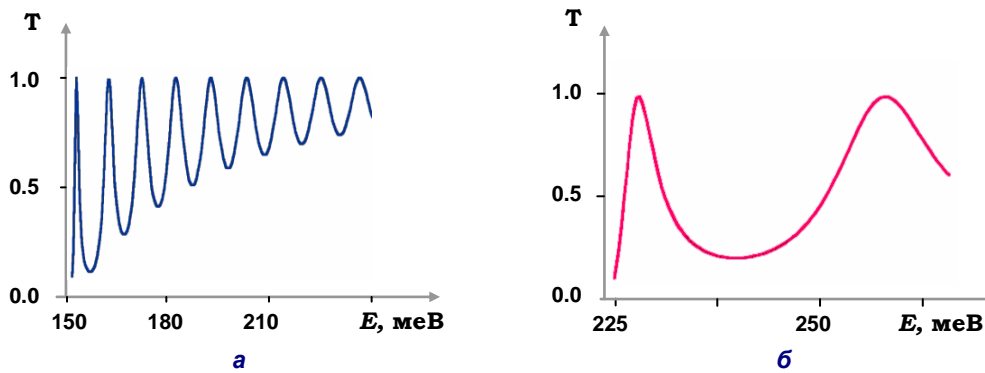


Рис. 16.12. Залежність від енергії коефіцієнта проходження у квадратній потенційній ямі шириною 25 нм у структурі AlGaAs/GaAs:

а – дірки: $V_b = 150$ меВ, $m^* = 0.48$ м;
 б – електрони: $V_b = 224$ меВ, $m^* = 0.67$ м

16.4.2. Щільність станів електронів у квантових ямах

Важливою характеристикою квантової системи є кількість станів, що приходиться на інтервал енергії $(E, E + \delta E)$,

$$\rho(E) = 2 \sum_n \delta(E - E_n). \quad (16.61)$$

Фактор 2 враховує, що кожен з n рівнів в ямі дворазово вироджений за спіном. Рух електронів у квантовій ямі є тривимірним. Цей факт віддзеркалює закон дисперсії (16.42), що пов'язує двовимірний рух у площині ями із рухом поперек ями: двовимірна підзона, що репрезентує рух у площині ями, може бути пов'язана із кожним із квантових зв'язаних станів в ямі. Щільністю станів у цьому випадку є

$$\rho(E) = 2 \sum_{k_x k_y} \sum_n \delta \left[E - \epsilon_n - \frac{\hbar^2 k_{\perp}^2}{2m^*} \right]. \quad (16.62)$$

Площу зразка приймаємо великою, тобто можемо використовувати циклічні умови Борна-Кармана для координат x та y . Знаходимо

$$k_x = m \frac{2\pi}{L_x}, \quad k_y = n \frac{2\pi}{L_y}. \quad (16.63)$$

З іншого боку, великі значення планарних розмірів ями дозволяють використати правило заміни суми на інтеграл

$$\sum_{k_x k_y} f(k_x, k_y) \rightarrow \frac{L_x L_y}{4\pi^2} \int dk_x dk_y f(k_x, k_y). \quad (16.64)$$

Тоді для щільності станів маємо

$$\rho(E) = \frac{S}{2\pi^2} \sum_n \int dk_x dk_y \delta \left[E - \epsilon_n - \frac{\hbar^2 (k_x^2 + k_y^2)}{2m^*} \right]. \quad (16.65)$$

Як обчислюються інтеграли типу (16.65)? Розглянемо інтеграл

$$j = \int dx dy \delta(\epsilon - ax^2 - ay^2). \quad (16.66)$$

По-перше, зазначимо, що при $\epsilon < 0$ і $a > 0$ інтеграл (16.66) тотожно дорівнює нулю. У випадку, коли $\epsilon > 0$ скористуємось відомими формулами

$$\delta(A^2 - X^2) = \frac{1}{2A} [\delta(X - A) + \delta(X + A)], \quad (16.67)$$

$$\text{та} \quad \delta(ax) = \frac{1}{a} \delta(x). \quad (16.68)$$

Тоді замість (16.66) маємо

$$\begin{aligned} j &= \int dy dx \frac{1}{a} \delta(\varepsilon/a - x^2 - y^2) = \\ &= \frac{1}{a} \int dy \frac{1}{2\sqrt{\varepsilon/a - y^2}} \int dx \left[\delta(x - \sqrt{\varepsilon/a - y^2}) + \delta(x + \sqrt{\varepsilon/a - y^2}) \right]. \end{aligned} \quad (16.69)$$

Оскільки аргумент дельта-функції у другому доданку підінтегральної функції у (16.69) може дорівнювати нулю лише за $x < 0$, а у першому доданку за $x > 0$, то обидва доданки при інтегруванні дають по одиниці, оскільки y змінюється у всьому діапазоні $[-\infty, +\infty]$. У результаті дістаємо

$$j = \frac{1}{a} \int dy \frac{1}{\sqrt{\varepsilon/a - y^2}}. \quad (16.70)$$

Для обчислення цього інтегралу необхідно визначити межі інтегрування. Виходитимемо з таких міркувань. Аргумент у дельта-функції у (16.69) може дорівнювати нулю лише за виконання умови

$$\varepsilon/a - y^2 \geq 0, \quad (16.71)$$

або $-\varepsilon/a \leq y \leq \varepsilon/a$. Тоді маємо

$$j = \frac{1}{a} \int_{-\sqrt{\varepsilon/a}}^{\sqrt{\varepsilon/a}} dy \frac{1}{\sqrt{\varepsilon/a - y^2}} = \frac{1}{a} \arcsin\left(\frac{y}{\sqrt{\varepsilon/a}}\right) \Big|_{-\sqrt{\varepsilon/a}}^{\sqrt{\varepsilon/a}} = \frac{\pi}{a}. \quad (16.72)$$

Пам'ятаючи, що у цій формулі $a = \hbar^2 / 2m^*$, отримаємо остаточно

$$\rho(E) = \frac{m^* S}{\hbar^2 \pi} \sum_n \vartheta(E - \varepsilon_n), \quad (16.73)$$

де $\vartheta(x)$ – функція Хевісайда

$$\vartheta(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases} \quad (16.74)$$

Таким чином, щільність станів носіїв у квантовій ямі є східцеподібною (рис. 16.13). Часто застосовують щільність станів на одиницю площі. При цьому залежність від геометричних розмірів системи у площині двовимірного шару зникає. Кожний доданок у (16.73) відповідає внеску від окремої підзони, тобто щільність станів двовимірного електронного газу являє собою східцеподібну залежність, яка із кожною сходинкою пов'язана з енергією ϵ_n . Порівнюючи щільність станів для тривимірного та двовимірного електронного газу, бачимо,

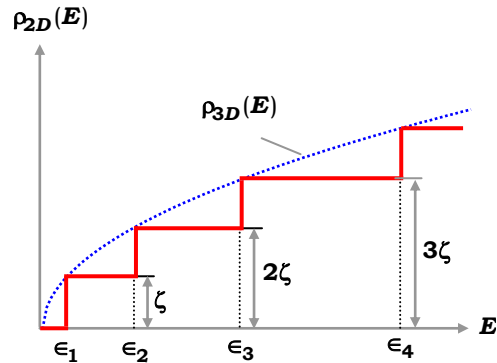


Рис. 16.13. Щільність станів двовимірного електронного газу у глибокій прямокутній потенціальній ямі. $\zeta = m^* / \pi \hbar^2$

що відмінність максимально виявляється для нижчих підзон. Зі зростанням номера підзони ця відмінність стає все меншою та майже зникає за великих значень n . Такі зміни у поведінці щільності станів двовимірного електронного газу є наслідками одновимірного конфайнменту або прямими наслідками низької розмірності системи. Ці зміни можуть спричинити істотні зміни у поведінці електронів. У реальних системах кожна сходинка функції (16.74) згладжується за рахунок розсіювання на дефектах. Цей ефект добре описує емпіричне правило, згідно з яким функція Хевісайда замінюється на

$$\vartheta(E - E_n) \rightarrow \frac{1}{\pi} \left\{ \frac{\pi}{2} + \text{Arctg} \left[\frac{(E - \epsilon_n)}{\Gamma_n} \right] \right\} \quad (16.75)$$

із параметром поширення Γ_n . Оскільки, як видно із (16.73), $\rho(E) = 0$ при $E < \epsilon_1$, і якщо для об'ємного матеріалу з енергетичною щілиною E_g поглинання світла можливе лише при $\hbar\omega \geq E_g$, то завдяки нерівності $\epsilon_1(L) \geq E_g$ для квантової ями край смуги поглинання буде зсунутий до синьої області. Цей зсув залежить від ширини ями та може варіюватись від нуля до кількох сотень міліелектронвольт.

■ 16.5. Поперечний транспорт через подвійну гетероструктуру

Розглянемо поперечний рух електронів у подвійній гетероструктурі GaAs-AlGaAs-GaAs, де два шари напівпровідника GaAs розділені шаром сполуки AlGaAs, яка характеризується широкою забороненою зоною так, що можна вважати, що електрон рухається у полі прямокутного потенціального бар'єру (рис. 16.14). Використовуючи метод Т-матриці (див. додаток I), обчислимо коефіцієнт проходження частинкою прямокутного бар'єру. Припустимо, що $E > V_b$. Оскільки потенціал ліворуч і праворуч бар'єра дорівнює нулю, то в обох цих областях хвильові вектори

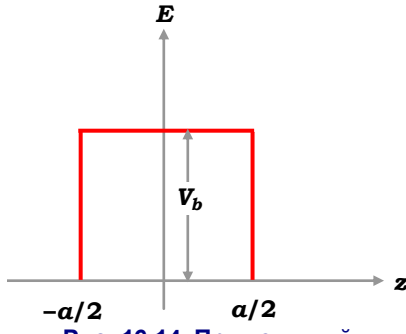


Рис. 16.14. Прямокутний потенціальний бар'єр

$$k = \sqrt{2m^*E/\hbar^2}. \quad (16.76)$$

Т-матрицю будемо у дві стадії. Спочатку запишемо матрицю переходу з області 1 ($z < -a/2$), що характеризується хвильовим вектором k , до області 2 ($-a/2 \leq z \leq a/2$), яка характеризується хвильовим вектором

$$k_b = \sqrt{2m^*(E - V_b)/\hbar^2}, \quad (16.77)$$

а далі, використовуючи головну властивість методу Т-матриці, матрицю переходу із області 1 до області 3 ($z > a/2$) як добуток матриць $\mathbf{T}^{(21)}$ і $\mathbf{T}^{(32)}$. Беручи до уваги, що границі бар'єрних областей лежать не на початку координат, використовуємо формули (I.11) і (I.18) із додатку I. Маємо

$$\begin{aligned} \mathbf{T}^{(31)} &= \begin{pmatrix} e^{-ika/2} & 0 \\ 0 & e^{ika/2} \end{pmatrix} \mathbf{T}^{(32)} \begin{pmatrix} e^{ik_b a/2} & 0 \\ 0 & e^{-ik_b a/2} \end{pmatrix} \times \\ &\times \begin{pmatrix} e^{ik_b a/2} & 0 \\ 0 & e^{-ik_b a/2} \end{pmatrix} \mathbf{T}^{(21)} \begin{pmatrix} e^{-ika/2} & 0 \\ 0 & e^{ika/2} \end{pmatrix}. \end{aligned} \quad (16.78)$$

Зауважимо, що

$$\begin{pmatrix} e^{ik_b a/2} & 0 \\ 0 & e^{-ik_b a/2} \end{pmatrix} \begin{pmatrix} e^{ik_b a/2} & 0 \\ 0 & e^{-ik_b a/2} \end{pmatrix} = \begin{pmatrix} e^{ik_b a} & 0 \\ 0 & e^{-ik_b a} \end{pmatrix}, \quad (16.79)$$

$$\mathbf{T}^{(21)} = \left(\frac{1}{2k_b} \right) \begin{pmatrix} k_b + k & k_b - k \\ k_b - k & k_b + k \end{pmatrix}, \quad (16.80)$$

та

$$\mathbf{T}^{(32)} = \left(\frac{1}{2k} \right) \begin{pmatrix} k_b + k & -k_b + k \\ -k_b + k & k_b + k \end{pmatrix}. \quad (16.81)$$

Виконуючи операцію прямого матричного добутку, отримаємо, що

$$\mathbf{T}^{(31)} = \frac{1}{2kk_b} \begin{pmatrix} (2kk_b \cos k_b a + i(k^2 + k_b^2) \sin k_b a) e^{-ika} & -i(k^2 - k_b^2) \sin k_b a \\ i(k^2 - k_b^2) \sin k_b a & (2kk_b \cos k_b a - i(k^2 + k_b^2) \sin k_b a) e^{ika} \end{pmatrix}$$

Неважко бачити, що для отриманої матриці $\det |\mathbf{T}| = 1$. Тоді, згідно із загальними принципами Т-матриці (див. додаток I), амплітуда хвильової функції частинки, що пройшла перешкоду, визначається величиною

$$t = \frac{\det |\mathbf{T}|}{T_{22}} = \frac{2kk_b e^{-ika}}{2kk_b \cos k_b a - i(k^2 + k_b^2) \sin k_b a}, \quad (16.82)$$

а коефіцієнт проходження (за потоком) – квадратом модуля t (з обох боків бар'єру хвильові числа однакові)

$$T = \left\{ 1 + \frac{V_b^2}{4E(E - V_b)} \sin^2 k_b a \right\}^{-1}, \quad E > V_b. \quad (16.83)$$

Формулу, аналогічну (16.83) легко отримати й для випадку $E < V_b$. Для цього у (16.83) необхідно замінити $k_b \rightarrow i\sqrt{2m^*(V_b - E)/\hbar^2} = i\tilde{k}_b$. Тоді

$$T = \left\{ 1 + \frac{V_b^2}{4E(V_b - E)} \sinh^2 \tilde{k}_b a \right\}^{-1}, \quad E < V_b, \quad (16.84)$$

У випадку, коли $k_b a$ є дуже великою, тобто або енергія бар'єру велика, або великою є його ширина, у (16.84) основну роль відіграватиме експонента гіперболічного синусу, що дасть

$$T \cong \frac{16E}{V_b} e^{-2k_b a} . \quad (16.85)$$

Таким чином, ми отримали загальновідому оцінку ймовірності тунелювання через будь-який бар'єр $\sim e^{-2k_b a}$. На рис. 16.15 подано залежність коефіцієнта проходження бар'єру висотою 0,3 еВ завширшки 10 нм у GaAs. Він якісно відрізняється від класичного випадку, де коефіцієнт проходження дорівнює 1 при $E > V_b$ і 0 при $E < V_b$. Звернемо увагу також на осциляції коефіцієнта, які особливо яскраво виявляються за енергій, що ненабагато перевищують V_b .

Таким чином, у планарних нанорозмірних системах слід очікувати специфічних ефектів у поперечному транспорті, що пов'язані з виявленням квантової природи руху носіїв у неоднорідних системах на відстанях порядку довжини хвилі де Бройля структурах – квантових ямах і бар'єрних структурах.

Далі розглянемо цікаві ефекти, обумовлені квантовою інтерференцією, які можна спостерігати у подвійних бар'єрних структурах.

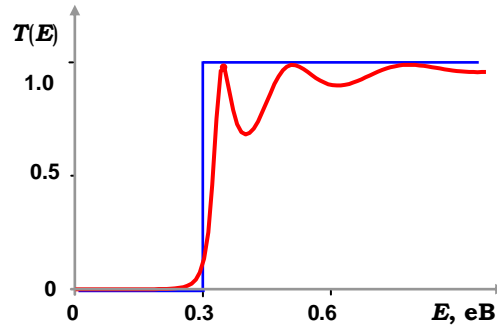


Рис. 16.15. Коефіцієнт проходження квантової частинки (червона крива) бар'єрної області (0.3 еВ) завширшки 10 нм у подвійній гетероструктурі GaAs-AlGaAs-GaAs.

Синім показано коефіцієнт проходження у класичному випадку

■ 16.6. Резонансне тунелювання

Розглянемо квантову яму між двох бар'єрів, що може, наприклад бути утворена к структурі із тонкого шару напівпровідника (GaAs), оточеного з обох боків товстими шарами AlGaAs, які відіграють роль бар'єрних структур. Згадаємо, що у випадку, коли обидві бар'єрні об-

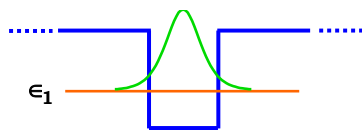


Рис. 16.16. Одиначна квантова яма

ласті є нескінченними у просторі (рис. 16.16), маємо одиничну квантову яму, в якій існують локалізовані рівні (перший із них і відповідну хвильову функцію подано на рисунку). Коли бар'єрні області починають стискатися і стають порівняними із розмірами ями, виникає істот-

но інша ситуація: замість фіксованого локалізованого рівня – вузька смуга невизначених станів, оскільки електрон може тунелювати через бар'єри і покинути яму. Однак якщо товщина бар'єрів є досить великою, електрон може лишатися в ямі досить довго. Такий стан електрона є резонансним або квазізв'язаним. Енергію такого стану неможна точно визначити, вона розмита у деяку зону $\Delta \epsilon$ шириною \hbar/τ , де τ – час життя електрона в ямі. Резонансні стани можна просто описати у термінах коефіцієнтів проходження. Загалом ймовірність проходження частинкою двох бар'єрів T грубо можна представити як добуток коефіцієнтів проходження кожного із бар'єрів окремо. Поблизу резонансу коефіцієнт проходження T значно зростає та досягає одиниці, якщо структура стає симетричною – маємо *резонансне тунелювання*. Цей ефект спостерігається експериментально у широкому діапазоні частот: від мікрохвильового до оптичного (широковідомий еталон Фабрі-Перо).

Для пояснення механізму резонансного тунелювання обчислимо коефіцієнт проходження частинки в одновимірній резонансній структурі (рис. 16.17). Нехай правий бар'єр характеризується амплітудами проходження та відбиття t_R і r_R , відповідно. Ці коефіцієнти залежать від хвильового вектора, отже – від параметрів ями та енергії частинки. У системі координат, початок якої збігається із серединою ями, матриця переходу за руху частинки із ями через бар'єр праворуч має вигляд

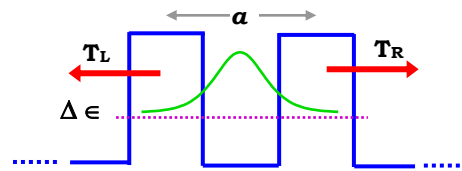


Рис. 16.17. Розмивання локалізованого рівня у вузьку зону у двобар'єрній структурі

$$\mathbf{T}_R^{(32)} = \begin{pmatrix} 1/t_R^* & -r_R^*/t_R^* \\ -r_R/t_R & 1/t_R \end{pmatrix}. \quad (16.86)$$

Оскільки ми розглядаємо електрон, що перебуває в ямі, та обчислюємо ймовірності для його переходу з області ями до області правого або лівого берегів приладу, то лівий бар'єр характеризуватимемо амплітудами переходів із ями через область лівого бар'єру. У такому разі для матриці переходу з лівого берега приладу до ямної області маємо використовувати інверсну матрицю переходу

$$\mathbf{T}_L^{(21)} = \begin{pmatrix} 1/t_L^* & r_L/t_L \\ r_L^*/t_L^* & 1/t_L \end{pmatrix}. \quad (16.87)$$

Тоді, маючи на увазі, що потенціальні бар'єри зсунуті на величину $\pm a/2$, запишемо, використовуючи матриці зсуву (див. додаток I), матрицю переходу із лівого берега системи на правий

$$\mathbf{T}^{(31)} = \begin{pmatrix} e^{-ika/2} & 0 \\ 0 & e^{ika/2} \end{pmatrix} \begin{pmatrix} 1/t_R^* & -r_R^*/t_R^* \\ -r_R/t_R & 1/t_R \end{pmatrix} \begin{pmatrix} e^{ika/2} & 0 \\ 0 & e^{-ika/2} \end{pmatrix} \\ \times \begin{pmatrix} e^{ika/2} & 0 \\ 0 & e^{-ika/2} \end{pmatrix} \begin{pmatrix} 1/t_L^* & r_L/t_L \\ r_L^*/t_L^* & 1/t_L \end{pmatrix} \begin{pmatrix} e^{-ika/2} & 0 \\ 0 & e^{ika/2} \end{pmatrix}$$

Прямий добуток цих матриць дає

$$\mathbf{T}^{(31)} = \begin{pmatrix} (1 - r_L^* r_R^* e^{-2ika})/t_L^* t_R^* & (r_L e^{ika} - r_R^* e^{-ika})/t_R^* t_L \\ (r_L^* e^{-ika} - r_R e^{ika})/t_L t_R^* & (1 - r_L r_R e^{2ika})/t_L t_R \end{pmatrix} \quad (16.88)$$

Амплітуда хвилі частинки, що пройшла бар'єрну область, дорівнює $1/T_{22}$ (див. додаток I), тобто

$$t = \frac{t_L t_R}{1 - r_L r_R e^{2ika}}. \quad (16.89)$$

Для подальшого аналізу запишемо комплексні амплітудні коефіцієнти відбиття у полярній формі, ввівши їхні фази ρ_R і ρ_L , тобто

$$t = \frac{t_L t_R}{1 - |r_L| |r_R| e^{2ika + i\rho_L + i\rho_R}} \quad (16.90)$$

Оскільки квадрати модуля амплітудних коефіцієнтів проходження та відбиття дають коефіцієнти проходження та відбиття за потоком, маємо із (16.90)

$$T = \left| \frac{t_L t_R}{1 - |r_L| |r_R| e^{2ika + i\rho_L + i\rho_R}} \right|^2 = \frac{T_L T_R}{1 + R_R R_L - |r_R| |r_L| (e^{i\phi} + e^{-i\phi})}, \quad (16.91)$$

де $\phi = 2ka + \rho_L + \rho_R$. Оскільки $\cos \phi = 1 - 2\sin^2 \phi/2$, то перепишемо цю формулу у вигляді

$$T = \frac{T_L T_R}{(1 - \sqrt{R_R} \sqrt{R_L})^2 + 4\sqrt{R_R} \sqrt{R_L} \sin^2 \phi/2}. \quad (16.92)$$

В отриманому виразі всі величини залежать від енергії, й аналізувати поведінку коефіцієнта проходження як функції енергії – досить непросте справа. Але якщо взяти до уваги, що поблизу резонансу найбільш швидко змінюється фаза хвилі між бар'єрами, можна припус-

тити, що коефіцієнт проходження досягає пікової величини, коли фаза задовольняє умові резонансу: $\phi = 2n\pi$. У цьому випадку

$$T = \frac{T_L T_R}{(1 - \sqrt{R_L R_R})^2}. \quad (16.93)$$

Використаємо той факт, що в реальних двобар'єрних системах коефіцієнти проходження є малими величинами. Маємо

$$(1 - \sqrt{R_L R_R})^2 = (1 - \sqrt{(1 - T_L)(1 - T_R)})^2 \approx \left(1 - \left(1 - \frac{1}{2}(T_L + T_R)\right)\right)^2 = \frac{1}{4}(T_L + T_R)^2$$

Тоді (16.93) можна переписати у вигляді

$$T = T_{ph} \approx \frac{4T_L T_R}{(T_L + T_R)^2}. \quad (16.94)$$

Зауважимо, що умова резонансного переходу через двобар'єрну область

$$2ka + \rho_L + \rho_R = 2n\pi \quad (16.95)$$

є умовою інтерференції хвилі в ямі. Дійсно, розглядаючи електрон як хвилю, що проходить через бар'єри, можна вважати, що він додає до своєї фази величину ka при проходженні ями в кожному напрямку із додаванням фази ρ_L і ρ_R за відбиття від бар'єрів. Якщо бар'єри є однаковими, величина коефіцієнту проходження може за резонансу досягати одиниці (див. (16.94) при $T_L = T_R$). Вважаючи індивідуальні коефіцієнти проходження через бар'єри малими, із (16.92) отримаємо

$$T \approx \frac{T_L T_R}{(T_L + T_R)^2 / 4 + 4 \sin^2 \phi / 2} = T_{ph} \left[1 + \frac{16}{(T_L + T_R)^2} \sin^2 \frac{\phi}{2} \right]^{-1}. \quad (16.96)$$

Аналізуючи цю формулу, зазначимо, що завдяки коефіцієнту перед синусом другий доданок у квадратних дужках може бути домінуючим, якщо синус не дуже малий. Покладемо, що типовою величиною для синуса є $1/2$. Тоді із (16.96) отримаємо, що $T \approx (T_L T_R) / 2$. Маємо тривіальний результат: коефіцієнт проходження через дві перешкоди є добутком коефіцієнтів переходу через кожен із перешкод. Але у випадку резонансу ситуація істотно змінюється. Нехай $\phi = 2n\pi + \delta\phi$. Розкладемо праву частину (16.96) за малим параметром $\delta\phi$

$$T \approx T_{ph} \left[1 + \frac{4(\delta\phi)^2}{(T_L + T_R)^2} \right]^{-1} = \frac{T_{ph}}{1 + \left(\frac{\delta\phi}{\phi_0/2} \right)^2}, \quad (16.97)$$

де $\phi_0 = T_L + T_R$. Видно, що при $\delta\phi = \pm\phi_0$ коефіцієнт проходження зменшується вдвічі порівняно із максимальним значенням T_{ph} , тобто ϕ_0 – напівширина Лоренцевого піку за резонансного тунелювання. Перейдемо до залежності від енергії і запишемо

$$\Gamma = \frac{dE}{dk} \frac{dk}{d\phi} \phi_0 = \frac{\hbar v}{2a} (T_L + T_R). \quad (16.98)$$

Вважаючи, що в основному фаза частинки змінюється при проходженні області ями, ми ввели швидкість електрона за його руху між бар'єрами. Тоді із (16.97) маємо

$$T(E) \approx T_{ph} \left[1 + \left(\frac{E - E_{ph}}{\Gamma/2} \right)^2 \right]^{-1}, \quad (16.99)$$

де E_{ph} – енергія, за якої відбувається резонансне тунелювання.

На рис. 16.18 подано поведінку коефіцієнта проходження двобар'єрної області за резонансного тунелювання для системи

AlGaAs-GaAs-AlGaAs

із двома однаковими бар'єрами, де $E_{ph} = 0.05$ еВ, $\Gamma = 0.01$ еВ. Видно, що поведінка коефіцієнта проходження має яскраво виражений резонансний характер. Резонансна залежність коефіцієнта проходження дозволяє сподіватись на можливість отримання системи з нелінійними вольт-амперними характеристиками. Такий ефект можна застосувати

при побудові електронних приладів, наприклад тунельного діода. Розглянемо гетероструктуру, що містить дві бар'єрні області, розділені шаром напівпровідника, який утворює квантову яму. Енергетичну діаграму розглянутої структури подана на рис. 16.19. У випадку незначного зміщення (16.19 а) резонансний рівень лежить вище за найвищий

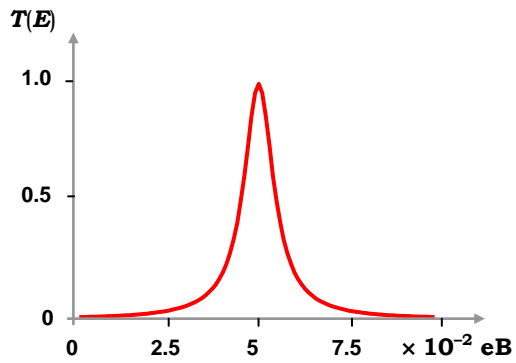


Рис. 16.18. Залежність коефіцієнту проходження частинкою двобар'єрної структури від енергії частинки. Резонансне тунелювання

рівень енергії заповнених станів, резонансне тунелювання відсутнє, а тунелювання через два бар'єри спричиняє дуже незначний струм (область *a* на рис. 16.19). За збільшення зміщення рівень усередині ями попадає до смуги енергій заповнених станів, і резонансне тунелювання

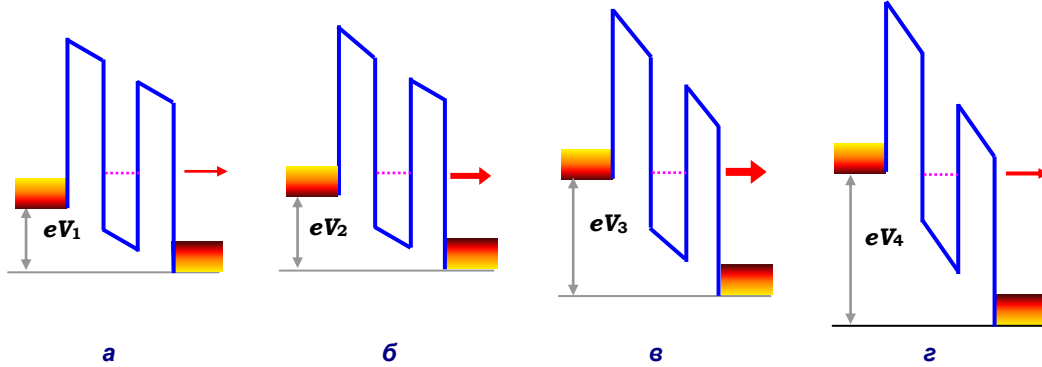


Рис. 16.19. Енергетичні діаграми, що пояснюють формування струму у резонансно-тунельному діоді

дає значний струм, що підвищується зі зростанням зміщення (область *б* на рисунку). Подальше зростання зміщення до величини, за якого рівень усередині ями розташовується нижче дна зони провідності лівого берега приладу, струм різко зменшується (*в*) і знову визначається тунелюванням через обидва бар'єри, але за досить великого потенціалу зміщення (рис. 16.19 *г*). У результаті дістаємо нелінійну вольт-амперну характеристику (рис. 16.20), що має одну (або кілька, якщо в області між бар'єрами формується кілька енергетичних "рівнів") *N*-подібну область. Праворуч від максимуму на кривій $I(V)$ спостерігається ділянка від'ємної диференціальної провідності. Завдяки цьому в електронних схемах тунельний діод може використовуватись не тільки для випрямлення струму, а й виконувати різноманітні функції. Оскільки вольт-амперна характеристика тунельного діода має *N*-подібний ви-

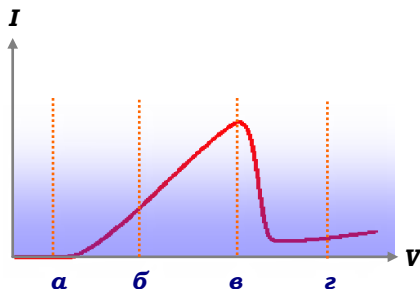


Рис. 16.20. Вольт-амперна характеристика тунельно-резонансного діода

гляд, то його, зокрема можна використовувати як генератор. Якщо ж до центральної області резонансного діоду підвести контакт, через який можна керувати положенням енергетичного рівня, то отримаємо новий прилад – тунельний резонансний транзистор. Резонансний тунельний діод являє собою перший приклад реального пристрою, що виконаний на основі структур із квантовими ямами

та бар'єрами, який базується на ефекті просторового квантування. Ідею такого діоду було вперше запропоновано Йогансеном в 1963 р. А в 1974 р. Єсакі та Чанг створили діючий прилад.

■ 16.7. Зв'язані квантові ями. Надґратки

Розглянемо електронні властивості структури, яка містить деяку кількість квантових ям, утворених тонкими шарами напівпровідника (напр., GaAs), що розділені шаром діелектрика або напівпровідника з великою забороненою зоною (напр., AlGaAs). Такі структури мають незвичайні властивості, що можуть бути цікавими не тільки з погляду фундаментальної науки, але й як систем, на основі яких будується велика кількість електронних та оптичних приладів. На рис. 16.21 схематично подано таку гетероструктуру, що містить по черезні шари GaAs-AlGaAs. Такі структури називають надґратками.

Розташування країв енергетичних зон при утворенні гетеропереходів обговорювалось в розд. 14, зокрема ми розглядали утворення гетерограниць різних типів. Згідно із типами гетерограниць утворюються й структури надґраток. Наприклад, якщо надґратка формується із напівпровідників, що утворюють гетероперехід першого типу, то виникає надґратка першого типу, де мінімум зони провідності E_{C1} і максимум валентної зо-

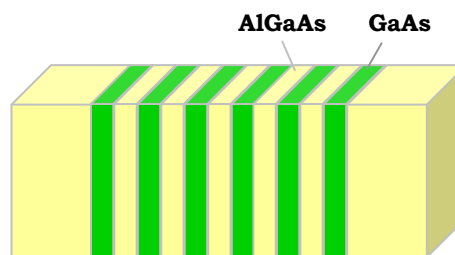


Рис. 16.21. Почережні квантові ями, розділені тонкими шарами діелектрика, які утворюють надґратку

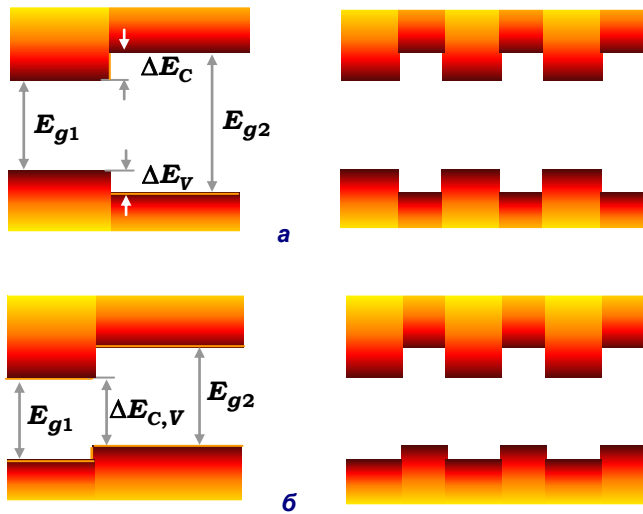
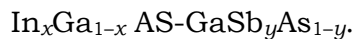


Рис. 16.22. Розташування країв енергетичних зон у гетеропереходах (ліворуч) і надґратках (праворуч) при утворенні надґраток: а – першого роду; б – другого роду

лентної зони ΔE_V напівпровідників, що утворюють надґратку. Широко розповсюдженою системою, що утворює надґратку першого типу, є $\text{Al}_x\text{Ga}_{1-x}\text{As}-\text{GaAs}$. У надґратках другого типу мінімум зони провідності одного напівпровідника розташований у забороненій зоні другого, а максимум валентної зони другого – у забороненій зоні першого. Таким чином, енергетична діаграма надґратки другого типу має східчастий хід зон (рис. 16.22 б). У таких надґратках модуляція країв зони провідності та валентної зони має один знак, що означає: просторово енергетичній ямі електрона відповідає бар'єр для дірки. Прикладом надґратки із такою енергетичною структурою є система



До надґраток другого типу часто зараховують і надґратки, утворені матеріалами із гетеропереходами третього типу, де мінімум зони провідності одного напівпровідника за енергією розташовується нижче, ніж максимум валентної зони другого напівпровідника. Такі надґратки часто називають надґратками другого типу із забороненими зонами, що не перекриваються. Прикладом такої надґратки є система $\text{InAs}-\text{GaSb}$.

Інший метод отримання надґраток – застосування технології поширювання легування. У легуваних надґратках використовується один напівпровідник, а періодичний потенціал утворюють чергуванням

ни E_{V1} одного напівпровідника за енергією розташовуються всередині забороненої зони іншого (рис. 16.22). У надґратках такого типу виникає періодична система квантових ям для носіїв у першому напівпровіднику, які відділені одна від одної бар'єрами, що утворюються в іншому напівпровіднику. При цьому для електронів глибина квантової ями визначається різницею ΔE_C між мінімумами зон провідності. А глибина квантових ям для дірок – різницею між стелями валентної зони ΔE_V напівпровідників, що утворюють надґратку.

шарів n - та p -типів одного й того самого напівпровідника. Іноді такі леговані шари відділяють один від одного тонким шаром діелектрика. Незалежно від способу утворення надґратки вона характеризується загальними рисами, головною з яких є наявність у структурі періодичного потенціалу, параметри якого можна задати наперед. Оскільки наявність потенціалу надґратки істотно змінює зонну енергетичну структуру висхідних напівпровідників, то, таким, чином, надґратка за сутністю являє собою новий тип напівпровідникового матеріалу, часто із характеристиками, аналогічних яким у природі не існує. Напівпровідникові надґратки характеризуються особливими властивостями, що зумовлюють їх широке практичне застосування. Наприклад, у них спостерігається:

- ♦ істотна зміна порівняно із висхідними матеріалами енергетичного спектру;
- ♦ наявність великої кількості енергетичних зон;
- ♦ значна анізотропія;
- ♦ пригнічення процесів електрон-діркової рекомбінації;
- ♦ концентрація носіїв може контролювано змінюватись у широкому діапазоні значень, і не визначається легуванням;
- ♦ широкі можливості перебудови зонної структури.

Для детальнішого розгляду властивостей таких структур спочатку розглянемо ефекти взаємодії електронних станів у двох квантових ямах. Коли ми розглядали одиничні квантові ями, то з'ясували, що електронна хвильова функція є відмінною від нуля в області бар'єра. Відповідно й електрон має ненульову ймовірність проникнення до підбар'єрної області. Цей ефект можна продемонструвати шляхом обчислення ймовірності перебування електрона у бар'єрному шарі гетероструктури, а саме

$$P_b = \int_{|z|>L/2} dz |\chi(z)|^2. \quad (16.100)$$

Цю ймовірність можна обчислити чисельно. Результат такого обчислення подано на рис. 16.23. Він є однією з основних відмінностей квантової поведінки частинки у тонкій плівці від класичного випадку. Таке тунелювання електрона до підбар'єрної області є основою для взаємодії двох і більше квантових ям і утворення так званих надґраток.

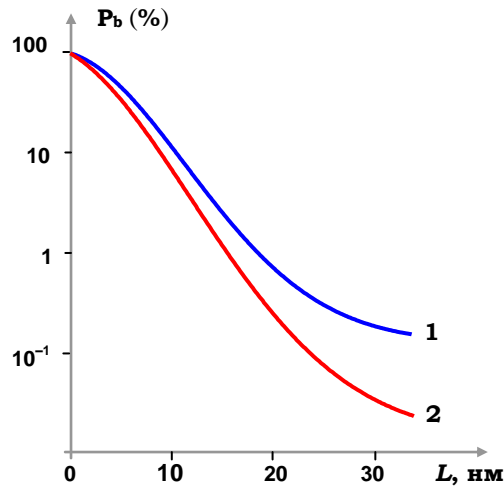


Рис. 16.23. Залежність ймовірності знаходження електрона у підбар'єрній області від товщини квантової ями L .

Електрон перебуває в основному енергетичному стані.

Криві: 1 – $V_b = 224$ меВ, $m^* = 0.067m_e$; 2 – $V_b = 150$ меВ, $m^* = 0.4m_e$

16.7.1. Дві зв'язані квантові ями

Розглянемо систему, що містить дві квантові ями, розділені бар'єром. Для спрощення обчислень вважатимемо що це – ідентичні прямокутні квантові ями з однаковою глибиною потенціалів V_b і шириною L . Нехай відстань між ямами дорівнює d , і ями розташовані симетрично щодо площини $z = 0$ (рис. 16.24). Розглянемо рух електрона у такій структурі

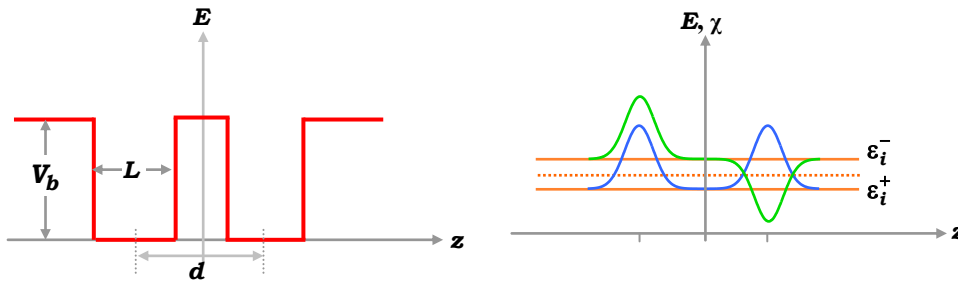


Рис. 16.24. Енергетичні рівні та хвильові функції для двоямної гетероструктури.

Штрихова лінія відповідає рівню енергії ϵ_i в одиничній квантовій ямі.

Для зв'язаних ям цей рівень розщеплюється на два – ϵ_i^\pm вздовж вісі OZ . Рівняння Шредингера у такій подвійній ямі вздовж вісі OZ має простий вигляд

$$\left[-\frac{\hbar^2}{2m^*} \frac{\partial^2}{\partial z^2} + V\left(z + \frac{d}{2}\right) + V\left(z - \frac{d}{2}\right) \right] \psi = \epsilon \chi(z). \quad (16.101)$$

Доданок $V(z + d/2)$ є потенціалом лівої ями із центром у точці $z = -d/2$. $V(z - d/2)$ відповідає правій ямі. При цьому потенціал

$$V(z) = \begin{cases} V_b > 0, & \text{всередині бар'єру} \\ 0, & \text{всередині ями} \end{cases} \quad (16.102)$$

Вважаючи, що електронні стани у такій системі є суперпозицією електронних станів одиничних незв'язаних ям, хвильову функцію електрона запишемо як суму за хвильовими функціями одноямних станів $\chi_\nu(z)$

$$\chi(z) = \sum_\nu \{A_\nu \chi_\nu(z - d/2) + B_\nu \chi_\nu(z + d/2)\}, \quad (16.103)$$

де ν нумерує всі зв'язані та незв'язані стани у випадку однієї ями. Якщо ж ями розділені досить великою відстанню, то тунелювання між станами у різних ямах є малоймовірним процесом. Тоді можна обмежитись розглядом тільки одного рівня, скажемо i -го, і хвильову функцію електрона можна записати у вигляді комбінації

$$\chi(z) = A_i \chi_i(z - d/2) + B_i \chi_i(z + d/2). \quad (16.104)$$

Ця функція відповідає енергії електрона, що наближено дорівнює енергії ϵ_i електрона в одиничній ямі. Оскільки відстань між ямами є досить великою, як перший крок можна вважати, що енергія електрона просто двічі вироджена енергія ϵ_i . Іншими словами, електрон у стані i може займати ліву або праву яму з ймовірністю одиниця. Для корекції енергії і знаходження коефіцієнтів A_i та B_i скористаємося процедурою, що базується на теорії збурень для вироджених станів, а також тим, що ϵ_i – енергія електрона одиничної ями.

$$\left[-\frac{\hbar^2}{2m^*} \frac{\partial^2}{\partial z^2} + V(z - d/2) \right] \chi_i(z - d/2) = \epsilon_i \chi_i(z - d/2), \quad (16.105)$$

та

$$\left[-\frac{\hbar^2}{2m^*} \frac{\partial^2}{\partial z^2} + V(z + d/2) \right] \chi_i(z + d/2) = \epsilon_i \chi_i(z + d/2). \quad (16.106)$$

Підставимо функцію (16.104) до (16.101) і скористаємось рівняннями (16.105) та (16.106). Отримаємо таке рівняння

$$[\varepsilon_i + V(z + d/2) - \epsilon] A_i \chi_i(z - d/2) + [\varepsilon_i + V(z - d/2) - \epsilon] B_i \chi_i(z + d/2) = 0. \quad (16.107)$$

Помножимо (16.107) спочатку на $\chi_i^*(z - d/2)$ а потім – на $\chi_i^*(z + d/2)$. Обидва тримані рівняння проінтегруємо за z .

$$\begin{aligned} & \left[\varepsilon_i + \int_{-\infty}^{\infty} dz \chi_i^*(z - d/2) V(z + d/2) \chi_i(z - d/2) - \epsilon \right] A_i + \\ & + \left[(\varepsilon_i - \epsilon) \int_{-\infty}^{\infty} dz \chi_i^*(z - d/2) \chi_i(z + d/2) + \right. \\ & \left. + \int_{-\infty}^{\infty} dz \chi_i^*(z - d/2) V(z - d/2) \chi_i(z + d/2) \right] B_i = 0 \end{aligned} \quad \text{та}$$

$$\begin{aligned} & \left[(\varepsilon_i - \epsilon) \int_{-\infty}^{\infty} dz \chi_i^*(z + d/2) \chi_i(z - d/2) + \int_{-\infty}^{\infty} dz \chi_i^*(z + d/2) V(z + d/2) \chi_i(z - d/2) \right] A_i + \\ & + \left[(\varepsilon_i - \epsilon) + \int_{-\infty}^{\infty} dz \chi_i^*(z + d/2) V(z - d/2) \chi_i(z + d/2) \right] B_i = 0 \end{aligned}$$

Введемо позначення

$$\int dz \chi_i^*(z - d/2) \chi_i(z + d/2) = \int dz \chi_i^*(z + d/2) \chi_i(z - d/2) = r_i \quad (16.108)$$

$$\int dz \chi_i^*(z \pm d/2) V(z \mp d/2) \chi_i(z \pm d/2) = s_i \quad (16.109)$$

$$\begin{aligned} \int dz \chi_i^*(z - d/2) V(z - d/2) \chi_i(z + d/2) = \\ = \int dz \chi_i^*(z + d/2) V(z + d/2) \chi_i(z - d/2) = t_i \end{aligned} \quad (16.110)$$

та отримаємо систему рівнянь для визначення невідомих коефіцієнтів A_i, B_i

$$[\varepsilon_i + s_i - \epsilon] A_i + [(\varepsilon_i - \epsilon) r_i + t_i] B_i = 0, \quad (16.111)$$

$$[(\varepsilon_i - \epsilon) r_i + t_i] A_i + [(\varepsilon_i - \epsilon) + s_i] B_i = 0. \quad (16.112)$$

Для існування нетривіальних розв'язків системи рівнянь (16.111)–(16.112) необхідно занулити детермінант цієї системи, тобто виконати умову

$$(\varepsilon_i + s_i - \epsilon)^2 - [(\varepsilon_i - \epsilon) r_i + t_i]^2 = 0. \quad (16.113)$$

Звідси маємо

$$(\varepsilon_i + s_i - \varepsilon) = \pm [(\varepsilon_i - \varepsilon)r_i + t_i], \quad (16.114)$$

тобто у системі двох зв'язаних квантових ям виникає зсув і розщеплення висхідного енергетичного рівня на два

$$\varepsilon_{(i)}^{\pm} = \varepsilon_i + \frac{s_i}{1 \pm r_i} \pm \frac{t_i}{1 \pm r_i}. \quad (16.115)$$

Таким чином, система енергетичних рівнів двоямної гетероструктури визначається набором інтегралів (16.108)–(16.110), причому перший із них (r_i) називається інтегралом перекривання та визначає міру перекривання хвильових функцій ям, другий (s_i) – називається інтегралом зсуву та характеризує зсув виродженого енергетичного рівня за рахунок взаємодії станів у різних ямах, третій (t_i) показує міру розщеплення зсунутої лінії. Якщо бар'єрна область (відстань між ямами) є досить великою, то хвильові функції ям перекриваються слабо. Це означає, що інтеграл перекривання є малим і ним можна знехтувати у виразі для енергії (16.115), тобто можна записати, що

$$\varepsilon_{(i)}^{\pm} \cong \varepsilon_i + s_i \pm t_i. \quad (16.116)$$

Оскільки значення інтегралів s_i та t_i зазвичай є негативними, то бачимо, що початковий енергетичний рівень зсувається у бік нижчих енергій, а потім розщеплюється (рис. 16.25). При цьому нижньому рівню ($\varepsilon_{(i)}^+$) відповідає симетрична хвильова функція ($A_i = B_i$), яку подано синім кольором. Стан

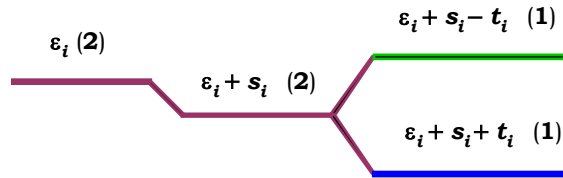


Рис. 16.25. Діаграма модифікації енергетичних рівнів двох квантових ям за їх взаємодії.

У дужках показано число станів із такою енергією

із верхнім рівнем енергії ($\varepsilon_{(i)}^-$) описано несиметричною хвильовою функцією ($A_i = -B_i$), яку схематично подано на рисунку зеленим кольором. Таким чином, тунельна взаємодія станів із двох квантових ям спричиняє істотну деформацію енергетичної структури висхідних електронних станів. Варто зазначити, що завдяки руху електронів у площині ями повну енергію електрона можна записати у вигляді

$$E_i^{(\pm)}(\vec{k}_{\square}) = \varepsilon_{(i)}^{\pm} + \frac{\hbar^2}{2m^*} k_{\square}^2, \quad (16.117)$$

тобто у структурі, що містить дві квантові ями, зберігається специфічний характер руху електронів: у площині ями вони поведуться як вільні частинки з ефективною масою m^* . Рух у поперечному напрямку

квантований і просторово обмежений областю ям. Із рис. 16.24 видно, що обидва (парний і непарний) стани є локалізованими на обох квантових ямах – і парна, і непарна хвильові функції мають екстремуми, що означає максимуми квадратів модулів $|\chi|^2$.

16.7.2. Енергія електронів у надґратках

Як згадувалось раніше, інтерес викликають гетеросистеми, що містять кілька квантових, у зв'язку із чим актуальним є аналіз електронних властивостей таких структур. Аналіз властивостей системи гетеропереходів, що утворюють багат шарову структуру, досить складний і в загальному випадку може проводитись тільки чисельними методами, причому методи можуть бути успішними тільки для небагатьох моделей. Але існують прості моделі, які дозволяють отримати результати, що в основному пояснюють властивості надґраток, тобто особливості формування енергетичного спектру електронів і шляхи впливу на зонну структуру з метою формування заздалегідь заданих її характеристик. Однією із таких моделей є модель Кроніґа-Пенні. Розглянемо систему, що складається із неперервного набору квантових ям завширшки b , які розділені бар'єром шириною a . Потенціальна енергія $V_b(z)$ є періодичною функцією z із періодом $d = a + b$ (рис. 16.26)

$$V_b(z) = \sum_n V_b(z - nd) \tag{16.118}$$

де
$$V_b(z) = \begin{cases} 0, & n(a+b) \leq z \leq n(a+b) + a \\ V_b, & n(a+b) - b \leq z \leq n(a+b) \end{cases} \tag{16.119}$$

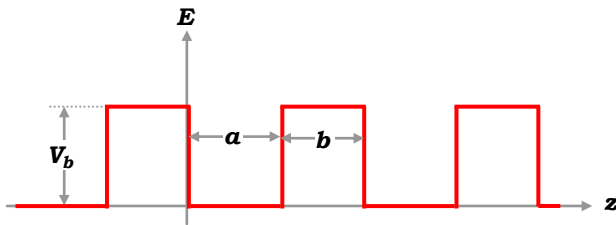


Рис. 16.26. Схематичне представлення потенціальної енергії надґратки із періодом $d = a + b$

Ця модель добре описує гетеротструктуру - надґратку першого типу (див. рис. 16.22). Детально модель Кроніґа-Пенні було розглянуто у розд. 3. У результаті було отримано, що закон дисперсії електронів у такій моделі визначається умовою

$$\cos(qd) = F(\epsilon), \tag{16.120}$$

де
$$\xi = k_w / k_b, \quad k_w = \sqrt{2m^* \epsilon / \hbar^2}, \quad k_b = \sqrt{2m^* (V_b - \epsilon) / \hbar^2},$$

$$F(\epsilon) = \cos(k_w a) \operatorname{ch}(k_b b) - \frac{1}{2}(-\xi + 1/\xi) \sin(k_w a) \operatorname{sh}(k_b b). \quad (16.121)$$

У цих формулах енергія руху електронів у площині ями

$$\epsilon = V_b + \frac{\hbar^2 k_b^2}{2m^*} = \frac{\hbar^2 k_w^2}{2m^*}. \quad (16.122)$$

У граничному випадку необмежено товстого бар'єру права частина (16.120) розходиться як $\exp(k_b b)$, якщо коефіцієнт перед цією експонентою не є нулем, тобто якщо не виконується умова

$$\cos(k_w a) - \frac{1}{2}(-\xi + 1/\xi) \sin(k_w a) = 0. \quad (16.123)$$

Це рівняння збігається із рівнянням зв'язаних станів ізольованої квантової ями. Дійсно, як було показано раніше ((16.34) і (16.35)), закон дисперсії в ізольованій квантовій ямі парного та непарного розв'язків має вигляд

парний розв'язок: $\operatorname{tg}(k_w a) = \frac{k_b}{k_w} \rightarrow \operatorname{ctg}(k_w a) = \frac{k_w}{k_b} \quad (16.124)$

непарний розв'язок: $\operatorname{ctg}(k_w a) = -\frac{k_b}{k_w}. \quad (16.125)$

Сума цих рівнянь дає $\operatorname{ctg}(k_w a) = \frac{1}{2} \left(\frac{k_w}{k_b} - \frac{k_b}{k_w} \right)$, що збігається із (16.123).

Таким чином, підзони надгратки за енергій електронів, менших від висоти потенціала бар'єра V_b з'являються як гібридизація зв'язаних станів ізольованих квантових ям завдяки тунельному зв'язку через бар'єр. Коли бар'єри є широкими (і/або високими), але кінцевими, добуток $k_b b$ є великим і підзони – вузькі. Ширина підзон пропорційна інтегралу передачі та швидко (експоненційно) зменшується зі зростанням b за фіксованого k_b . Функція $F(\epsilon)$ у правій частині (16.120) значно змінюється зі зміною ϵ , та енергетичний сегмент, де $|F(\epsilon)| < 1$ є вузьким. Для отримання наближених, але зручних виразів для дисперсійних співвідношень позначимо через $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ зв'язані стани ізольованої квантової ями. Розкладемо $F(\epsilon)$ поблизу деякої енергії ϵ_s , скажімо ϵ_j у першому порядку за $\epsilon - \epsilon_j$ та отримаємо дисперсійні співвідношення для j -ї підзони

$$\epsilon_j(q) = \epsilon_j + s_j + 2t_j \cos(qd), \quad (16.126)$$

де
$$s_j = \frac{-F(E_j)}{[F'(E)]_{E=E_j}}, \quad 2t_j = \frac{1}{[F'(E)]_{E=E_j}}. \quad (16.127)$$

Таким чином, (16.126) показує, як формуються мінізони спектру електронів надґратки, рух яких локалізований в ямі (рис. 16.27).

Розглянемо тепер випадок, коли $\epsilon > V_b$. Тоді рівняння, що визначає енергетичний спектр електронів надґратки, стає таким.

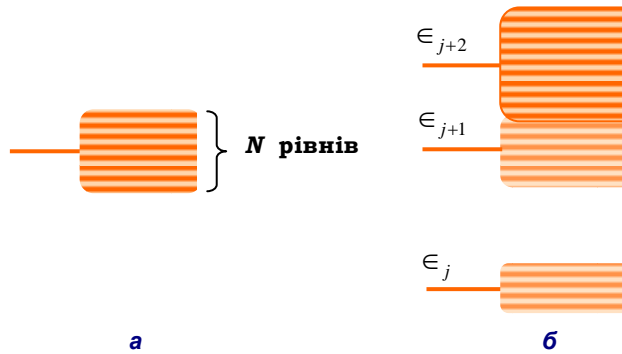


Рис. 16.27. Формування мінізон у надґратці:
а – розщеплення одного рівня в окремій квантовій ямі в енергетичну зону завдяки тунельному зв'язку;
б – мінізони, утворені із різних енергетичних рівнів окремої ями

$$\cos(qd) = \cos(k_w a) \cos(k_b b) - \frac{1}{2} \left(\xi + \frac{1}{\xi} \right) \sin(k_w a) \sin(k_b b). \quad (16.128)$$

Рівняння (16.128) є дисперсійним і пов'язує енергію із хвильовим вектором q . Для даної величини q існує нескінченна множина розв'язків рівняння (16.128). Для розрізнення розв'язків індекс підзони позначає енергію E та хвильову функцію χ . Одновимірний Блохівський стан залежить від двох орбітальних квантових чисел n та q і позначається як $|nq\rangle$. Дозволені підзони розділяються щілиною (можливе також перекриття міні-зон із формуванням ширшої мінізони). Для енергій, що відповідають цим енергетичним щілинам, не існує дозволених станів. Величина щілини між двома підзонами $|nq\rangle$ та $|n+1, q\rangle$ зменшується за зростання n . Для великих значень енергії ($\epsilon \gg V_b$), $k_b \sim k_w$ рівняння (16.128) зводиться до $k_w = q + 2j\pi/d$. Дійсно, при $k_b \approx k_w$, маємо, що (16.128) можна переписати

$$\cos qd \cong \cos k_w a \cos k_w b - \sin k_w a \sin k_w b = \cos k_w (a + b), \quad (16.129)$$

звідки $k_w d = qd + 2j\pi$. Або $k_w = q + (2j\pi)/d$. Потенціал надґратки тепер майже не відчувається електронами із високими енергіями, тобто довжина хвилі електрона $2\pi/k_w$ стає набагато меншою, ніж характерний період надґратки d , таким чином ефекти дифракції зникають. Зазначимо, що енергії, за якими $k_w a = p\pi$, завжди відповідають до-

зволеним станам надґратки. Дійсно, аналогічно до випадку, коли енергія електрона в ямі менша за висоту бар'єра, при $\epsilon > V_b$ можна розв'язок рівняння (16.128) записати у вигляді, подібному до (16.126), тобто

$$\epsilon_i(q) = \epsilon_i + s_i + 2t_i \cos qd. \quad (16.129)$$

Дисперсійне співвідношення (16.129) схематично подано на рис. 16.28 а.

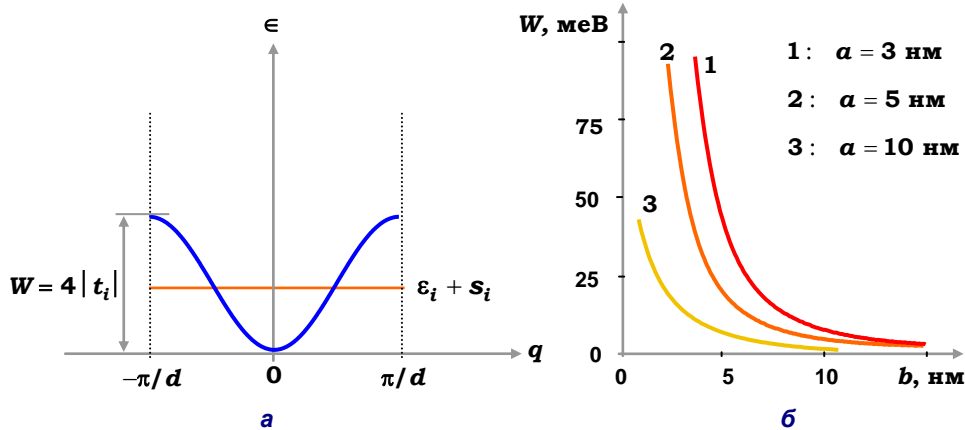


Рис. 16.28. Закон дисперсії мінізони у надґратці (а). Залежність ширини мінізони від товщини бар'єрної області у гетероструктурі, що утворює надґратку для різних величин ширини квантової ями (б)

Дисперсійна крива $\epsilon_i(q)$ виявляє невеликий зсув рівнів, який тим не менше вдвічі більшим за зсув для двоямної задачі, адже тепер у ями є ще дві сусідні, які за рахунок тунельної взаємодії формують цей зсув. Замість розщеплення рівня на два, що розділені енергією $2|t_i|$, як відбувалось у випадку двоямної задачі, тепер виникає неперервна зона $2t_i \cos qd$ завширшки $4|t_i|$, причому ширина зони не залежить від числа ям у надґратці. Останній доданок у (16.129) дає дисперсійний закон для i -ї енергетичної зони. Ці енергетичні зони часто називають мінізонами. Ширина мінізон обмежена, визначається інтегралом передачі, тобто значно залежить від параметрів надґратки (рис. 16.28 б). Повна енергія електрона у надґратці включає енергію поперечного руху (16.129) та енергію руху у площині шарів. Її можна записати як

$$E_i(\mathbf{k}, q) = \epsilon_i + s_i + \frac{\hbar^2 k^2}{2m^*} - 2|t_i| \cos qd, \quad (16.130)$$

де \mathbf{k} – **двовимірний** хвильовий вектор у площині напівпровідникового шару XOY. Ця формула показує, як істотно змінюється закон дисперсії електрона у штучній структурі – надґратці. Поблизу дна або стелі мі-

нізони енергії (16.130) можна спростити, розкладаючи косинус у ряд. Таким чином, поблизу $q = 0$ маємо

$$E_i(\mathbf{k}, q) = \epsilon_i + s_i - 2|t_i| + \frac{\hbar^2 q^2}{2M_i} + \frac{\hbar^2 k^2}{2m^*}, \quad (16.131)$$

де z -компонента тензора ефективної маси визначається інтегралом передачі відповідно до тунельного характеру поперечного руху

$$M_i = \frac{\hbar^2}{2d^2 |t_i|}. \quad (16.132)$$

Формула (16.131) демонструє, що рух електронів у надґратці є надзвичайно анізотропним. Анізотропія визначається шириною шарів і висотою потенціальних бар'єрів, що утворюють надґратку. Іншою характерною рисою енергії (16.130) є існування ділянок дисперсійної кривої, яким відповідає негативна ефективна маса, тобто друга похідна енергії за q змінює знак при $q = \pi/2d$. Поблизу стелі мінізони ефективна маса, що пов'язана із поперечним рухом, дорівнює M_i . Оцінимо ширину мінізони та величину ефективної маси. Припустимо, що ширина ями є 5 нм, а період надґратки – 10 нм. Знаходимо $\Delta \epsilon_i \approx 15 \text{ meV}$ і $M_i = 4.4 m_e$. Поперечна маса електрона у надґратці на порядок більша за типові значення ефективної маси електрона в об'ємному напівпровіднику. На рис. 16.29 подано утворення мінізон над ґратки з енергетичних

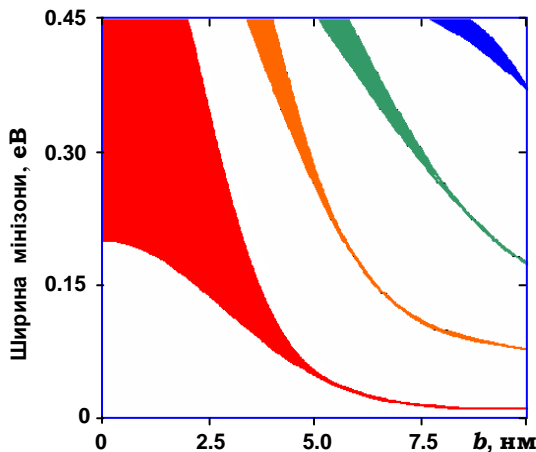


Рис. 16.29. Утворення міні зон над ґратки з енергетичних рівнів квантових ям

рівнів квантової ями залежно від ширини бар'єру b . Результати розрахунків отримано в моделі Кроніна-Пенні для надґратки, де ширини бар'єрів та ям збігаються. Глибина потенціальних ям надґратки дорівнювала 0,4 еВ. Звертає увагу, що для бар'єрів, ширших за 5 нм, мінізони, що утворюються із низько розташованих енергетичних рівнів, дуже вузькі. При цьому щілини між мінізонами існують навіть для випадків, коли мінізони розташовані вище потенціальних бар'єрів.

16.7.3. Щільність станів надґратки

Оскільки в площині ями носії можуть рухатись вільно, введемо такі позначення для стану носія у надґратці: n – індекс підзони, q – хвильовий вектор надґратки, ($|q| < \pi/d$), $k_{\perp} = (k_x, k_y)$ – хвильовий вектор носія у площині шару, σ_z – орієнтація електронного спіну ($\sigma_z = \pm 1/2$). Власні енергії

$$E(n, q, k_{\perp}, \sigma_z) = \frac{\hbar^2 k^2}{2m^*} + \epsilon_n(q) \quad (16.133)$$

не залежать від спінового індексу та є сумою енергій, що відповідають паралельному $\epsilon_n(q)$ та перпендикулярному $\hbar^2 k^2 / (2m^*)$ руху до вісі z . Якщо N_d – товщина надґратки, а $S = L_x L_y$ її площа, то щільність станів надґратки запишемо у стандартному вигляді

$$N(E) = \sum_{n, q, k_{\perp}, \sigma_z} \delta \left(E - \epsilon_n - s_n + 2 |t_n| \cos(qd) - \frac{\hbar^2 k^2}{2m^*} \right). \quad (16.134)$$

Проводячи підсумовування за \mathbf{k} та σ так само, як при обчисленні щільності станів одиначної квантової ями, маємо

$$N(E) = \frac{Sm^*}{\pi \hbar^2} \sum_{n, q} \theta(E - \epsilon_n - s_n + 2 |t_n| \cos qd). \quad (16.135)$$

Від суми за q перейдемо до інтегрування за q згідно із правилом $\sum_q \dots \rightarrow (d/\pi) \int dq$

$$N(E) = \frac{Sdm^*}{\pi^2 \hbar^2} \sum_n \int_{-q_E}^{q_E} dq \cdot \theta(E - \epsilon_n - s_n + 2 |t_n| \cos qd). \quad (16.136)$$

Граничні значення хвильових векторів, що визначають межі інтегрування, визначено з умови $E - \epsilon_n - s_n + 2 |t_n| \cos qd > 0$, тобто

$$q_E = \frac{1}{d} \arccos \frac{\epsilon_n + s_n - E}{2 |t_n|} \quad (16.137)$$

при

$$\epsilon_n + s_n - 2 |t_n| < E < \epsilon_n + s_n + 2 |t_n|. \quad (16.138)$$

За умови виконання нерівності $\epsilon_n + s_n - 2 |t_n| > E$ тета-функція у (16.135) тотожно дорівнює нулю. Якщо виконується нерівність

$E > \epsilon_n + s_n + 2|t_n|$, то тета-функція дорівнює одиниці та (16.135) дає $N(E) = (Sm^* N) / (\pi \hbar^2)$. Коли виконується нерівність (16.138), тета-функція у (16.136) дорівнює одиниці та інтегрування за q дає

$$\int_{-q_E}^{q_E} dq \cdot \theta(E - \epsilon_n - s_n + 2|t_n| \cos qd) = \arccos \frac{\epsilon_n + s_n - E}{2|t_n|}.$$

Таким чином, маємо для щільності станів

$$N(E) = \frac{L_x L_y m^*}{\pi \hbar^2} N \begin{cases} 0, & E < \epsilon_n^{(-)} \\ (1/\pi) \arccos((\epsilon_n + s_n - E)/(2|t_n|)), & \epsilon_n^{(-)} < E < \epsilon_n^{(+)} \\ 1, & E > \epsilon_n^{(+)} \end{cases}, \quad (16.139)$$

де $\epsilon_n^{(+)} = \epsilon_n + s_n + 2|t_n|$ і $\epsilon_n^{(-)} = \epsilon_n + s_n - 2|t_n|$ верхня та нижня границі n -ї мінізони, відповідно. Отже величина плато щільності станів дорівнює щільності зв'язаних станів в одиночній ямі, що збільшена в N разів. Із (16.139) видно, що коли енергія попадає до зонної щілини дисперсійних співвідношень надґратки, щільність станів квантується величиною $N_0 = Nm^* S / (\pi \hbar^2)$. На рис. 16.30 подано приведену щільність станів надґратки для значень енергії поблизу n -ї мінізони.

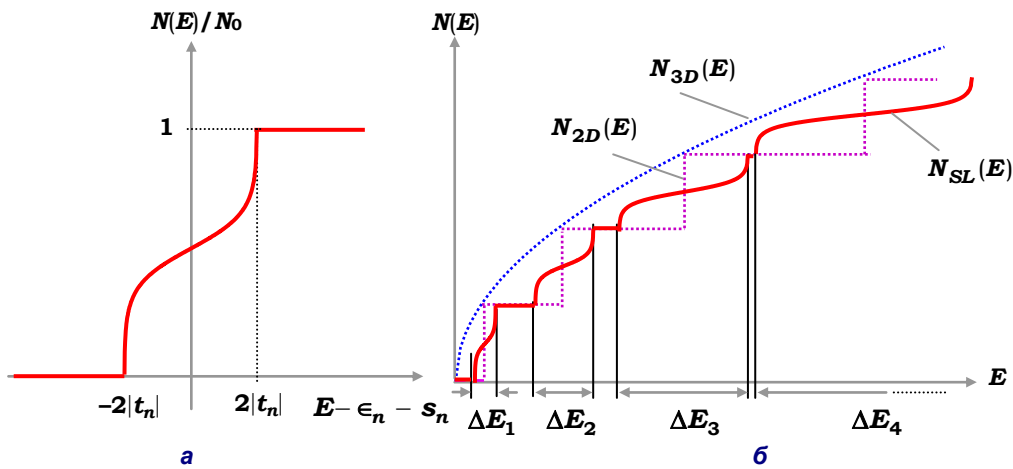


Рис. 16.30. Щільність станів надґратки для n -ї мінізони (а) та перших чотирьох мінізон (б). Штриховими лініями позначено щільність станів об'ємного зразка (синій) та квантової ями (бузковий)

■ 16.8. Квантові нитки (голки)

У попередньому розділі ми з'ясували, що перехід від тривимірного електронного газу до двовимірного спричиняє квантування руху електронів у напрямку, нормальному до площини плівки, де формується двовимірний електронний газ. У результаті електрон характеризується двома ступенями свободи. При цьому створення двовимірного газу відбувалося завдяки наявності одновимірного потенціалу $V(z)$, який формувався плоскою гетероструктурою. Для переходу від двовимірного електронного газу до одновимірного необхідно обмежити рух електронів ще в одному вимірі, тобто тепер обмеження руху електронів відбуватиметься у двох вимірах, і потенціал конфайнменту буде двовимірним $V(y,z)$. Тепер тільки вздовж напрямку x може відбуватись вільний рух електрона. Існує кілька сучасних технологій формування таких структур. Однією із найпростіших є безпосереднє "вирізання" вузької смужки з гетерограниці за допомогою літографії. При цьому для отримання квазіодновимірних електронних структур завширшки порядку десятків нанометрів, тобто структур, де ефект просторового квантування буде помітним, виготовляти смужки шириною такого самого порядку не обов'язково – немає необхідності використовувати літографічну техніку надвисокої роздільної здатності. На бокових гранях витравленої смужки аналогічно вільній поверхні напівпровідника утворюються поверхневі стани, що зазвичай створюють шар збіднення (рис. 16.31 б). Цей шар викликає додаткове звуження провідного каналу. У результаті виникає можливість спостерігати ефекти просторового квантування у смужках досить великої ширини, $\sim 0,1$ мкм. Можна також витравити квазіодновимірну структуру із тришарової структури, що утворює квантову яму (рис. 16.31 а). Іншою можливою структурою із двовимірним конфайнментом є так звана структура із розщепленим затвором, яку можна просто утворити шляхом витравлювання вузької смужки в металічному плоскому контакті структури Шотткі (рис. 16.31 в). Тут обмеження поперечного руху електронів виникає завдяки дії електричного поля, що виштовхує електрони з під

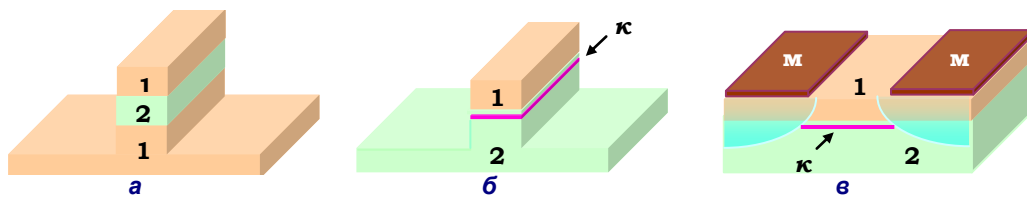


Рис. 16.31. Напівпровідникові гетероструктури із квантовими нитками, що утворюються за допомогою субмікронної літографії шляхом витравлювання вузької смужки із планарної гетероструктури: **а – квантової ями; б – вузького електронного каналу поблизу гетеро границі;**

в – витравлювання вузької щілини у затворі Шотткі.

1 – напівпровідник із широкою забороненою зоною (AlGaAs);

2 – напівпровідник із вузькою забороненою зоною (GaAs); м – металічний затвор,

к – електронний канал. Затемнення – області збіднення електронами

контактної області у середину області під протравленою смужкою. Прикладом квазіодновимірних структур є молекулярні конгломерати, що утворюються за агрегації органічних молекул. Подібний приклад ми розглядали у розд. 3.

16.8.1. Хвильові функції та енергетичні підзони електронів у квантових нитках

Розглянемо загальні риси поведінки електронів у гетероструктурі з обмежуючим потенціалом типу $V(y, z)$. Згідно із методом розділення змінних, розв'язок рівняння Шредингера можна представити у вигляді

$$\psi(x, y, z) = e^{ik_x x} \chi(y, z) \quad (16.140)$$

так, що рівняння Шредингера для поперечної хвильової функції $\chi(y, z)$ виглядатиме як

$$\left[-\frac{\hbar^2}{2m^*} \left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V(y, z) \right] \chi(y, z) = \epsilon \chi(y, z), \quad (16.141)$$

де
$$\epsilon = E - \frac{\hbar^2 k_x^2}{2m^*}. \quad (16.142)$$

Нехай $\chi_j(y, z)$ – деякий розв'язок (16.141), якому відповідає дискретний рівень енергії ϵ_j . Тоді повна енергія електрона у квантовій нитці

$$E = \epsilon_j + \frac{\hbar^2 k_x^2}{2m^*}. \quad (16.143)$$

Таким чином ми бачимо, що рух електрона обмежений у площині YOZ поблизу мінімуму V_0 потенціалу $V(y, z)$, при цьому електрон вільно розповсюджується в напрямку вісі OZ. Розглянемо найпростіший випадок потенціалу $V(y, z)$ як нескінченно глибокої прямокутної ями

$$V(y, z) = \begin{cases} 0, & 0 \leq y \leq L_y, \quad 0 \leq z \leq L_z, \\ \infty, & y \leq 0, \quad z \leq 0, \quad y \geq L_y, \quad z \geq L_z, \end{cases} \quad (16.144)$$

де L_y та L_z – поперечні розміри нитки. У цьому випадку хвильову функцію можна записати як добуток функцій, що окремо залежать від y та z

$$\chi(y, z) = \chi_{n1}(y) \chi_{n2}(z). \quad (16.145)$$

Для кожного із напрямків уздовж вісі ОУ та ОZ розв'язки одновимірних рівнянь Шредингера мають вигляд

$$\chi_{n_1}(y) = \sqrt{\frac{2}{L_y}} \sin \frac{\pi y n_1}{L_y}, \quad \chi_{n_2}(z) = \sqrt{\frac{2}{L_z}} \sin \frac{\pi y n_2}{L_z}, \quad n_1, n_2 = 1, 2, 3, \dots \quad (16.146)$$

При цьому енергія квантової частинки запишеться як

$$\epsilon_{n_1 n_2} = \frac{\pi^2 \hbar^2}{2m^*} \left(\frac{n_1^2}{L_y^2} + \frac{n_2^2}{L_z^2} \right). \quad (16.147)$$

Деякі інші методи застосовуються для розгляду квантової нитки, що утворена із двовимірного електронного газу за допомогою додаткового обмеження двовимірного руху електронів під дією зовнішнього поля. Така ситуація може реалізуватись, якщо на двовимірний електронний газ діятиме електричне поле, одновісні механічні напруження тощо. Цей випадок відомий як split-gate-техніка (див. рис. 16.29 в). У цьому випадку обмежувальний зовнішній потенціал можна наближено вважати параболічним, тобто

$$V(y, z) = V_1(y) + V_2(z), \quad V_1(y) = \frac{1}{2} \left(\frac{\partial^2 V_1}{\partial y^2} \right)_{y=0} y^2. \quad (16.148)$$

Якщо потенціал $V_2(z)$ забезпечує високі бар'єри при $|z| = L_z/2$, то розв'язок задачі такий самий, як й у випадку одновимірного квантового осцилятора. Тоді хвильову функцію можна записати через функцію Ерміта у вигляді

$$\chi_{n_1 n_2}(y, z) = C \cdot e^{-\alpha^2 y^2} H_{n_1}(\alpha y) \sin \frac{n_2 \pi z}{L_z}. \quad (16.149)$$

Енергія при цьому у напрямку ОZ квантуватиметься як для нескінченно глибокої прямокутної ями, а у напрямку ОУ – як для квантового осцилятора

$$\epsilon_{n_1 n_2} = \hbar \omega (n_1 + 1/2) + \frac{\pi^2 \hbar^2 n_2^2}{2m^* L_z^2}. \quad (16.150)$$

У формулах (16.149) та (16.150) ми використали позначення

$$\omega = \sqrt{\frac{2}{m^*}} \left(\frac{\partial^2 V_1}{\partial y^2} \right)_{y=0}, \quad \alpha = \frac{m^* \omega}{\hbar} \quad (16.151)$$

і поліном Ерміта

$$H_n(y) = (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2}. \quad (16.152)$$

Таким чином, можна стверджувати, що енергетичні рівні у квантових нитках значно залежать від форми обмежуючого потенціалу. Але можна також стверджувати, що загальною рисою поведінки квантових ниток є той факт, що додаткове просторове обмеження руху електронів (щодо їхнього руху у квантовій ямі) спричиняє зростання енергії найнижчого рівня та виникнення подвійного квантування енергетичних рівнів. Тепер два квантових числа n_1 і n_2 характеризуватимуть кожен енергетичний рівень при розгляді поперечного руху.

16.8.2. Щільність станів квазіодновимірного електронного газу

Як і в попередньому випадку двовимірного електронного газу запишемо щільність станів у вигляді

$$N(E) = 2 \sum_{n_1 n_2} \sum_{k_x} \delta \left(E - \epsilon_{n_1 n_2} - \frac{\hbar^2 k_x^2}{2m^*} \right), \quad (16.153)$$

де знову множник 2 відповідає двом різним значенням спіну для електронів, що характеризуються однією енергією. Вважаючи квантову нитку досить довгою, перейдемо від суми за хвильовим числом до інтегрування

$$N(E) = \sum_{n_1 n_2} \frac{L_x}{\pi} \int_{-\infty}^{\infty} dk_x \delta \left(E - \epsilon_{n_1 n_2} - \frac{\hbar^2 k_x^2}{2m^*} \right). \quad (16.154)$$

Тепер знов застосуємо властивості дельта-функції

$$\delta(A^2 - x^2) = [\delta(A - x) + \delta(A + x)] / 2A, \quad \text{та} \quad \delta(ax) = \delta(x) / a.$$

Із врахуванням цих формул здійснюємо перетворення з інтегралом у (16.154)

$$j = \int_{-\infty}^{\infty} dk_x \frac{1}{2\sqrt{E - \epsilon_{n_1 n_2}}} \left[\delta(\sqrt{E - \epsilon_{n_1 n_2}} - ax) + \delta(\sqrt{E - \epsilon_{n_1 n_2}} + ax) \right], \quad (16.155)$$

де $a = \sqrt{\hbar^2 / 2m^*}$. Інтегрування із δ -функцією дасть $2/a$. Таким чином маємо

$$j = \sqrt{\frac{2m^*}{\hbar^2}} \frac{1}{\sqrt{E - \epsilon_{n_1 n_2}}} \vartheta(E - \epsilon_{n_1 n_2}). \quad (16.156)$$

Тета-функція тут з'явилася завдяки тому, що при $E - \epsilon_{n_1 n_2} < 0$ інтеграл тотожно є нулем. Підставляючи (16.156) до (16.154), дістаємо

$$N(E) = \frac{L_x}{\pi} \sqrt{\frac{2m^*}{\hbar^2}} \sum_{n_1 n_2} \frac{1}{\sqrt{E - \epsilon_{n_1 n_2}}} \vartheta(E - \epsilon_{n_1 n_2}). \quad (16.157)$$

Поведінку щільності станів квазіодновимірного електронного газу подано на рис. 16.32. Видно, що ця поведінка істотно відрізняється від $N(E)$ для двовимірного газу. Замість східцеподібної вона стає розривною, де у межах однієї підзони зростання енергії спричиняє швидкий спад щільності станів. Перехід до нової підзони із подальшим зростанням енергії супроводжується появою нескінченних стрибків у $N(E)$. Така специфічна поведінка щільності станів електронів у квантовій нитці спричиняє появу

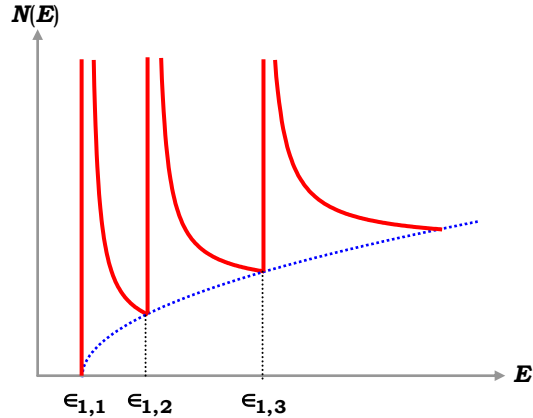


Рис. 16.32. Щільність електронних станів квантової нитки

цілої низки оптичних ефектів, які можна використати у практичних застосуваннях. Оптичні властивості низьковимірних систем буде розглянуті далі, а тепер ми обговоримо специфічні ефекти, пов'язані із провідністю квантових ниток. Найбільш цікавим випадком є провідність коротких (довжиною L , меншою за довжину вільного пробігу у матеріалі) квантових ниток – такі короткі квазіодновимірні структури часто називають квантовими голками. При цьому електрон рухається від одного контакту до іншого без зіткнень. Такий режим струму називають балістичним. Електронний транспорт у таких квазіодновимірних балістичних структурах зручно розглядати у більш загальному випадку квантового приладу, який ми і розглянемо далі.

16.8.3. Квантовий електронний транспорт. Формула Ландауера

Розглянемо деякий прилад, характерні лінійні розміри якого набагато менші всіх класичних характерних величин, розмірності довжини (напр., довжини екранування, довжини вільного пробігу), а також довжини хвилі де Бройля. На рис 16.33 подано схему та енергетичну ді-

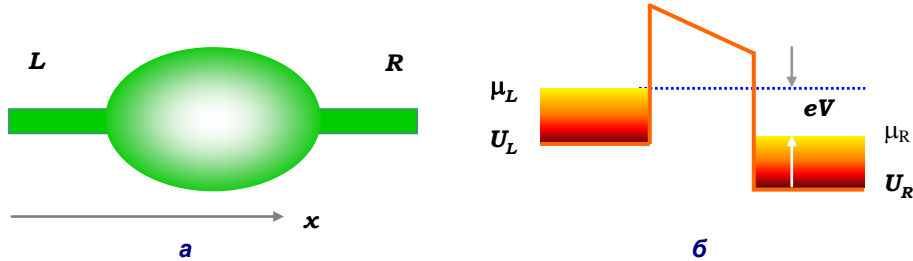


Рис. 16.33. Квантовий прилад (а) та його енергетична діаграма (б)

аграму такого приладу. Тут L та R – лівий і правий береги приладу, відповідно, V – потенціал зміщення. Вважатимемо, що прилад працює за досить низьких температур, тобто напівпровідник вироджений. Нехай на прилад діє електричне поле, спрямоване вздовж вісі Ox . У результаті лівий берег приладу знаходиться за енергії, піднятій на $eV/2$ щодо рівня Фермі, а правий – опущений на енергію $eV/2$. Оскільки через нелокальність – хвильові функції електронів з лівого берега приладу розповсюджуються на правий берег і навпаки – ми маємо характеризувати систему як ціле. Характеризуватимемо її функцію розподілу носіїв. При цьому лівий берег приладу характеризуватиметься функцією розподілу $F_F(E + eV/2 - E_F)$, а правий – функцією розподілу $F_F(E - eV/2 - E_F)$. Для визначеності вважатимемо, що квантовий рух відбувається тільки вздовж вісі Ox . Таким чином можна припустити, що рух носіїв відбувається у потенціалі

$$U(x, y, z) = U_1(x) + U_2(y, z) \tag{16.158}$$

так, що хвильова функція системи як функція координат може бути факторизованою

$$\Psi(x, y, z) = \Psi_{||}(x)\Psi_{\perp}(y, z). \tag{16.159}$$

У площині, перпендикулярній вісі Ox , стани системи характеризуватимемо хвильовими функціями $\Psi_T = \Psi_{m,n}(y, z)$. Оскільки електронний транспорт відбувається вздовж вісі Ox , розглядатимемо поведінку тільки повздовжньої частини хвильової функції. Зауважимо, що носії з правого та лівого берегів приладу взаємодіють. Це означає, що стан носіїв на лівому та правому берегах приладу характеризуватиметься комбінацією хвильових функцій лівого та правого берегів приладу

(вони відрізняються значеннями хвильового вектора). Запишемо граничні вирази для хвильових функцій на інтерфейсі між берегами та приладом, тобто у перерізах L і R . Для хвильової функції електронів, що рухаються з лівого берега, маємо

$$\Psi_{\square}^{(L)}(x) = \begin{cases} e^{ik_L(x-x_L)} + r_L e^{-ik_L(x-x_L)}, & x \sim x_L \\ t_R e^{ik_R(x-x_r)}, & x \sim x_R \end{cases}, \quad (16.160)$$

де r_L – коефіцієнт відбиття носіїв від бар'єру, а коефіцієнт t_R описує проходження носіїв через бар'єр приладу за руху частинки зліва направо. Аналогічно для руху електронів із правого берега граничні умови для хвильової функції запишемо як

$$\Psi_{\square}^{(R)}(x) = \begin{cases} e^{-ik_R(x-x_R)} + r_R e^{ik_R(x-x_R)}, & x \sim x_R \\ t_L e^{-ik_L(x-x_l)}, & x \sim x_L \end{cases}. \quad (16.161)$$

Для коефіцієнтів t_L , r_L , t_R і r_R можна записати кілька фундаментальних співвідношень, що мають загальний характер і не залежать від специфіки квантового приладу. Наприклад, вимога неперервності хвильових функцій (16.160) та (16.161) дає співвідношення

$$k_L(1 - |r_L|^2) = k_R |t_R|^2, \quad (16.162)$$

$$k_R(1 - |r_R|^2) = k_L |t_L|^2. \quad (16.163)$$

Запишемо тепер потік частинок через правий і лівий перерізи контактів приладу. Для цього використаємо формулу, що пов'язує потік квантових частинок із хвильовою функцією Ψ

$$\mathbf{j} = -\frac{i\hbar}{2m} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*). \quad (16.164)$$

Підставимо сюди хвильову функцію (16.160). Отримаємо потік частинок, що втікають у прилад з лівого берега

$$i_{in} = \frac{\hbar k_L}{m} = v_L. \quad (16.165)$$

Потік електронів, що пройшли на правий берег приладу,

$$i_{out} = \frac{\hbar k_R}{m} |t_R|^2 = v_R |t_R|^2. \quad (16.166)$$

Коефіцієнт пропускання за руху електронів з лівого на правий берег приладу визначається як відношення потоку частинок, що пройшли

бар'єрну область, до загального потоку частинок, що рухаються зліва направо

$$T_{L-R} = \frac{i_{out}}{i_{in}} = \frac{k_R}{k_L} |t_R|^2. \quad (16.167)$$

Аналогічно, коефіцієнт пропускання за руху електронів з правого на лівий берег приладу

$$T_{R-L} = \frac{k_L}{k_R} |t_L|^2. \quad (16.168)$$

Згідно із принципом детальної рівноваги, потоки електронів рівноважної системи, що перебуває у енергетичному стані E , мають бути рівними

$$T_{L-R} = T_{R-L} = T(E). \quad (16.169)$$

Тоді, записуючи струм зліва направо, як (зауважимо, що числом електронів, яке приходить на одиницю довжини, є $2F_F / L_C$; коефіцієнт 2 відбиває факт подвійного виродження по спіну)

$$J_L = \frac{2e}{L_C} \sum_{m,n} \sum_{k>0} vT(E)F_F(E(k,m,n) + \frac{1}{2}eV - E_F), \quad (16.170)$$

і аналогічно струм із правої частини приладу до лівої

$$J_R = -\frac{2e}{L_C} \sum_{m,n} \sum_{k>0} vT(E)F_F(E(k,m,n) - \frac{1}{2}eV - E_F), \quad (16.171)$$

отримаємо сумарний струм

$$J = J_L + J_R. \quad (16.172)$$

У формулах (16.170) та (16.171) через L_C позначено ефективну довжину приладу, а через n та m – номери підзон (завдяки руху електронів у площині YOZ енергетичні рівні розмиваються у підзони!). Підставимо до (16.172) вирази для струмів із (16.171) і (16.170) та перейдемо від підсумовування за хвильовими векторами до інтегрування за стандартними правилами²,

$$\sum_k (...) \rightarrow L_C \int \frac{dk}{2\pi} (...) \rightarrow L_C \int \frac{dE}{2\pi\hbar v} (...). \quad (16.173)$$

Отримаємо повний струм через прилад

² При виведенні формул враховано, що $k = \hbar p = \hbar mv$ і використано перетворення:

$$dk = (dk/dE)dE, \quad (dE/dk)^{-1} dE = (d/dk(k^2/2m)) dE = \frac{1}{\hbar v} dE.$$

$$J = e \sum_{m,n} \int \frac{dE}{2\pi\hbar} \cdot T(E) \left[F_F(E(k, m, n) + \frac{1}{2}eV - E_F) - F_F(E(k, m, n) - \frac{1}{2}eV - E_F) \right] \quad (16.174)$$

За низьких температур функція розподілу Фермі майже не відрізняється від ϑ -функції. Таким чином, у випадку незначних зміщень V можна записати

$$F_F(E + \frac{1}{2}eV - E_F) - F_F(E - \frac{1}{2}eV - E_F) \approx eV\delta(E - E_F). \quad (16.175)$$

Виконуючи інтегрування у (16.174), отримаємо

$$I = e^2 \int \frac{dE}{2\pi\hbar} T(E) V \delta(E - E_F) = \frac{e^2}{h} VT(E_F), \quad (16.176)$$

звідки знайдемо струм через прилад

$$J = \frac{e^2}{h} 2V \sum_{m,n} T(E_F). \quad (16.177)$$

Для квантових приладів із перенесенням заряду зручною характеристикою є кондактанс – квантовий аналог класичної провідності. Кондактанс G визначається як коефіцієнт пропорційності між струмом і різницею потенціалів між берегами квантового приладу

$$J = G \cdot V. \quad (16.178)$$

Тоді із (16.177) для кондактансу квантового приладу маємо *формулу Ландауера*

$$G = \frac{2e^2}{h} \sum_{m,n} T(E_F, m, n), \quad (16.179)$$

із якої випливає нетривіальний висновок про яскраву відмінність між квантовими та класичними приладами: навіть у випадку, коли квантовий прилад є ідеальним (тобто розсіювання носіїв відсутнє, це означає, що коефіцієнт пропускання приладу дорівнює одиниці), прилад характеризується кінцевою величиною кондактансу, тобто має опір. При цьому величина опору залежатиме від кількості електронних підзон N , з яких електрони беруть участь у квантовому транспорті

$$G = G_0 N. \quad (16.180)$$

Величина
$$G_0 = \frac{e^2}{h} = 39.6 \text{ [мкСим]} \quad (16.181)$$

називається квантом кондактансу (обернена величина $(G_0)^{-1}$ відповідає опору 25,2 кОм), а

$$N = \sum_{m,n} T(E_F, m, n), \quad T(E_F, m, n) = 1. \quad (16.182)$$

Розглянемо тепер, як згідно із (16.174) відбуватиметься транспорт у макроскопічно великому у поперечних напрямках пристрої. Попере-чний рух електронів відбувається зі поперечними квантовими числами, відповідними хвильовими векторами – $n = k_y$ і $m = k_z$. Тоді енергія поперечного руху

$$E_{n,m} = \frac{\hbar^2(k_y^2 + k_z^2)}{2m}. \quad (16.183)$$

Функцію розподілу електронів

$$F_F(E) = 2 \sum_{n,m} \frac{1}{1 + e^{\frac{E + E_{n,m}}{k_B T}}} \quad (16.184)$$

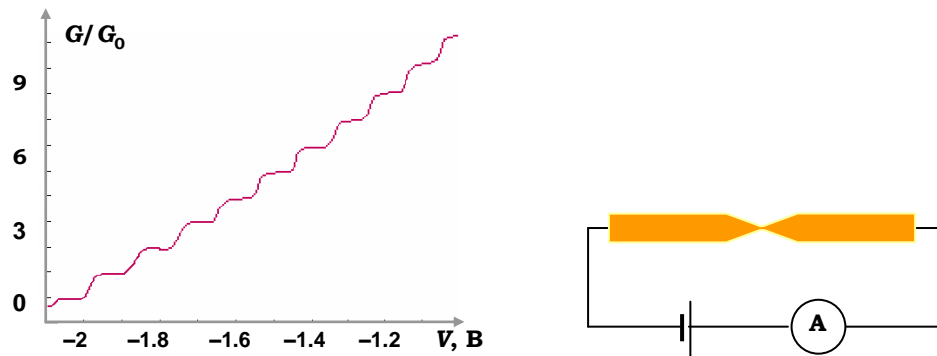
можна легко обчислити

$$F_F(E) = S \frac{mk_B T}{\pi \hbar^2} \ln \left(1 + e^{\frac{E}{k_B T}} \right), \quad (16.185)$$

де S – площа поперечного перерізу приладу. Тоді щільність струму у пристрої $j = J/S$

$$j = e \frac{mk_B T}{\pi \hbar^2} \int \frac{dE_{\square}}{2\pi \hbar} T(E_{\square}) \ln \left(\frac{1 + e^{\frac{E_F - E_{\square} + e\Phi/2}{k_B T}}}{1 + e^{\frac{E_F - E_{\square} - e\Phi/2}{k_B T}}} \right). \quad (16.186)$$

Ця формула має велике практичне значення, оскільки дозволяє обчислювати у наноприладах вольт-амперні характеристики та їх залежність від концентрації електронів, температури тощо. Експериментальні вимірювання квантування провідності у коротких квантових нитках зазвичай проводяться на структурах, що являють собою не нитку, а так званий точковий контакт – вузьку перемичку, яка з'єднує дві ділянки двовимірного електронного газу. Із формального погляду таку структуру можна уявити як квантову нитку із довжиною, порівнюваною з її шириною. Квантування провідності має спостерігатись і в таких структурах, оскільки в кінцевих формулах (16.180-16.182) не міститься жодних конкретних параметрів щодо розмірів і форми нитки. Експериментальна крива залежності провідності від напруги на затворі, що визначає концентрацію носіїв, має вигляд східцеподібної кривої, висота кожної сходинки дорівнює одному кванту кондактанса (рис. 16.34).



16.34. Залежність провідності квантової нитки від напруги на точковому контакті, що визначає концентрацію носіїв (а). Схема експерименту (б)

■ 16.9. Квантовий ефект Холла

Як ми бачили, класичний ефект Холла полягає у виникненні електричного поля, перпендикулярного струму, що протікає за поперечно направленою магнітним полем \mathbf{B} . Це поле Холла є перпендикулярним і струму, і магнітному полю: при накладенні магнітного поля перпендикулярно електричному струму у напівпровіднику з'являється ненульовий недиагональний компонент тензора провідності, тому для недиагонального компоненту тензора опору можна написати

$$\rho_{xy} = \frac{B}{en}. \quad (16.187)$$

У 1980 р. фон Клітцингом винайдено, що за низьких температур в інверсному шарі носіїв заряду у напівпровідниках, що вміщені у магнітне поле, перпендикулярне площині квазидвовимірного електронного газу, недиагональний компонент тензора опору змінюється дис-

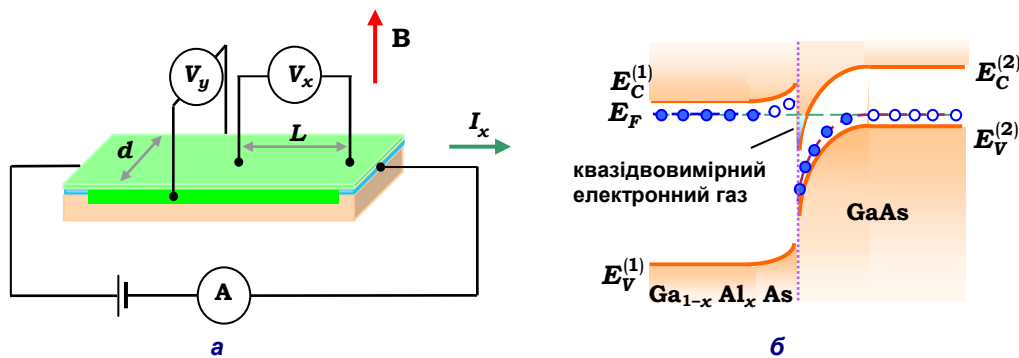


Рис. 16.35. Схема установки для вимірювання квантового ефекту Холла.

Канал квазидвовимірних електронів подано синім кольором.

Утворення каналу квазидвовимірних електронів на межі розподілу модуляційно-легованої гетероструктури. Синіми кульками позначено домішкові рівні, заповнені електронами, білими – дірками. Рівень Фермі збігається із рівнем донорної домішки в AlGaAs

кретно $\rho_{xy} = \text{const}/l$ із цілими значеннями параметра l . Це явище отримало назву квантового ефекту Холла. Пізніше, у 1982 р., незалежно групами дослідників керованими Тсуї, Штермером і Госсардом винайдено, що в МДН-структурах і на гетеропереходах AlGaAs-GaAs можна спостерігати ефект, коли параметр l набуває дробових значень: $l = 1/3, 2/3, 4/3, 5/3, 7/3, 8/3, 1/5, 2/5, 3/5, 4/5, 6/5, 2/7$ та $5/2$.

Розглянемо детальніше квантовий ефект Холла. На рис. 16.35 подано схему вимірювання квантового ефекту Холла. Вимірювались по-вздовжний

$$\rho_{xx} = \frac{d}{L} \cdot \frac{V_x}{I_x} \quad (16.188)$$

і поперечний (недіагональний компонент) опір, що визначається як відношення виникаючої поперечної напруги до струму, що протікає вздовж структури (див. рис.16.35)

$$\rho_{xy} = \frac{V_y}{I_x}. \quad (16.189)$$

Результати вимірювання подані на рис. 16.36. Видно немонотонну поведінку зміни опору за збільшення магнітного поля. Холловський опір (ρ_{xy}) має вигляд східцеподібної структури із висотою сходинок, що визначається опором

$$R_H = h/(e^2 l), \text{ де } l = 1, 2, 3, \dots$$

Константа h/e^2 наближено дорівнює 25 кОм (відповідає оберненому кванту кондактанса). Зауважимо, що відкриттю квантового ефекту Холла передувало відкриття іншого цікавого ефекту – зникнення електричного опору двовимірного металу в сильному магнітному полі (див. нижній графік рис. 16.36). Для з'ясування витоків цього ефекту згадаємо, що електрон у магнітному полі рухається вздовж кругової орбіти із радіусом $R_B = \sqrt{\hbar/(eB)}$. Цей результат ми отримували при розгляді руху електрона в квантуючому магнітному полі та виникнення рівнів Ландау

$$E_l = \hbar\omega_C(l + 1/2), \quad (16.190)$$

де $\omega_C = eB/m$ – частота обертання електрона за руху вздовж кругової орбіти у площині, перпендикулярній напрямку магнітного поля \mathbf{B} . На один електрон у

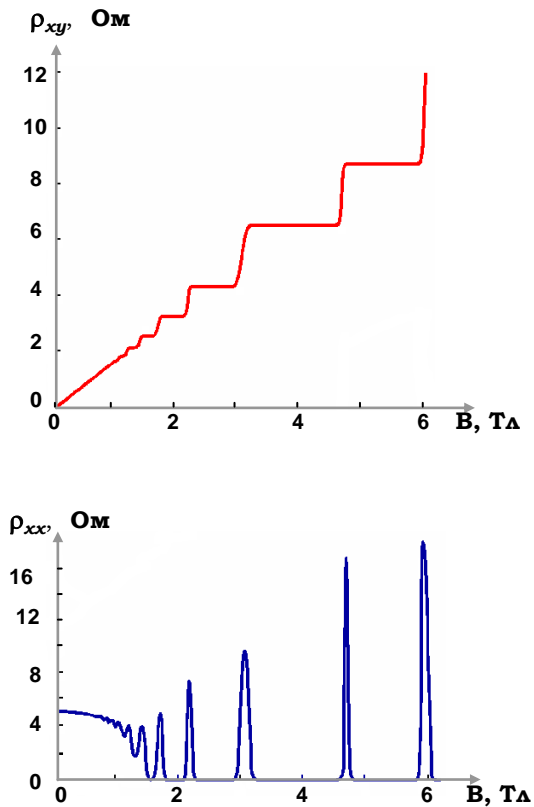


Рис. 16.36. Залежності недиагональної і діагональної компоненти питомого опору квазидвовимірного електронного газу від магнітного поля за сталої концентрації носіїв

ташовуватись удвічі менше електронів, ніж без урахування такого розщеплення. Оскільки це розщеплення є досить великим, то спінові підрівні не перекриваються. Урахування спінового розщеплення рівнів Ландау не надасть нічого нового при розгляді квантового ефекту Холла, тому у подальшому говоритимемо про рівні Ландау, маючи на увазі підрівень із визначеною орієнтацією спіна. За зміни концентрації носіїв або величини магнітного поля \mathbf{B} змінюється положення рівня Фермі щодо системи рівнів Ландау. Уявимо, що рівень Фермі розташований у проміжку між рівнями l та $l+1$. Щільність станів для таких енергій дуже мала, і при $T \rightarrow 0$ усі стани на рівнях Ландау із номерами, меншими за l , заповнені (рис. 16.37). Отже, концентрація електронів дорівнюватиме

площині квантування приходить ся площа $S_B = \pi R_B^2$. Кількість електронів, що можуть розташуватися на кожному із рівнів Ландау, легко визначити із простих міркувань. Вважатимемо, що площа двовимірного металу є одиничною. Тоді число електронів, що розташовані на кожному із рівнів Ландау, є відношенням площини, що займає двовимірна система електронів, до площини електронної орбіти, яка визначається величиною магнітного поля

$$N_B = \frac{eB}{h}. \quad (16.191)$$

Оцінка для поля $B = 10^2$ Тл, показує, що на кожному квадратний сантиметр квазидвовимірного газу електронів приходить ся 10^{12} електронних станів. Зауважимо, що (16.191) ми отримали, взявши до уваги зееманівське розщеплення рівнів у сильному магнітному полі: ми вважали, що в сильному магнітному полі на кожному зі спінових підрівнів Ландау може роз-

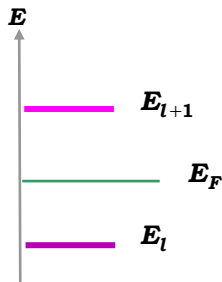


Рис. 16.37. Заповнення електронами рівнів Ландау

$$n = l \cdot N_B . \tag{16.192}$$

Розглянемо стандартну геометрію ефекту Холла (рис. 16.38) і визначимо зв'язок між струмом (j_x), що протікає через зразок, і напругою, яка виникає на бічних гранях зразка. Повторюючи міркування, що ми використовували при розгляді класичного ефекту Холла, запишемо баланс сил, які діють на електрон $-eE_y = ev_x B$, звідки визначимо зв'язок між швидкістю електрона та

поперечним електричним полем $v_x = -E_y / B$. Електричний струм $j_x = -ev_x n$ буде таким чином пов'язаний із поперечним полем співвідношенням

$$j_x = \frac{en}{B} E_y = \rho_{xy}^{-1} E_y ,$$

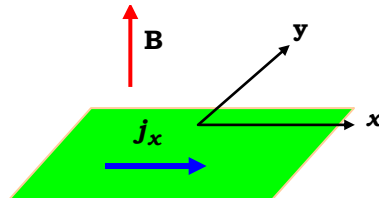


Рис. 16.38. Геометрія ефекту Холла

з якого знайдемо недіагональний компонент тензора опору

$$\rho_{xy} = \frac{B}{en} . \tag{16.193}$$

Підставляючи сюди вираз (16.192), із врахуванням (16.191) дістанемо

$$\rho_{xy} = \frac{h}{l \cdot e^2} . \tag{16.194}$$

Таким чином, недіагональний компонент електричного опору квазідвовимірного електронного газу в магнітному полі змінюється дискретно, обернено пропорційно цілому числу l . Зауважимо, що цей формальний вивід лише вказує на можливість квантування недіагональної провідності й аж ніяк не пояснює квантовий ефект Холла.

Якісна інтерпретація квантового ефекту Холла може базуватись на перколяційній моделі провідності, в якій розглядаються локалізовані та рухливі стани реального двовимірного електронного газу. Розглянемо випадок фіксованої величини магнітного поля та заданого спектру рівнів Ландау зі щільною рухливості 2Δ (рис. 16.39). Локалізовані стани не дають внеску до формування повздовжнього струму. До формування повздовжньої провідності дають внесок лише рухливі стани, тобто такі, що розташовуються поза щільною рухливості. Оскільки заповнення енергетичних станів електронами визначає положення рівня Фермі, то можливі дві принципово різні ситуації у формуванні провідності

двовимірного електронного газу. Перша – рівень Фермі розташовується у щілині рухливості між серединами сусідніх рівнів Ландау. Усі рухливі стани розташовані нижче рівня Фермі, концентрація носіїв в областях, що зайняті рухливими станами, дорівнює максимально можливій N_B . У цьому випадку кожен із l заповнених рівнів Ландау дає внесок до формування поперечного опору

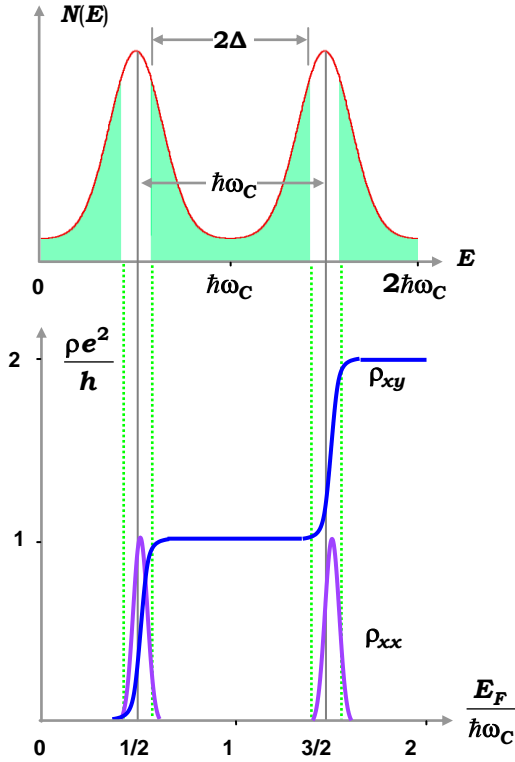


Рис. 16.39. Виникнення квантування опору у перколяційній моделі.

Зеленим позначено області щілини рухливості рухливих станів на l -му рівні змінюється від нуля до максимального значення N_B при проходженні рівнем Фермі області рухливих станів. Цим значенням рівня Фермі відповідає перехідна ділянка між плато холлівського опору із номерами l та $l + 1$. Таке пояснення магнетотранспорту у системі якісно відповідає на питання про походження холлівських плато та виникнення областей бездисипативного повздовжнього руху електронів. Але питання за рахунок чого поперечний опір в області плато квантується із високою точністю, залишається поза межами простої моделі, розглянутої раніше. Пояснення факту квантування поперечного опору ґрунтується на ідеї калібровочної інваріантності, яка у даному контексті означає, що додавання кванта магнітного потоку не змінює енергетичний спектр носіїв, а спричиняє

формування поперечного опору $R_B = h/e^2$, а всі рівні разом – $\rho_{xy} = h/e^2 l$. Ця ситуація відпо-

відає бездисипативному протіканню повздовжнього струму, оскільки за руху електрони не можуть втратити енергію за рахунок процесів розсіювання, адже всі рівні зайняті, а вільний підрівень Ландау, відділений енергетичною щілиною, виникає завдяки зеєманівському розщепленню. Друга – рівень Фермі розташовується в області рухливих станів поблизу піка l -го рівня Ландау. Під дією електричного поля відбувається протікання рухливими станами у смузі шириною $\sim kT$ поблизу рівня Фермі, і оскільки рівні вище рівня Фермі незайняті, струм супроводжується дисипативними процесами. Концентрація електронів в області рух-

лише збудження або дезбудження системи. Якісно застосування цієї ідеї можна зрозуміти, розглянувши уявний експеримент, в якому стрічка двовимірного електронного газу зігнута в петлю, і магнітне поле пронизує її, маючи всюди напрямок за нормаллю до поверхні. Між двома краями кільця прикладено напругу V_H . У разі відсутності дисипації енергія зберігається, тому можна записати закон індукції Фарадея як $I = dE/d\Phi$, де I – струм у петлі, а $dE/d\Phi$ – адіабатична похідна повної енергії системи за магнітним потоком через петлю. Якщо потік зміниться на один квант $\Phi_0 = h/e$, то, як вимагає калібровочна інваріантність, енергетичний спектр має залишатись незмінним. При цьому всі носії зміщуються в сусідні стани так, що один носій на кожному із рівнів Ландау виходить за один край кільця, а інший – входить з іншого краю. Таким чином, через систему переноситься l носіїв, по одному з кожного заповненого рівня Ландау. Якщо рівень Фермі розташований у щілині рухливості, то дисипація енергії відсутня, і повна зміна енергії відповідає переходу l електронів від одного краю кільця до іншого, тобто $\Delta E = leV_H$. Звідси, пам'ятаючи, що $\Delta\Phi = \Phi_0$, знайдемо, що

$$I = \frac{\Delta E}{\Delta\Phi} = l \frac{e^2}{h} V_H, \quad (16.195)$$

звідки отримуємо (16.194). У цій інтерпретації основною причиною виникнення квантування холловського опору є квантування магнітного потоку на елементарні кванти $\Phi_0 = h/e$ та електричного заряду на елементарні заряди e . Незалежність кванту опору від типу матеріалу та умов експерименту тепер здається цілком слушною. Цілочисельний квантовий ефект Холла таким чином ми зрозуміли. Пояснення дробного квантового ефекту Холла неможливе за напівінтуїтивного підходу, воно вимагає застосувань спеціальних методів теоретичної фізики і виходить за межі цього підручника.

Оскільки відношення h/e^2 пов'язане зі сталою тонкої структури α формулою $h/e^2 = 1/\alpha c$, то, вимірюючи поперечний опір ρ_{xy} (із точністю $\sim 10^{-7}$), можна дуже точно вимірювати сталу тонкої структури. Таким чином, квантовий ефект Холла знайшов практичне застосування у метрології.

■ 16.10. Оптичні властивості квантово-розмірних напівпровідникових структур

Як зазначалося в розд. 15, оптичні властивості напівпровідників базуються на можливості поглинання та емісії фотонів напівпровідниками. А характерними особливостями нанорозмірних напівпровідникових структур є: по-перше, можливість конструювання штучних напівпровідникових матеріалів із наперед заданими електронними властивостями, у тому числі й зумовленою зонною структурою; по-друге, нанорозмірні структури є ідеальними для застосувань у приладах генерації світла. Крім того, у таких структурах може спостерігатись так званий ефект підсилення локального поля, що робить нанорозмірні структури цікавими з погляду нелінійно-оптичних застосувань.

16.10.1. Використання гетероструктур для побудови приладів підсилення світла.

У розділі, що присвячений оптичним властивостям напівпровідників, зазначалося, що для виникнення інверсії заселеності ширина активної області є критичним параметром. Відомо, що наявність надлишкової концентрації носіїв в активній області спричиняє дифузійне вимивання носіїв із цієї області. Цьому витіканню носіїв запобігає кінцевий час їхнього життя. Для електрон-діркових пар цей час життя визначається часом рекомбінації τ . Якщо коефіцієнт дифузії носіїв в активній області D , то характерна ширина d області надлишкової концентрації нерівноважних носіїв не має бути меншою за дифузійну довжину $L_D = \sqrt{D\tau}$ (L_D – середня відстань, що проходить електрон (дірка) перед тим, як рекомбінувати). Оскільки дифузійна довжина носіїв у сполуках типу A_3B_5 має порядок кількох мікрометрів, неможливо зробити активну область меншою за ці кілька мікрометрів – типову дифузійну довжину гомоструктури p - n переходу. Для локалізації нерівноважних носіїв у малій активній області необхідно використати структуру із двома гетеропереходами. Основна ідея використання подвійної гетероструктури – створення потенціальних бар'єрів по обидва боки p - n переходу – що запобігає запобіганню дифузії як електронів, так і дірок. При цьому зовнішній потенціал Φ вибирається таким, щоб відбувалось спрямлення зон, а це ефективно спричиняє затягування полем електронів із n -області, а дірок – із p -області p - n -переходу до області квантової ями. Профіль потенціалу подвійної гетероструктури схематично подано на рис. 16.40. Гетероструктура складається із трьох матеріалів, що характеризуються шириною за-

бороненої зони E_{g1} , E_{g2} та E_{g3} . Зонні уступи вибрані так, щоб виникла структура із бар'єром для електронів у лівій частині перед p -областю і бар'єр для дірок – у правій частині перед n -областю. Середина область що характеризується шириною забороненої зони E_{g2} , доступна для обох типів носіїв і служить активною областю. У цьому випадку не довжини

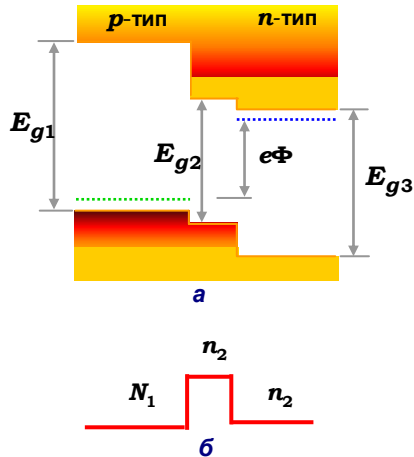


Рис. 16.40: а – енергетична діаграма подвійної гетероструктури, що може служити підсилювачем світла; б – схема індексу рефракції структури із подвійною гетеро границею. Прикладений потенціал Φ задовольняє умові плоских зон

дифузії, а відстань між бар'єрами визначає ширину активної зони. У результаті ця ширина може бути меншою за 0,1 нм, що приведе до того, що критичний струм може бути зменшений принаймні на порядок порівняно з відповідною гомоструктурою p - n переходу.

Іншою перевагою при використанні гетероструктур для підсилення світла є особливості розповсюдження світла у вузькій активній області. Неважко зрозуміти, що у просторово неодноріних середовищах концепція простої плоскої хвилі втрачає сенс. Розповсюдження та підсилення світла тепер сильно залежатиме від величини та геометрії неоднорідностей, так само як і від напрямку розповсюдження світла. Крім то-

го, неоднорідність середовища може спричиняти формування електромагнітних мод в області високих значень індексу рефракції. Ці моди можна розглядати як локалізовані фотони, і як наслідок – збільшення концентрації таких локалізованих фотонів стане підсиленням цих мод. Якщо розмір активної області є меншим за розміри локалізації моди, підсилення моди зменшується пропорційно відношенню об'ємів цих областей. Це явище можна описати у термінах фактора оптичного конфайнменту. У випадку, якщо активна зона є плоским шаром, перпендикулярним до вісі OZ , цей фактор

$$\Gamma \approx \frac{V_{AR}}{V_S} \frac{\int |E(z)|^2 dz}{\int |E(z)|^2 dz}, \quad (16.196)$$

де V_{AR} та V_S є – об'єми активної області та всього зразка, відповідно. Для плоскої хвилі цю величину можна наближено оцінити як

$$\Gamma \approx \frac{d}{L_z}, \quad (16.197)$$

де L_z – розмір локалізації світлової моди, яка для одноперехідного приладу може збігатися із розмірами зразка у напрямку z . Фактор Γ фактично являє собою частку електромагнітної енергії, що розповсюджується в активній області щодо енергії "закачаної" до всього зразку. Зрозуміло, що для вузької активної області цей фактор є малим. І зменшення підсилення завдяки малому фактору Γ є критичним. Наступним негативом, що пов'язаний із використанням вузької активної зони, є поглинання світла в неактивних областях, сусідніх з активною зоною. Навіть якщо таке поглинання не є істотним, воно може спричинити суттєве зменшення аж до зникнення ефекту оптичного підсилення. Із використанням подвійної гетероструктури цих проблем вдається уникнути. Для з'ясування деталей звернемося ще раз до рис. 16.40. Шар 1 – p -допований і характеризується щільною E_{g1} та індексом рефракції n_1 . Середній шар структури характеризується зонною щільною E_{g2} та індексом рефракції n_2 . Зазвичай цей шар є нелегованим. Третій шар структури легований донорами, тобто характеризується електронною провідністю (n -тип), зонною щільною E_{g3} і рефрактивним індексом n_3 . Якщо енергії заборонених зон задовольняють умові $E_{g2} < E_{g1}, E_{g3}$, індекс рефракції середнього шару є більшим за рефрактивні індекси зовнішніх шарів структури. Таким чином, за створення такої структури:

- ♦ зонна дискретність у валентній зоні та зоні провідності спричиняє конфайнмент нерівноважних електронів і дірок усередині активного шару. Порівняно зі звичайним p - n -переходом ця подвійна гетероструктура акумулює обидва типи носіїв у більш вузькій області;

- ♦ великий індекс рефракції у середньому вузькому шарі спричиняє оптичний конфайнмент у цьому шарі: завдяки значній неоднорідності системи в ній можуть виникати локалізовані на активній області моди, що ефективно збільшує коефіцієнт підсилення.

Як приклад реалізації двобар'єрної гетероструктури можна навести GaAs/AlGaAs – діод з активною зоною виконаною на базі GaAs із $E_{g2} = 1.42$ eV та $n_2 = 3.6$. Оточуючі шари виконані із $\text{Al}_x\text{Ga}_{1-x}\text{As}$ з $0.35 \leq x \leq 0.5$. Вони характеризуються забороненою зоною, більшою за 1.43 eV і рефрактивним індексом, меншим за 3,6 на 5–10 %. Такий діод може підсилювати світло у діапазоні довжин хвиль 0.82–0.88 мкм.

За зменшення товщини активного шару у структурі починають все більшу роль відігравати квантоворозмірні ефекти. Але необхідно усвідомлювати, що зменшення товщини активної зони у системі, що

подана на рис. 16.40, не приведе до істотних переваг від внесків просторового квантування. Це твердження легко зрозуміти, якщо згадати, що для виникнення ефекту підсилення світла потрібно мати конфайнмент як для електронної, так і для діркової компоненти електронно-діркової системи. У такому випадку необхідно застосовувати інші конструкції квантоворозмірних систем, наприклад такі, як на рис. 16.41. Для таких принципових схем електрони та дірки, що генеруються зовнішнім випромінюванням або інжекцією із p і n областей, рухаються у бар'єрних шарах і захоплюються в активній зоні квантової ями. Характерний час цього захоплення менший за 1 пкс. Протилежні процеси вимивання носіїв з активної зони потребують додаткової енергії і мають дуже низьку ймовірність реалізації. Носії у квантовій ямі релаксують на найнижчі можливі рівні енергії, що сприяє накопиченню носіїв у надзвичайно вузькій активній області, яка має типові розміри ≤ 10 нм. Схожа ситуація може реалізуватись і в квантових нитках, і квантових точках.

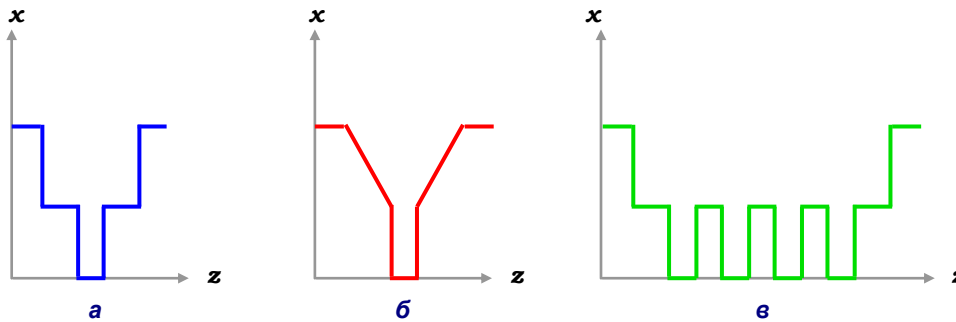


Рис. 16.41. Профілі квантоворозмірних структур на основі $(\text{Al}_x\text{Ga}_{1-x})_{0.5}\text{In}_{0.5}\text{P}$, в яких одночасно реалізується як конфайнмент носіїв, так й оптичний конфайнмент:
 а – одинична квантова яма зі східцеподібним розподілом індексу рефракції;

б – одинична квантова яма з лінійним розподілом індексу рефракції;

в – надгратка зі східцеподібним розподілом індексу рефракції.

За віссю OZ відкладено ширину відповідного шару гетероструктури,

а за віссю OX – частку алюмінію у сполуці $(\text{Al}_x\text{Ga}_{1-x})_{0.5}\text{In}_{0.5}\text{P}$

16.10.2. Умови виникнення інверсії заселеності рівнів у квантовій ямі

У розд. 15 було розглянуто умови виникнення інверсної заселеності рівнів в об'ємному напівпровіднику. Ці умови зводились до нерівності $E_g < \hbar\omega < E_{Fn} - E_{Fp}$. Аналогічні умови можна записати й для квантової

ями. Відмінності полягатимуть тільки в тому, що у лівій частині нерівності необхідно записати енергію оптичної зонної щілини

$$E_g + \epsilon_{C1} + \epsilon_{V1} < \hbar\omega < E_{Fn} - E_{Fp}, \quad (16.198)$$

де ϵ_{C1} та ϵ_{V1} – енергії дна та стелі найнижчих електронної і діркової підзон у квантовій ямі, відповідно. Як відомо, оптична щільність станів і величина інверсної заселеності рівнів визначають частотну залежність коефіцієнта підсилення світла

$$\alpha \propto \omega \cdot \rho^{QW}(\omega) \cdot [\mathfrak{I}(E_C) - \mathfrak{I}(E_V)]. \quad (16.199)$$

Використовуючи функцію щільності станів квантової ями, а також вважаючи, що різниця чисел заповнення є майже східцеподібною функцією енергії, отримуємо спектральну поведінку коефіцієнта підсилення квантової ями (рис. 16.42). Для порівняння подано також спектральну характеристику коефіцієнта підсилення світла для об'ємного зразка. Із поведінки $\alpha(\omega)$ для квантової ями видно, що тільки найнижчі енергетичні стани електронів і дірок дають внесок до формування інверсної заселеності рівнів. Крім того, східцеподібна поведінка щільності станів спричиняє звуження спектральної області підсилення світла. Цей факт використовується в лазерах на квантових ямах для генерації меншого числа оптичних мод. Таким чином, застосування квантових ям у багатьох випадках дає переваги для досягнення поліпшення спектральних характеристик лазерів.

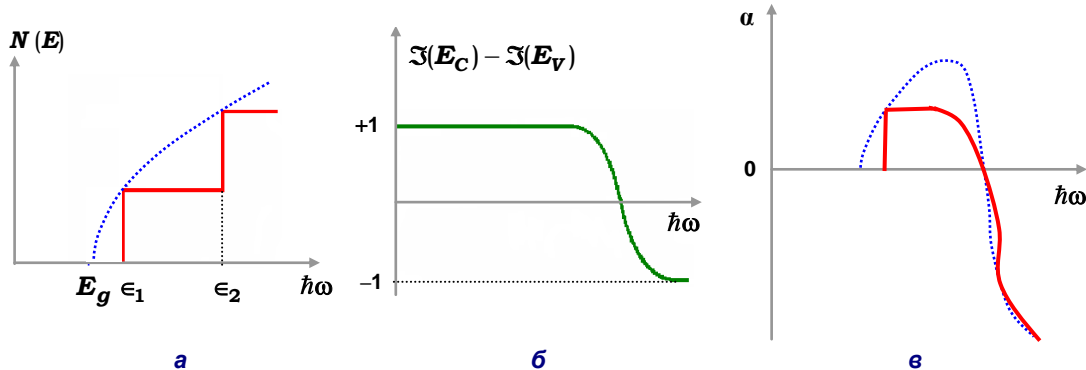


Рис. 16.42. Вплив щільності станів (а) та фактора заповнення (б) на формування коефіцієнта підсилення світла (в) квантовою ямою (червона крива).

Для порівняння подано результат для об'ємного зразка (синя крива)

Треба мати на увазі також наявність ефекту оптичного конфайнменту у тонких шаруватих структурах (локалізація оптичних мод на

просторових неоднорідностях). Таким чином, коефіцієнт підсилення світла квантоворозмірної системи можна записати як

$$\alpha(\omega) = \Gamma\alpha_0(\omega), \tag{16.200}$$

де $\alpha_0(\omega)$ - коефіцієнт підсилення об'ємним зразком матеріалу. Вузькозонний шар, що відіграє роль квантової ями, може одночасно служити хвильовідною планарною структурою для світла. На рис. 16.43 подано розподіл інтенсивності в найнижчій оптичній моді, локалізованій сідцеподібною структурою. Видно, що ступінь оптичної локалізації зменшується зі зменшенням товщини обмежуючої структури. Таким чином, одинична квантова яма не може спричинити значну локалізацію світлової хвилі. Відповідно й коефіцієнт Γ буде у такій ямі незначним. Для локалізації світла всередині квантової ями з одночасним високим рівнем захвату світла можна запропонувати спеціальний тип активної області гетероструктури. Схему двох таких типів активної зони зі змінним індексом рефракції для структур InGaAsP/InGaP подано на рис. 16.44. Немонотонну поведінку $\Gamma(d/2)$ можна зрозуміти з таких міркувань. При $d \rightarrow 0$ та $d \rightarrow \infty$ маємо конфайнмент моди у простій одиничній квантовій ямі. Зрозуміло, що для обмежених значень d реалізуватимуться проміжні (із більшою локалізацією світла)

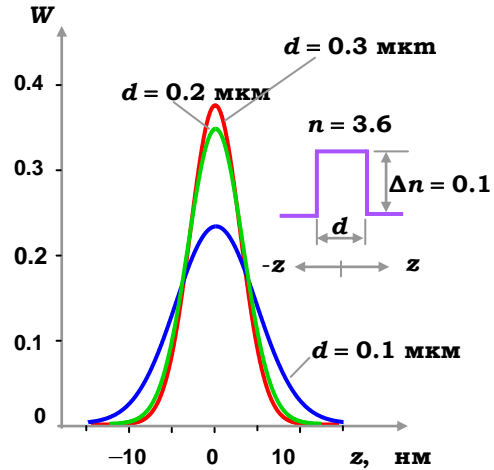


Рис. 16.43. Розподіл інтенсивності поля (відн. од.) нижньої оптичної моди ($\lambda = 0.9$ нм), локалізованої у квантовій ямі, шириною d зі сідцеподібним індексом рефракції (див. урізання)

подано на рис. 16.44. Немонотонну поведінку $\Gamma(d/2)$ можна зрозуміти з таких міркувань. При $d \rightarrow 0$ та $d \rightarrow \infty$ маємо конфайнмент моди у простій одиничній квантовій ямі. Зрозуміло, що для обмежених значень d реалізуватимуться проміжні (із більшою локалізацією світла)

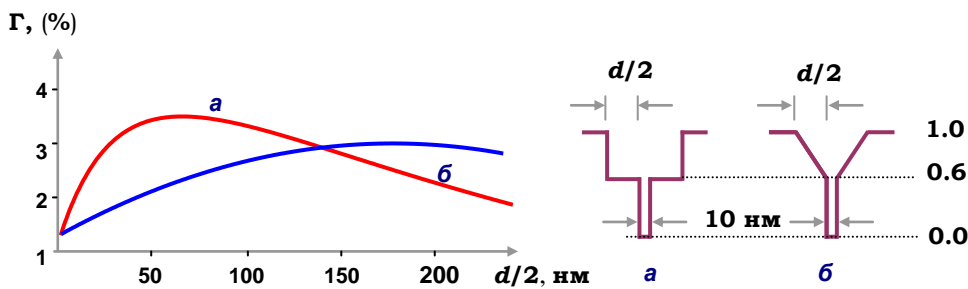


Рис. 16.44. Коефіцієнт оптичного конфайнменту лазерної структури типу а та б на основі $(Al_xGa_{1-x})_{0.5}In_{0.5}P/Ga_{0.5}In_{0.5}P$ одиничної квантової ями

значення Γ . У випадку накачування шляхом інжекції носіїв квантова яма та структура, де відбувається локалізація світла (хвилевод) мають бути включені до області інжекції. У цьому випадку для квантової ями InGaAs/InP із товщиною вузькозонного шару 5 нм вже за струмів інжекції $23 \text{ A}\cdot\text{см}^{-2}$ коефіцієнт підсилення α_0 досягає 100 см^{-1} . Для квантової ями AlGaAs/GaAs такої самої ширини цієї величини коефіцієнт підсилення світла досягає за струму інжекції, більшого за $60 \text{ A}\cdot\text{см}^{-2}$.

16.10.3. Світловипромінюючі діоди та лазери

Виконуючи комплекс заходів для підвищення ефективності лазерної генерації, згідно із міркуваннями, що обговорювались вище, можна суттєво поліпшити характеристики світловипромінюючого пристрою. При цьому варто мати на увазі, що при утворенні певних квантово-розмірних структур можна досягти потрібних, наперед заданих частотних характеристик приладу, яких у принципі неможливо досягти за допомогою використання природних напівпровідникових матеріалів. Наприклад, світловипромінюючі діоди покривають широкий діапазон частот – від інфрачервоного $\sim 8,0 \text{ мкм}$ (на основі сполуки InGaAsP) до ближнього ультрафіолету – $0,4 \text{ мкм}$ (на основі сполуки GaN): фіолетовий ($\lambda = 0,4 \text{ мкм}$) – GaN; зелений ($\lambda = 0,58 \text{ мкм}$) – GaP:N; жовтий ($\lambda = 0,6 \text{ мкм}$) – GaAs_{0,14}P_{0,86}; червоний ($\lambda = 0,7 \text{ мкм}$) – GaAs_{0,6}P_{0,4}; інфрачервоний ($\lambda = 0,9 \text{ мкм}$) – GaAs, ($\lambda = 1,23 \text{ мкм}$) – In_{0,72}Ga_{0,28}As_{0,6}P_{0,4}. Світловипромінюючі діоди виконуються у двох типах конструкцій – поверхневовипромінюючі (а) і випромінюючі з бічної грані (б) як схематично подано на рис. 16.45.

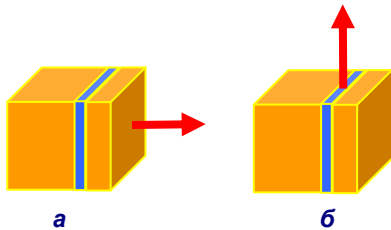


Рис. 16.45. Світлодіоди на основі структури із квантовою ямою

Випромінювання світла у потрібному напрямку забезпечується за допомогою утворення поглинаючого шару із протилежного до напрямку випромінювання або за допомогою відбиття від металічних контактів.

Інжекція носіїв до активної області напівпровідникового діода – найбільш прийнятний метод генерації нерівноважних електронів і дірок для створення інверсії заселеності та стимульованої емісії. Незважаючи на те, що стимульована емісія з області інжекції діода є важливим ефектом, підпороговий стан світловипромінюючого діода, коли випромінювання відбувається тільки завдяки спонтанній емісії, у багатьох випадках є більш прийнятним. Для таких діодів не потребу-

ється утворення зворотного зв'язку для контролю вихідної потужності, вони можуть працювати у широкому температурному діапазоні та є недорогими приладами. Розглянемо принцип дії таких діодів. Для p - n -діодів із прямим зміщенням швидкість спонтанної емісії (тобто повний потік фотонів)

$$\gamma = \frac{n}{\tau_R} V_0, \quad (16.201)$$

де n – концентрація електрон-діркових пар в активній області (об'ємом V_0) діода, τ_R – час життя пари перед рекомбінацією із випромінюванням світла. З іншого боку, швидкість інжекції електронів і дірок в одиничний об'єм за одиницю часу

$$\mathfrak{R} = \frac{I}{eV_0}, \quad (16.202)$$

де I – струм, e – заряд електрона. Таким чином, концентрація носіїв визначатиметься повним часом життя нерівноважних електрон-діркових пар τ

$$n = \tau \cdot \mathfrak{R}. \quad (16.203)$$

Відношення часів життя τ_R та τ характеризує власну ефективність емісії з p - n переходу

$$\eta_i = \frac{\tau}{\tau_R}, \quad (16.204)$$

звідки маємо

$$\gamma = \eta_i \frac{I}{e}. \quad (16.205)$$

Зрозуміло, що не всі фотони, які емітують із активної області, дають внесок до формування вихідного світлового потоку. Дійсно, поглинання світла всередині активної області, внесок безвипромінювальної рекомбінації тощо приводять до того, що не всі емітовані фотони формують вихідний світловий потік. Тому зручно ввести величину, що характеризує вихідну ефективність діода. За аналогією із (16.205) запишемо

$$\gamma_{out} = \eta_{ext} \frac{I}{e}. \quad (16.206)$$

Величина η_{ext} є відношенням вихідного потоку фотонів до потоку інжекттованих електронів. Очевидно, що вона залежить як від фізичних механізмів емісії фотонів, так і від конструкції діода. Оптична потужність світловипромінюючого діода на частоті ω , таким чином,

$$P_{opt} = \hbar\omega\gamma_{out} = \eta_{ext}\hbar\omega\frac{I}{e} \quad (16.207)$$

Але пряма пропорційність потужності струму інжекції справедлива за невеликих струмів інжекції. За великих рівнів інжекції зростання потужності відбувається значно повільніше за I .

Відомо, що лазерний ефект спостерігається за умови наявності двох складових – підсиленню світла завдяки стимульованій емісії та оптичному зворотному зв'язку. Найпростіший спосіб отримання оптичного зворотного зв'язку – створення активної зони підсилення між двома дзеркалами, що являють собою оптичний резонатор. Тоді, проходячи через активну область, світлова хвиля збільшує амплітуду. Далі, відбиваючись від поверхні дзеркала, вона втрачає частину інтенсивності – амплітуда хвилі трохи зменшується. Знов хвиля проходить через активну область, де підсилюючись, збільшує амплітуду й знов при відбиванні за рахунок втрат амплітуда зменшується – і так багаторазово. Зрозуміло, якщо підсилення перевершує оптичні втрати, то амплітуда світлової хвилі всередині резонатора зростає. Але зростання не може бути необмеженим за рахунок ефекту насичення підсилення. Фізична причина такого насичення полягає в тому, що підвищення інтенсивності світла приводить до додаткової електрон-діркової рекомбінації через випромінювальні канали і тим самим – до зменшення інверсної заселеності рівнів. А зменшення інверсної заселеності спричиняє зменшення коефіцієнта підсилення. Таким чином, стаціонарний стан

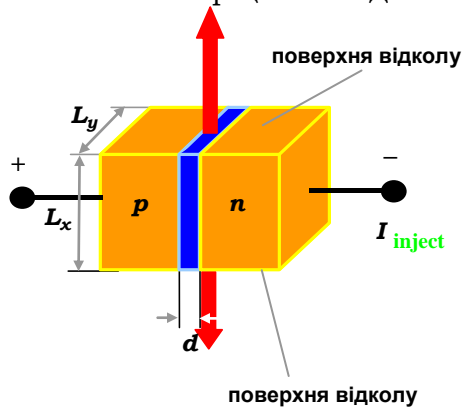


Рис. 16.46. Схема інжекційного лазера з оптичним резонатором, що утворений двома поверхнями кристалічного вигнутого шару (d – товщина активного шару)

лазерної генерації встановлюється, якщо ефект підсилення дорівнює втратам. Досягнення такого стану можна вважати критерієм сталості амплітуди в оптичному резонаторі. Для визначення порогового рівня накачування розглянемо напівпровідниковий інжекційний лазер, де оптичний зворотній зв'язок реалізується шляхом утворення кристалічних площин вигнутого шару перпендикулярних до площини p - n -контакту (рис. 16.46). Визначимо коефіцієнт відбиття від площин вигнутого шару як відношення інтенсивності відбитого світла до інтенсивності падаючого

$$r = \frac{I_r}{I_{in}}, \quad (16.208)$$

де I_r та I_{in} – інтенсивності відбитого та падаючого на площину відколу світла, відповідно. З іншого боку коефіцієнтом відбиття для межі розподілу напівпровідник–повітря є

$$r = \left(\frac{n_{ri} - 1}{n_{ri} + 1} \right)^2, \quad (16.209)$$

де n_{ri} – коефіцієнт рефракції. Тоді інтенсивність світла, що проходить через дзеркало,

$$I_{out} = (1 - r)I_{in}. \quad (16.210)$$

Нехай поверхні відколу характеризуються двома коефіцієнтами відбиття r_1 та r_2 . Після двох проходів через прилад інтенсивність світла зменшується в $r_1 \cdot r_2$ разів. Визначимо ефективний коефіцієнт оптичних втрат як

$$\alpha_r = \frac{1}{2L_x} \ln \frac{1}{r_1 r_2}. \quad (16.211)$$

Як видно із рис. 16.46, L_x є відстанню між поверхнями, що утворюють оптичний резонатор. Зазначимо, що область оптичної комірки $L_x \times L_y$ може бути значно більшою за активну область приладу. Це означає, що інші механізми втрат відмінні від тих, що пов'язані з ефектами рекомбінації, можуть додатково визначати коефіцієнт поглинання. Нехай ці додаткові втрати визначаються коефіцієнтом α_s . Тоді повний коефіцієнт оптичних втрат

$$\alpha_{ls} = \alpha_r + \alpha_s. \quad (16.212)$$

Якщо коефіцієнт підсилення деякої оптичної моди у резонаторі перевищує коефіцієнт оптичних втрат, виникає лазерний ефект. Таким чином, критерій лазерної генерації запишемо як

$$\alpha \geq \alpha_{ls} = \alpha_r + \alpha_s. \quad (16.213)$$

З іншого боку, коефіцієнт підсилення можна виразити через порогове значення щільності струму та щільності струму інжекції як

$$\alpha = \gamma \left(\frac{j}{j_{th}} - 1 \right). \quad (16.214)$$

Тоді критерій початку лазерної генерації запишемо як співвідношення між пороговим струмом, що спричиняє появу інверсної заселеності j_{th} , і пороговим струмом лазерної генерації j_{th}^l

$$j_{th}^l = \frac{\alpha_{ls} + \gamma}{\gamma} j_{th}. \quad (16.215)$$

Звідси видно, якщо $\alpha_{ls} \ll \gamma$, то порогові густини струмів майже збігаються. Наприклад, лазерний діод, побудований на структурі InGaAsP, із розмірами $d \times L_x \times L_y = 2 \times 10 \times 200 \text{ нм}^3$, $\alpha_{ls} = 120 \text{ см}^{-1}$ та $\gamma = 600 \text{ см}^{-1}$ при температурі 300 К має співвідношення між пороговими струмами $J_{th}^l = 1.2 \cdot j_{th}$. Оскільки пороговий струм виникнення інверсної заселеності для цієї структури становить $j_{th} = 4.6 \cdot 10^4 \text{ А} \cdot \text{см}^{-2}$, то маємо $J_{th}^l = 5.6 \cdot 10^4 \text{ А} \cdot \text{см}^{-2}$, а повний струм $I = 1.1 \text{ А}$. Такий великий струм через систему з однорідним p - n -переходом вимагає застосування спеціальних заходів для відведення тепла. На противагу цьому у випадку приладу із подвійним гетероконтактом порогові струми можуть бути набагато меншими, таким чином лазерні діоди можуть ефективно працювати й за кімнатних температур. Наприклад, для активної області приладу із подвійним гетероконтактом $d = 0,1 \text{ мкм}$ і коефіцієнтом оптичного конфайнменту $\Gamma = 1$, порогове значення щільності струму для виникнення інверсної заселеності $j_{th} = 2.3 \cdot 10^3 \text{ А/см}^2$, а повне значення порогу щільності струму $J_{th}^l = 55 \text{ мА}$ що є цілком прийнятною величиною.

Таким чином, явно видно переваги, що надає використання квантової ями як активної області інжекційного лазера. Розглянемо детальніше можливості оцінювання порогового струму для виникнення лазерної генерації. Почнемо із того, що перепишемо (16.200) у термінах коефіцієнта оптичного конфайнмента квантової ями із хвилеводною структурою і коефіцієнта підсилення одиничної квантової ями

$$\alpha(\omega) = \Gamma_0 \alpha_0(\omega). \quad (16.216)$$

Умови початку лазерної генерації перепишемо у вигляді

$$\Gamma_0 \alpha_0(\omega) > \alpha_s + \frac{1}{2L_x} \ln \frac{1}{r_1 r_2}. \quad (16.217)$$

Внутрішні оптичні втрати, які описує коефіцієнт α_s , включають втрати за рахунок розсіювання поза хвилеводною структурою – α_{sc} , втрати в пасивній області хвилеводу – α_{pp} , а також втрати в активній області – в області квантової ями – α_{qw} . Ці втрати можна записати у вигляді

$$\alpha_s = \alpha_{sc} + \Gamma_0 \alpha_{qw} + (1 - \Gamma_0) \alpha_{pp}. \quad (16.218)$$

Зауважимо, що головним механізмом втрат у пасивній області є поглинання на вільних носіях. Цей механізм стає важливим за високих рівнів інжекції, коли велика кількість носіїв рухається через бар'єрну

область і формує тим самим величину α_{pp} . Далі перепишемо (16.214) у вигляді

$$\alpha_0 = \gamma \left(\frac{j}{j_{th}} - 1 \right). \quad (16.219)$$

Підставляючи сюди α_0 із (16.217) і використовуючи (16.218), отримаємо умову для величини струму інжекції, при якого виникає лазерний ефект

$$j > j_{th} \left(1 + \frac{\alpha_{sc}}{\gamma \Gamma_0} + \frac{1}{2\gamma \Gamma_0 L_x} \ln \frac{1}{r_1 r_2} \right). \quad (16.220)$$

Як видно, суттєву роль у формуванні порогового струму відіграє коефіцієнт оптичного конфайнмента. Оцінимо величину порогового струму для квантової ями InGaAsP де $j_{th} \approx 20 \text{ А} \cdot \text{см}^{-2}$, $\gamma = 10^3 \text{ см}^{-3}$. Коефіцієнт оптичного конфайнмента $\Gamma_0 = 0.03$. Типовою величиною $\alpha_{sc} = 3 \text{ см}^{-1}$. Тоді для порогового струму маємо нерівність $j < 20 + 2 + (1/3L_x) \ln(1/r_1 r_2)$. Для подальшої оцінки приймемо, що довжина оптичної області $L_x = 200 \text{ мкм}$, а коефіцієнти відбиття від поверхонь кристалічного сколу $r_1=r_2=0.31$. Тоді отримуємо для порогового струму цілком прийнятне значення $j = 60 \text{ А} \cdot \text{см}^{-2}$.

Таким чином, можна стверджувати, що застосування квантово-розмірних структур у розробці нових типів світловипромінюючих пристроїв привело до справжньої технологічної революції. Розроблено і випускається у промислових масштабах велика кількість нових типів дешевих та ефективних лазерів і світлодіодів, що випромінюють у широкому діапазоні частот від інфрачервоного до ультрафіолетового.

16.10.4. Фотоприймальні пристрої на основі квантоворозмірних структур

Фотоприймачі та фотоприймальні пристрої є важливими елементами будь-якої оптико-електронної системи і за сутністю визначають її основні характеристики. Одним із найбільш перспективних методів реєстрації інфрачервоного випромінювання є ІЧ-детектування на основі квантово-розмірних структур, зокрема структур із багатьма квантовими ямами. Головною перевагою таких систем при побудові фотоприймальних пристроїв є можливість, комбінуючи склад (матеріали) і геометрію систем, досягати найбільш ефективних параметрів:

ширини робочого діапазону частот за прийнятних значень коефіцієнта перетворення, що характеризує частку енергії потоку випромінювання, яке падає на приймач і перетворюється на електричний струм. Наприклад, на основі структур із квантовими ямами InGaAs/InP, InGaAs/InAlAs, AlGaInAs/InP (рис. 16.47) створюють фотоприймачі інфрачервоного діапазону, що працюють в інтервалі довжин хвиль 3-20 мкм.

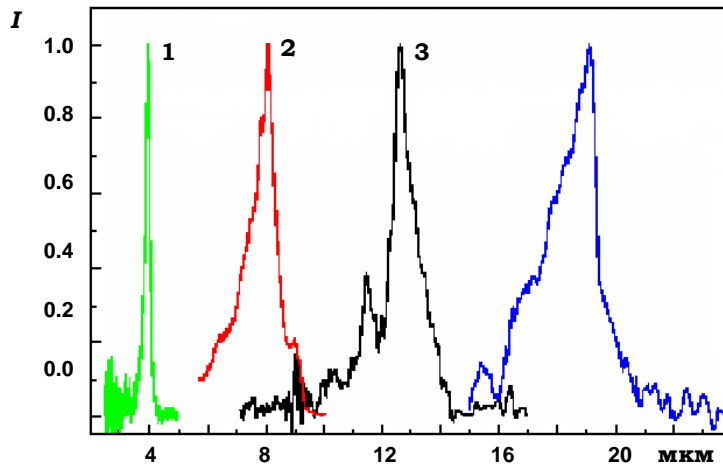


Рис. 16.47. Спектри поглинання структур із квантовими ямами на основі: 1 – InGaAs/InP; 2 – InGaAs/InAlAs; 3 – AlGaInAs/InP

■ 16.11. Задачі

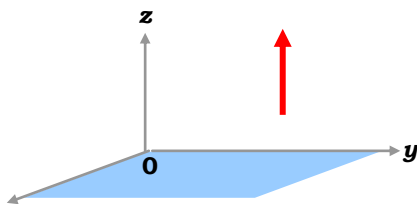


Рис. 16.48

1. Обчислити щільність станів двовимірного електронного газу у квантуючому магнітному полі.

Розв'язок. Розглянемо двовимірний електронний газ, локалізований в площині XOY, на який діє магнітне поле, направлене за нормаллю до цієї площини (рис. 16.48). У результаті впливу зовнішнього магнітного поля у зазначеній геометрії рух електронів у площині XOY

стане квантованим з енергією $E_n = \hbar\omega_C(n + 1/2)$. Це означає, що щільність станів одиниці площі такого газу матиме вигляд

$$N(E) = \sum_n \delta(E - \hbar\omega_C(n + 1/2)), \quad (16.221)$$

При цьому, згідно із (16.73) має виконуватись граничний перехід

$$\lim_{B \rightarrow 0} N_{2D-B}(E) = \frac{m^*}{\hbar^2 \pi}. \quad (16.222)$$

Маючи це на увазі, запишемо щільність станів у вигляді

$$N_{2D-B}(E) = \sum_n N_B^{(n)}(E), \quad (16.223)$$

$$N_B^{(n)}(E) = g(B) \delta(E - \hbar\omega_C(n + 1/2)), \quad (16.224)$$

де $g(B)$ – коефіцієнт, що забезпечує виконання умови (16.222). При $B \rightarrow 0$ відстань між двома енергетичними рівнями E_n та E_{n+1}

$$\Delta E = \hbar \frac{eB}{m^*} \rightarrow 0. \quad (16.225)$$

Це означає, що за підсумовування у (16.223) крок підсумовування стає нескінченно малим. Отже можна перейти від підсумовування до інтегрування. Таким чином, замість (16.223) запишемо

$$N_{2D-B}(E) = g(B) \sum_n \delta(E - \hbar\omega_C(n + 1/2)) = \frac{g(B)}{\hbar\omega_C} \int_0^\infty d\xi \delta(E - \hbar\omega_C/2 - \xi). \quad (16.226)$$

За умови $E - \hbar\omega_C/2 > 0$ межі інтегрування можна розповсюдити на область від'ємних ξ , що дозволить безпосередньо виконати інтегрування та дістати

$$\lim_{B \rightarrow 0} N_{2D-B}(E) = \frac{g(B)}{\hbar\omega_C} = \frac{m^*}{\pi\hbar^2}. \quad (16.227)$$

Звідси знайдемо
$$g(B) = \frac{eB}{\pi\hbar}. \quad (16.228)$$

Підставляючи цей вираз до (16.224), остаточно запишемо

$$N_{2D-B}(E) = \frac{eB}{\pi\hbar} \sum_n \delta(E - \hbar\omega_C(n + 1/2)). \quad (16.229)$$

2. Обчислити щільність станів квазідвовимірному електронного газу у квантуючому магнітному полі. Порівняти результат із результатом попередньої задачі.

3. Обчислити енергетичні стани сферичної квантової GaAs точки у середовищі $Al_{0.3}Ga_{0.7}As$. Яким має бути мінімальний радіус такої квантової точки, для виникнення зв'язаних станів у ній?

4. Обчислити енергетичні стани у квантовій ямі GaAs шириною 20 нм, що сформована на підкладці $Al_{0.3}Ga_{0.7}As$. Плівка GaAs покрита товстим шаром $Al_{0.25}Ga_{0.75}As$.

5. Обчислити коефіцієнти проходження поперечного руху електрона у структурі, що описана у попередній задачі.

Список літератури

1. *Davis J.* The Physics of Low-Dimensional Semiconductors. An Introduction (Cambridge University Press, 1998).
2. *Mitin V.V., Kochelap V.A., Strosio M.A.* Quantum Heterostructures (Cambridge University Press, 1998)
3. *Херман М.* Полупроводниковые сверхрешетки. – М.: Мир, 1990.
4. Physics of Low-Dimensional Semiconductor Structures / Ed. by *Butcher P., March N.H.* and *Tosi M.P.* (Plenum Press, N.Y. and London, 2001).
5. *Göbel E.O., Ploog K.* Fabrication and Optical Properties of Semiconductor Quantum Wells and Superlattices, *Prog.Quant.Electr.* – 1990.- v.14.- p.289-356.

Розділ 17

НЕВПОРЯДКОВАНІ НАПІВПРОВІДНИКОВІ СТРУКТУРИ

Структура та властивості неупорядкованих конденсованих систем викликають значний інтерес дослідників. Неупорядковані матеріали, крім того, мають практичні застосування. Неупорядковані системи – сильно леговані кристалічні напівпровідники, сплави, аморфні напівпровідники тощо – є, можна сказати, об'єктами загального типу, а упорядковані структури типу кристалічної ґратки являють собою ідеалізовані об'єкти, що в чистому вигляді не зустрічаються у природі. Незважаючи на це, основними об'єктами теоретичних побудов твердих тіл були, в основному, властивості ідеальних кристалічних систем, трактування яких спрощувалось завдяки трансляційній інваріантності кристалічної ґратки та перетворюванню симетрії відповідної точкової групи кристалу. Таким чином, теорія кристалічних твердих тіл істотно використовує ідеальність їхньої структури й не може без суттєвих змін бути перенесена на неупорядковані системи. До цього часу досягнуто істотного прогресу у розумінні фізики неупорядкованих систем. Установлено закономірності, що є загальними для широкого класу об'єктів і розвиваються відповідні методи теоретичного аналізу. Фізика неупорядкованих систем (і фізика некристалічних напівпровідників як її складова частина) являє собою широку область фізики конденсованого стану із розмаїттям об'єктів, методів і тематики робіт у різних напрямках.

■ 17.1. Модельні уявлення про електронні процеси у некристалічних напівпровідниках

При побудові моделей електронних властивостей некристалічних твердих тіл важливим аспектом є визначення понять порядку та безпорядку, оскільки наявність упорядкованого розташування атомів у кристалічних твердих тілах є характерною ознакою останніх, що відрізняє їх від неупорядкованих систем.

17.1.1. Далекий порядок і безпорядок у конденсованих середовищах

Наявність далекого порядку означає, що між значеннями деякої фізичної величини у довільній точці системи та її (величини) значенням у нескінченно віддаленій точці існує кореляція. Різні типи далекого порядку у природі виявляються як структурний, магнітний, надпровідний тощо порядки. Найбільш широко виявляється ідея далекого порядку у кристалічному твердому тілі. У цьому випадку існує кореляція густини. У рідині рух атомів спричиняє флуктуацію густини, яка руйнує далекодіючі кореляції. У кристалах атоми прив'язані до вузлів, що утворюють кристалічну ґратку, і кореляції типу **густина–густина** у двох різних точках не зникають. Кристали характеризуються як трансляційним, так й орієнтаційним далеким порядком.

Для визначення позиційного атомного порядку зручно ввести функцію, що описує кореляцію величини, яка характеризує порядок у системі – кореляційну функцію **густина–густина**

$$G(\mathbf{r}) = \langle \rho(0)\rho(\mathbf{r}) \rangle, \quad (17.1)$$

де $\langle \dots \rangle$ – термодинамічне усереднення. Таким чином, завданням є визначення рівноважного значення кореляційної функції $G(\mathbf{r})$ за заданої температури. У **кристалах**, де існує далекий порядок, кореляційна функція не занулюється на будь-яких масштабах. Якщо кореляційна функція наближується до нуля при $|\mathbf{r}| \rightarrow \infty$, то далекого порядку у системі немає та можливий лише близький порядок у розташуванні атомів. Така ситуація реалізується в аморфних системах, стеклах і сильно легованих матеріалах. У них позиційний порядок зберігається тільки в межах одного чи кількох відстаней між атомами, і кореляційна функція спадає з відстанню експоненційно, тобто

$$G(\mathbf{r}) \sim \exp(-|\mathbf{r}|/\xi), \quad (17.2)$$

де ξ – кореляційна довжина або відстань, на якій ще існує порядок. У деяких випадках кореляційна функція спадає за ступеневим законом

$$G(\mathbf{r}) \sim |\mathbf{r}|^{-\alpha}, \quad (17.3)$$

що відповідає квазідалекому порядку. Такого типу порядок може реалізуватись у двовимірних структурах – плівках, надпровідних шарах тощо.

Неповнота порядку характеризує безпорядок у системі, який завжди порушує ту чи іншу симетрію та може реалізуватись різноманітними шляхами.

Будуючи моделі неупорядкованих систем, не можна безпосередньо користуватись математичним апаратом теорії, що розроблений для кристалів. Однак існують деякі загальні характеристики, які можна використати для будь-якої конденсованої системи. Такою характеристикою є, наприклад щільність станів $N(E)$. Коли йдеться про структурний безпорядок, необхідно, перед усім зазначити, що завжди існує відносно жорсткий каркас, що складається з атомів або іонів. У довершеному кристалі таким каркасом є ґратка, у вузлах якої розташовані атоми, на фоні яких реалізується динаміка більш швидких ступенів свободи – електронних збуджень, коливань атомів у вузлах каркасу тощо. Каркас не обов'язково має бути жорстко фіксований, але час його перебудови має бути великим порівняно із характерним часом швидких процесів. У рідині миттєвий каркас також може задовольняти цій вимозі.

Ідеальний кристал, що складається у загальному випадку з атомів різних сортів, характеризується як геометричним упорядкуванням у всіх точках каркасу, наприклад трансляційною симетрією, так і регулярністю розташування атомів різних сортів – композиційним порядком. Невпорядковані системи залежно від виду та ступеню безпорядку у побудові їх каркасу зазвичай поділяються на два класи. Одні з них характеризуються відсутністю регулярності у розташуванні атомів різних сортів. Це, наприклад можуть бути неупорядковані тверді розчини або сплави заміщення, коли система являє собою ідеальну (або майже ідеальну) кристалічну ґратку, а неупорядкованість полягає в тому, що вузли цієї ґратки займають атоми одного із компонентів сплаву з ймовірністю, що визначається концентрацією цього компоненту. Якщо концентрацією речовини, що розчинена у матриці, є f , то ймовірність знайти у вузлі атом розчиненої речовини дорівнює f , а ймовірність знайти у довільному вузлі атом розчинника – $(1-f)$. Це є безпорядком заміщення або композиційним безпорядком (рис. 17.1). Іншим класом неупорядкованих систем є такі системи, в яких відсутня трансляційна симетрія каркасу, тобто далекий порядок у розташуванні атомів відсутній і може існувати лише ближній порядок. Такий тип безпорядку називається топологічним безпорядком. Він характерний для аморфних, рідких і газоподібних

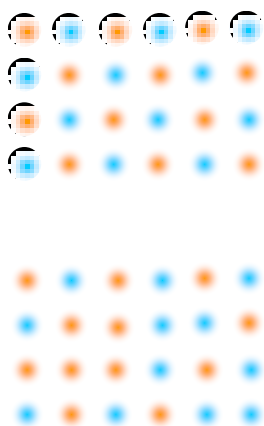


Рис. 17.1. Композиційний безпорядок у сплаві заміщення системи:
а – упорядкованої,
б – неупорядкованої

середовищ (рис. 17.2). У твердих тілах такий безпорядок може бути реалізованим за рахунок сильних спотворень, що обумовлені різноманітними дефектами – точковими, дислокаційними петлями, міжзеренними границями тощо, що порушують топологію структури. Структурний безпорядок може співіснувати з топологічним. Наприклад, в аморфних сплавах та стеклах, що отримуються шляхом швидкого охолодження з розплаву або при розпиленні на підкладку виникає структурний безпорядок. Для таких систем використовується модель випадкової щільної упаковки куль.

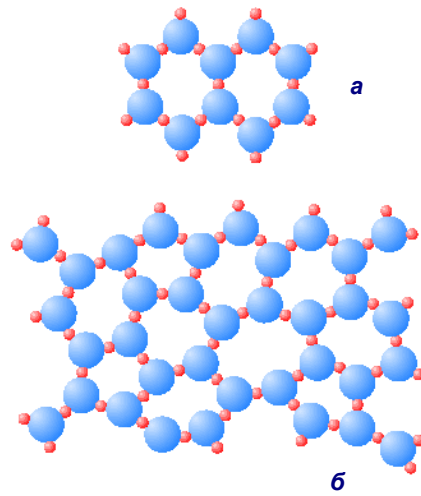


Рис. 17.2. Топологічний безпорядок у двовимірному кремнеземі SiO_2 : а – упорядкований стан (кварц), б – неупорядкований стан (скло)

17.1.2. Електропровідність у некристалічному середовищі. Формула Кубо-Грінеуда.

Як відомо (див. розд. 1-2), у кристалах електрони описує блохівська хвильова функція

$$\psi(\mathbf{r}) = u(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (17.4)$$

У некристалічних речовинах хвильова функція електрона не є у загальному випадку блохівською. Тим не менш, розв'язок рівняння Шредингера має існувати, тому електронний стан принаймні може характеризуватись визначеною енергією: властивості системи може описувати щільність станів $N(E)$. Величина $N(E)$ визначається таким чином, щоб за температури T число електронів в інтервалі енергій dE для кожного із напрямків спінів дорівнювало $N(E)f(E)dE$, де

$$f(e) = \left\{ 1 + \exp\left(\frac{E - E_F}{kT}\right) \right\}^{-1} \quad (17.5)$$

– функція розподілу Фермі. При розрахунках щільності станів і відповідних хвильових функцій необхідно розрізняти два граничних випадки – у першому, наближення вільних носіїв є прийнятним та еле-

ктрони розсіюються слабо. При цьому можна ввести поняття ефективної маси і зв'язок між хвильовим вектором та енергією має звичайний характер квадратичної функції $E = \hbar^2 k^2 / 2m^*$, а щільність станів – $N(E) \sim \sqrt{E - E_C}$. Доки середня довжина вільного пробігу L має кінцеву величину та виконується умова $kL \gg 1$, вищезгаданий граничний випадок є припустимим. Коли L зменшується настільки, що $kL \sim 1$, зміни у щільності станів набувають незворотного характеру і користуватись квадратичним законом дисперсії неможливо.

Доки у некристалічних речовинах розупорядкування не приведе до подібного зменшення довжини вільного пробігу, що $kL \sim 1$, можуть існувати такі ситуації:

- ♦ **розсіювання на кожному атомі є незначним.** Тоді хвильовий вектор k є прийнятним квантовим числом, і невизначеність хвильового вектора Δk є такою, що $\Delta k / k \ll 1$. У цьому випадку ізоенергетичні поверхні є сферичними, і можна використовувати квадратичне наближення для закону дисперсії електронів. Така ситуація реалізується у рідких металах для енергії Фермі та в деякому околі поблизу енергії Фермі;

- ♦ **розсіювання на кожному атомі є значним.** Тоді $\Delta k / k \sim 1$ і хвильовий вектор не є прийнятним квантовим числом. У цьому випадку, як можна довести, довжина вільного пробігу $L \sim 1/k$ і не може бути меншою за $1/k$. Щільність станів при цьому значно відрізняється від тієї, що дає наближення квадратичного закону дисперсії;

- ♦ якщо взаємодія електронів із розсіювачами стає більшою, виникає нове явище, відсутнє у кристалах: **за заданої енергії хвильові функції електронів стають локалізованими.** Це означає, що кожна хвильова функція ψ_E зосереджена у малій області простору, спадаючи експоненційно з відстанню як $\exp(-\alpha R)$. Енергетичні рівні при цьому квантовані, а щільність станів обмежена та неперервна. При цьому стани локалізовані, хоча хвильові функції можуть перекриватись. Цей ефект називається локалізацією Андерсона і буде детальніше розглянутий далі.

Якщо за заданої енергії деякі стани локалізовані, то всі стани із цією енергією мають бути локалізовані. Експериментальним доведенням того, що локалізовані та нелокалізовані стани не співіснують за заданої енергії є факт, що у домішковій зоні кристалічного кремнію, утвореній за легування його фосфором, ширина лінії комбінаційного розсіювання світла, пов'язана із переходами електронів у зоні провідності, розширюється завдяки доплерівському зсуву. При цьому лінія, що визначається переходами із домішкової зони, не розширюється. Екс-

периментально показано, що розширена та нерозширена лінії не співіснують за однакової концентрації донорів.

У будь-якій некристалічній системі можлива велика кількість конфігурацій атомів, що відповідають одній енергії. Таку сукупність конфігурацій називають ансамблем. Це означає, що будь-яка розрахункова величина, яку необхідно порівняти з експериментом, має бути усереднена за всіма конфігураціями ансамблю. Серед них мають бути й такі, для яких стани нелокалізовані, наприклад конфігурації у вигляді кристалів. Якщо існує така провідність σ_E , що для всіх заповнених станів ($E \leq E_F$, E_F належить до області енергій, де стани локалізовані) $\sigma_E \rightarrow 0$ при $T \rightarrow 0$, то усереднення за всіма конфігураціями системи величини σ_E дасть нуль: $\langle \sigma_E \rangle = 0$. У цьому сенсі найбільш задовільним визначенням локалізації хвильових функцій з енергією E буде занулення усередненого значення величини σ_E , тобто

$$\langle \sigma_E \rangle = 0. \quad (17.6)$$

Для розрахунку σ_E розглядається взаємодія електромагнітної хвилі низької частоти ω із середовищем, де всі стани заповнені до деякого граничного значення енергії E . Тоді провідність на постійному струмі при $T = 0$ визначається як

$$\sigma_E = \sigma_E(0) = \lim_{\omega \rightarrow 0} \sigma_E(\omega). \quad (17.7)$$

У неупорядкованому напівпровіднику провідність за ненульової температури та нехтування взаємодією із коливаннями каркасу (із фононами) визначається через σ_E

$$\sigma = - \int \sigma_E \frac{\partial f}{\partial E} dE. \quad (17.8)$$

Вираз для σ_E , що можна отримати за допомогою (17.7), називається формулою Кубо-Грінвуда (її виведення див. у додатку J)

$$\sigma_E(0) = \frac{2\pi e^2 \hbar^3 \Omega}{m^2} \langle |D_E|^2 \rangle_E N^2(E), \quad (17.9)$$

де

$$D_E = \int \psi_{E'}^* \frac{\partial \psi_E}{\partial x} d^3x,$$

Ω – об'єм системи. $\langle \dots \rangle_E$ – усереднення за всіма станами з енергією E та за всіма станами з енергією E' із подальшим припущенням, що $E = E'$.

17.1.3. Локалізація Андерсона

Андерсонівською локалізацією називають явище, що виникає за розповсюдження хвиль у просторово-неоднорідному середовищі: завдяки багаторазовому розсіюванню на неоднорідностях та інтерференції розсіяних хвиль стає неможливим розповсюдження біжучих хвиль. Коливання набувають характеру стоячої хвилі, що локалізована в обмеженій області простору. Андерсонівська локалізація можлива для хвиль довільної природи, але особливо яскраво вона виявляється у випадку хвиль де Бройля для частинок при вивченні кінетичних властивостей, наприклад електропровідності та теплопровідності неупорядкованих твердих тіл (аморфні та сильно леговані напівпровідники), оскільки за наявності ефекту андерсонівської локалізації рухливість частинок є нульовою. Експериментальним виявом локалізації електронних станів є рівність нулю статичної провідності в умовах низьких температур. Якщо рівень Фермі лежить в області локалізованих станів, то провідність σ_T за кінцевих температур має бути активованою, що автоматично означає існування границі $\lim_{T \rightarrow 0} \sigma_T = 0$.

Явище андерсонівської локалізації найбільш інтенсивно вивчалось на модельному гамільтоніані такого типу

$$H = \sum_j \varepsilon_j |j\rangle\langle j| + \sum_{j'} V_{jj'} |j\rangle\langle j'|, \quad (17.11)$$

де ε_j - енергія j -го незбуреного стану, $V_{jj'}$ - збурення, що виникає завдяки неперіодичному потенціалу. У другому доданку сума береться за найближчими сусідами. Незважаючи на те, що андерсоновська локалізація є властивістю найпростішого гамільтоніана електронів, що майже не взаємодіють, виявилось, що вона являє собою надзвичайно складне явище. Андерсон, очевидно, був першим, хто усвідомив, що гамільтоніан за наявності безпорядку може характеризуватись подовженими або локалізованими станами залежно від величини розупорядкування. Розглянемо виникнення локалізованих станів у неупорядкованій системі, що складається з періодично повторюваних у просторі потенціальних ям різної глибини.

Як відомо, розв'язком рівняння Шредингера є

$$\nabla^2 \psi + \frac{2m}{\hbar^2} (E - V) \psi = 0 \quad (17.12)$$

для періодичного потенціалу V (див., напр. задачу Кроніга-Пенні) може бути утворення вузької енергетичної зони (рис. 17.3).

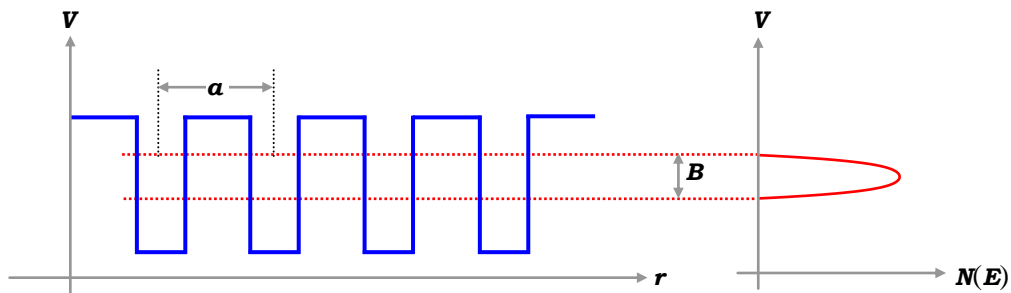


Рис. 17.3. Потенціальна енергія та щільність станів у регулярній структурі

Припустимо тепер, що ями розташовані досить далеко одна від одної так, що перекриття між атомними хвильовими функціями $\phi(\mathbf{r})$ на сусідніх вузлах є незначним. Нехай \mathbf{R}_n – координата n -ї ями. Тоді блохвиську хвильову функцію електрона у кристалі можна записати як

$$\psi_k(x, y, z) = \sum_n e^{i\mathbf{k}\cdot\mathbf{R}_n} \phi(\mathbf{r} - \mathbf{R}_n). \quad (17.13)$$

Функції ϕ вважатимемо сферично-симетричними. Якщо тепер W_0 – енергія рівня в ізолюваній ямі, то енергії електрона у простій кубічній ґратці, що відповідають хвильовим функціям (17.13),

$$E = W_0 + W_k, \quad (17.14)$$

з

$$W_k = -2I \sum_{j=x,y,z} \cos k_j a, \quad (17.15)$$

де інтеграл перекриття

$$I = \int d\mathbf{r} \phi^*(\mathbf{r} - \mathbf{R}_n) H \phi(\mathbf{r} - \mathbf{R}_{n+1}), \quad (17.16)$$

H – гамільтоніан системи. Інтеграл перекриття залежить від вигляду потенціалу, що формує ями і майже завжди може бути представлений у вигляді

$$I = I_0 e^{-\alpha R}. \quad (17.17)$$

При цьому параметр α визначається таким чином, що $\exp(-\alpha r)$ дає швидкість спадання хвильової функції ізолюваної ями із відстанню

$$\alpha = \sqrt{\frac{2mW_0}{\hbar^2}}. \quad (17.18)$$

Для воднеподібних атомних хвильових функцій оцінка передекспоненційного фактора I_0 дає

$$I_0 = \alpha e^2 [3(1 + \alpha R)/2 + (\alpha R)^2 / 6]. \quad (17.19)$$

При цьому ефективна маса електрона на дні зони провідності – $m^* = \hbar^2 / (2I\alpha^2)$, а ширина зони – $B = 2zI$, де z – координаційне число.

Що відбудеться з енергетичною зоною, якщо потенціал стане неперіодичним? Необхідно пам'ятати, що неперіодичний потенціал можна отримати із періодичного у два способи:

- ♦ зсуваючи кожний центр потенціальної ями, що утворюють регулярну структуру, на випадкову відстань;
- ♦ додаючи випадковий потенціал $<V/2$ до кожної ями.

Зрозуміло, що комбінація цих методів також дасть неупорядковану систему. Але для якісного розуміння ефекту андерсонівської локалізації можна користуватись найпростішим методом побудови неупорядкованої системи. Андерсон припускав, що V набуває будь-яких значень з інтервалу $\pm(1/2)V_0$ таким чином, що V_0 являє собою область розмиття енергій (рис. 17.4). Якщо величина V_0 є малою, то довжина вільного пробігу буде великою. Борнівське наближення дає, що

$$\frac{1}{L} = \frac{\pi}{\hbar} \left(\frac{V_0}{2} \right)^2 a^3 \frac{N(E)}{u}, \quad (17.20)$$

де значення швидкості u береться за енергії $E = E_F$. У випадку, що розглядається, можна вважати: щільність станів незначно відрізняється від такої у впорядкованій структурі, тобто має вигляд

$$N(E) = \frac{1}{\pi^2 \hbar^3} \left(\frac{m^3 E}{2} \right)^{1/2}, \quad (17.21)$$

та із (17.20) отримаємо

$$\frac{a}{L} = \frac{1}{32\pi} \left(\frac{V_0}{I} \right)^2. \quad (17.22)$$

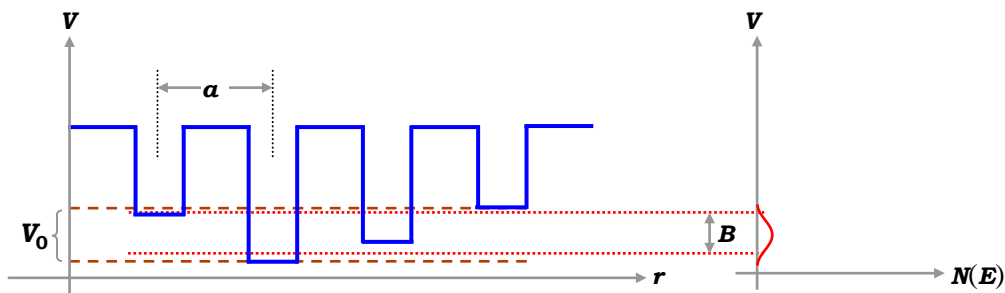


Рис. 17.4. Модельний потенціал неупорядкованої структури

Згідно із правилом Іоффе-Ригеля у неупорядкованих системах неможливою є така довжина вільного пробігу, що $kL < 1$. Це правило означає: найменша величина вільного пробігу реалізується у випадку, коли хвильова функція "втрачає пам'ять" про свою фазу на відстанях, що дорівнює міжатомним. Тоді довжину вільного пробігу можна оцінити як $L \approx a$, що означає згідно із (17.22)

$$\frac{V_0}{I} \approx \sqrt{32\pi}. \quad (17.23)$$

У випадку, коли ширина зони $B = 2zI$ і координаційне число $z = 6$, із (17.23) маємо, що втрата пам'яті фази хвильової функції на відстанях порядку міжатомних настає, якщо область розмиття потенціалу $V_0 \approx 0.83B$.

Для розуміння того, що відбувається, коли V_0 перевищить величину (17.23), слідуючи за Андерсоном, розглянемо дві потенціальні ями – a та b , розташовані на відстані R одна від одної. Енергії ям відрізняються від середньої на величини V_a і V_b , відповідно. Хвильові функції двох електронів в ямах мають вигляд

$$\begin{aligned} \psi_1 &= A\phi_a + B\phi_b, \\ \psi_2 &= B\phi_a - A\phi_b, \end{aligned} \quad (17.24)$$

де ϕ_a і ϕ_b – атомні хвильові функції. Коефіцієнти A , B та енергії E_1 , E_2 можна знайти за допомогою мінімізації інтеграла енергії. Не вдаючись до деталей отримання громіздких результатів такої мінімізації, розглянемо якісно деякі граничні випадки:

- ♦ I – якщо $|V_a - V_b| \ll I$, то $A \approx B$, $E_1 - E_2 \approx 2I$ і різниця енергій не може бути меншою за $2I$;
- ♦ II – якщо $|V_a - V_b| \gg I$, то $A/B \approx |V_a - V_b|/I$.

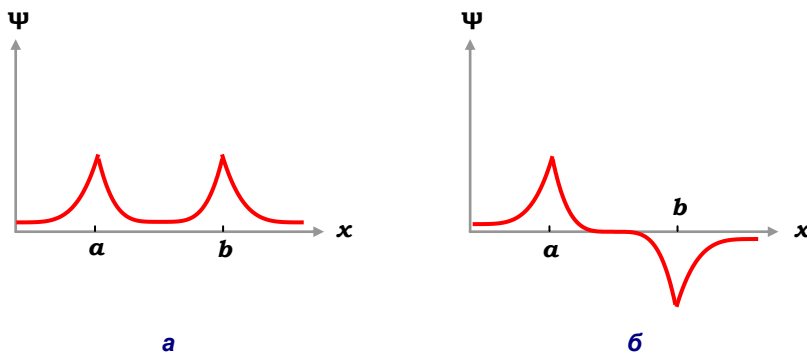


Рис. 17.5. Хвильова функція двох однакових потенціальних ям для станів: а – симетричних; б – антисиметричних

Хвильові функції для однакових ям (випадок I) подано на рис. 17.5, а хвильові функції для випадку II – на рис. 17.6.

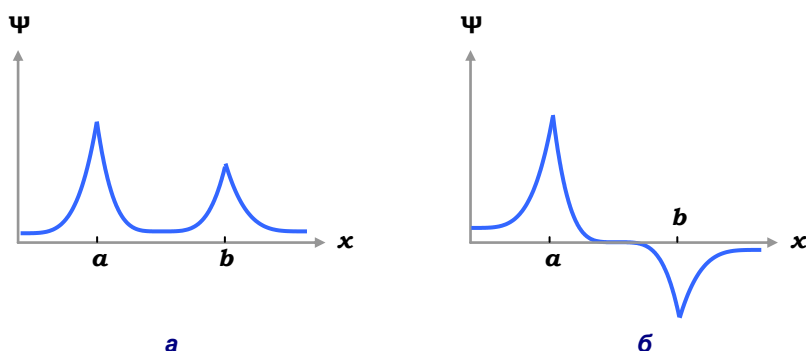


Рис. 17.6. Хвильова функція двох неоднакових потенціальних ям для станів: а – симетричних; б – антисиметричних

Уявимо тепер, що система складається із великої кількості ям. Тоді хвильова функція системи являтиме суперпозицію хвильових функцій окремих ям

$$\psi_i = \sum_j C_j^i \phi_j, \quad (17.25)$$

причому у результаті того, що коефіцієнти C_j^i , які характеризують внесок до i -го стану системи від j -ї ями, є різними, огинаюча матиме максимум на ямі із більшою глибиною. У результаті виникатиме ситуація, коли ймовірність знайти частинку локалізованою на цій ямі є набагато більшою, ніж в іншому місці системи.

Якщо кількість ям збільшувати до нескінченності, то при переході від однієї ями до іншої виникають флуктуації амплітуди та фази хвильової функції. Флуктуації збільшуються зі зростанням безпорядку у системі, тобто величини V_0/B . Якщо відношення V_0/B дуже велике, то хвильові функції кожної окремої ями будуть слабко збурені всіма іншими ямами та експоненційно спадатимуть із відстанню (рис. 17.7) – виникатиме локалізація станів. Для з'ясування, за яких значень параметру неупорядкованості V_0/B відбуватиметься така локалізація, Андерсон розглянув потенціал типу поданого на рис. 17.4 і переформулював задачу так: якщо у момент часу $t=0$ електрон локалізований на одній із ям, то що з ним відбуватиметься за $t \rightarrow \infty$? Чи існує при $T=0$ ненульова ймовірність того, що електрон продифундує на велику відстань. Андерсон знайшов, що дифузія не відбувається (відбувається локалізація), якщо $V_0/B \geq C_L$, де стала C_L залежить від координацій-

ного числа z . У цьому випадку хвильова функція електрона буде експоненційно затухати із відстанню r від n -ої ями, тобто хвильові функції локалізованих на ямі j станів матимуть вигляд

$$\psi_j(\mathbf{r}) = \sum_{n \neq j} A_n \phi(\mathbf{r} - \mathbf{R}_n) e^{-\alpha|\mathbf{r}_j - \mathbf{r}|}, \quad (17.26)$$

де коефіцієнти A_n характеризуються випадковими фазами.

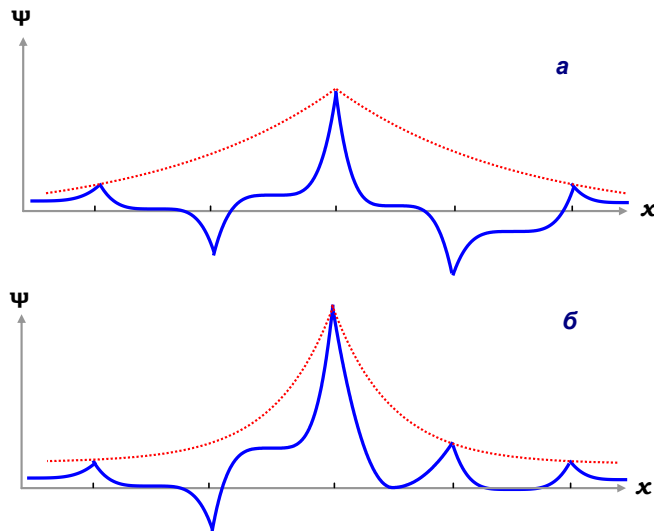


Рис. 17.7. Хвильова функція у моделі Андерсена:
а – локалізований стан поблизу переходу Андерсона (слабка локалізація);
б – сильна локалізація

За координаційного числа $z = 6$ Андерсон отримав, що $C_L = 5.5$. Таким чином, завдяки можливості андерсонівської локалізації у неупорядкованих системах електронний транспорт вже не визначатиметься процесами розсіювання на недосконалостях ґратки, а має формуватись іншими процесами. Одним із механізмів формування електронного транспорту у неупорядкованих системах завдяки наявності локалізованих станів є стрибковий механізм.

17.1.4. Стрибкова провідність

За ненульових температур локалізовані стани дають внесок до процесів перенесення, який називається стрибковою провідністю, оскільки він виникає як результат переміщення носіїв заряду з одного локального центру на інший (стрибків з однієї потенціальної ями до

іншої). Важливим аспектом концепції стрибкової провідності є припущення, що всі центри локалізації характеризуються різними енергіями, тобто два центри з однаковою енергією локалізації розташовані на нескінченній відстані один від одного. Оскільки одним із прикладів неупорядкованих систем є домішкові напівпровідники, характерні властивості стрибкової провідності розглянемо, виходячи із моделі домішкової зони за слабкого легування.

Нехай домішкові центри характеризуються координатами \mathbf{r}_i , а хвильові функції локалізованих на них електронів згасають на нескінченності згідно із моделлю воднеподібної домішки

$$\psi(\mathbf{r} - \mathbf{r}_i) \propto \left(\frac{\varepsilon \hbar^2}{m^* e^2} \right)^{-3/2} e^{-|\mathbf{r} - \mathbf{r}_i|/a_B}, \quad (17.27)$$

де введено перенормований борівський радіус зонного електрона $a_B = \varepsilon \hbar^2 / m^* e^2$ із діелектричною проникністю напівпровідника ε та ефективною масою m^* ¹. За невеликої концентрації домішок N ($a_B^3 N \ll 1$) домішкові стани можна вважати локалізованими. Але завдяки експоненційним хвостам хвильових функцій перекривання хвильових функцій сусідніх центрів є помітним, і за мірою цього перекривання існує ймовірність переходу (стрибка) електрона з одного центра на інший

$$\frac{1}{\tau_{ij}} \propto F(\varphi_{ij}, f_i, f_j) \int d^3q |M_q|^2 \delta(\hbar q s - \Delta_{ij}), \quad (17.28)$$

У цій формулі s – швидкість звука, а δ -функція забезпечує закон збереження енергії і відбирає фонони, що випромінюються або поглинаються електроном, для компенсування різниці енергій $\Delta_{ij} = \varepsilon_i - \varepsilon_j$ початкового j та кінцевого i станів. Функція F описує статистичні умови переходу з j -го до i -го станів і пропорційна ймовірності того, що початковий стан заповнений, а кінцевий – незаповнений, тобто добутку функції розподілу Фермі початкового стану $f_j = [\exp(\varepsilon_j - \mu) / kT + 1]^{-1}$ і кінцевого стану $(1 - f_i)$, а також залежить від функції розподілу фононів за енергіями Δ_{ij} : $\varphi_{ij} = [\exp(\Delta_{ij} / kT) - 1]^{-1}$. Для якісних висновків можна користуватись оцінкою цієї функції

¹ Зрозуміло, що ефективну масу можна ввести лише для кристалічного напівпровідника. Введення ефективної маси у (17.27) означає, що розглядається кристалічний стан напівпровідника із достатньо великою концентрацією домішкових центрів, але такою, що дозволяє розглядати систему як кристалічну.

$F \approx \exp(-\varepsilon_{ij}/kT)$ із деякою характеристичною енергією ε_{ij} . У першому наближенні інтеграл у (17.28) оцінюється як експонента $\exp(-2r_{ij}/a_B)$, де r_{ij} – відстань між двома домішковими центрами, на яких може локалізуватися електрон.

Таким чином, ймовірність переходу $1/\tau_{ij}$ між центрами j та i пропорційна добутку двох експонент. Якщо з'єднати кожен пару домішкових центрів уявними опорами R_{ij} , пропорційними τ_{ij} ,

$$R_{ij} = R_0 \exp\left[(2r_{ij}/a_B) + (\varepsilon_{ij}/kT)\right], \quad (17.29)$$

то можна отримати мережу випадкових опорів, що моделюють ізолятор. Така мережа називається мережею Абрахамса-Міллера (рис. 17.8).

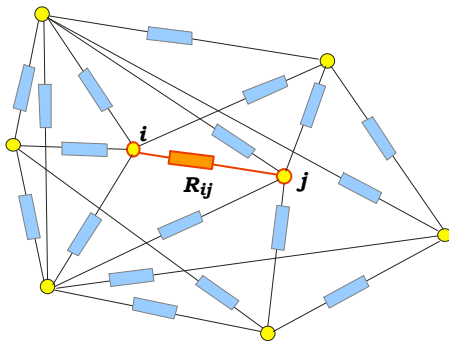


Рис. 17.8. Мережа Абрахамса-Міллера

Згідно із принципом детальної рівноваги, в умовах термодинамічної рівноваги електронні переходи між вузлами мережі відбуваються з однаковою ймовірністю в обидва боки. Під дією зовнішнього електричного поля з'являється направлений потік електронів, тобто провідність. Для її обчислення необхідно розв'язати задачу про провідність мережі Абрахамса-Міллера.

17.1.4.1. Переходи між найближчими центрами

Найбільш простим видом стрибкової провідності є провідність, що визначається переходами між найближчими центрами. Щільність станів у домішковій зоні за малої компенсації має максимум у точці, що знаходиться поблизу енергії іонізації ізолизованого донора E_D . Зрозуміло, що найбільш ймовірно стрибки відбуватимуться, якщо початковий і кінцевий стани мають енергії поблизу максимуму щільності станів. Для того, щоб стрибок міг бути реалізований, необхідно, щоб кінцевий стан був незайнятий. Ймовірність такої події залежить від відстані енергії стану до рівня Фермі μ і пропорційна експоненті $K = e^{-|\mu - E_D|/kT}$. Це й є множником, що визначає функцію F у (17.28). Фактор $\exp(-\Delta_{ij}/kT)$, що визначає кількість фононів, які беруть участь

у переході завдяки нерівності $\Delta_{ij} \ll |\mu - E_D|$, є набагато більшим за K . Якщо перехід відбувається за такої температури, що K є малою, але не зникаючою величиною, то $\exp(-\Delta_{ij}/kT) \approx 1$. Таким чином, величина $\xi_{ij} = (2r_{ij}/a_B) + (\varepsilon_{ij}/kT)$, що визначає опір між сусідніми вузлами мережі Абрахамса-Міллера, для всіх стрибків являє собою відстань від максимуму щільності станів до рівня Фермі $\varepsilon_{ij} = |\mu - E_D|$.

Оскільки фактор K входить до усіх ξ_{ij} , він не дає внеску до поширення R_{ij} та спричиняє просте перенормування передекспоненційного фактору у (17.29). Таким чином, за подальшого аналізу властивостей мережі визначатимемо опір між вузлами i та j простішим виразом

$$R_{ij} = R_0 \exp(2r_{ij}/a_B), \quad (17.30)$$

який, між іншим, показує: чим ближче розташовані вузли, тим меншим буде опір між ними у модельній мережі.

Таким чином, задача звелась до визначення перколяційного радіуса r_C у системі випадково розташованих із концентрацією N вузлів. Теорія перколяції стверджує, що

$$\frac{4\pi}{3} r_C^3 N = 2.7. \quad (17.31)$$

Звідси випливає, що $r_C = 0.865N^{-1/3}$. Із врахуванням цього (17.30) дає

$$R = R_0 e^{1.73/(a_B N^{1/3})}. \quad (17.32)$$

Експериментально стрибкову провідність можна спостерігати при вивченні температурної залежності опору легованого напівпровідника. Для цього зручно використовувати трансмутований матеріал. Це дозволяє, по-перше, забезпечити введення домішок у широкому діапазоні концентрацій (майже 3 порядки) за відсутності кореляцій в їхньому розташуванні за об'ємом зразка; по-друге, при цьому можна спостерігати гігантські зміни вимірюваного опору (аж до 12 порядків).

Опромінювання чистого германію потоком нейтронів приводить до того, що у результаті ядерних реакцій із ядер одних ізотопів Ge з'являються ядра Ga, що відіграють роль мілких акцепторів, а з ядер інших ізотопів Ge – ядра As, що є мілкими донорами. Відношення кількості отриманих домішок визначається перерізом відповідних ядерних реакцій і відносними концентраціями ізотопів у первісному матеріалі. У результаті опромінювання природного Ge виникає германій p -типу із коефіцієнтом компенсації $N_A/N_D = 0.4$. При цьому концентрації ак-

цепторних N_A і донорних N_D домішок і, як результат, концентрація носіїв заряду $p = N_A - N_D$ залежать тільки від часу опромінювання.

Для опису сукупності кривих необхідно припустити наявність двох паралельних провідних каналів, провідність кожного з яких залежить від температури експоненційно

$$\sigma = \sigma_b + \sigma_h = \sigma_{b0} e^{-\varepsilon_b/kT} + \sigma_{h0} e^{-\varepsilon_h/kT}, \quad \varepsilon_b \gg \varepsilon_h. \quad (17.33)$$

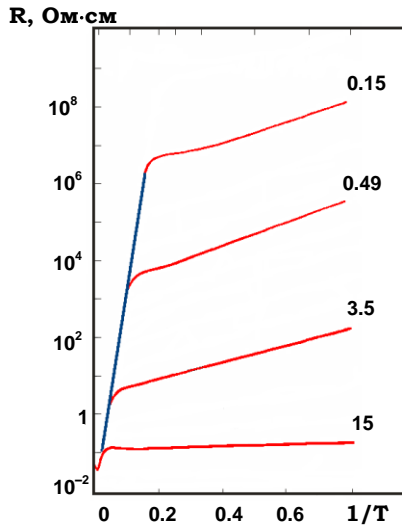


Рис. 17.9. Залежність опору трансмутованого р-германію від температури для різних ступенів легування N_A (числа над правою частиною кривих, $\times 10^{16} \text{ см}^{-3}$)

механізмом формування провідності стають стрибки електронів з одного центра на інший без участі делокалізованих (зонних) носіїв. Температура переходу від одного механізму провідності до іншого (коли σ_b та σ_h зрівнюються) буде тим вищою, чим більшою є концентрація локальних центрів N .

Величина σ_{h0} суттєво залежить від концентрації домішкових центрів. Припустимо, що ця залежність описується експонентою, тому

$$\sigma_h(T) = \sigma_{h0}^* e^{-f(N)/kT} e^{-\varepsilon_h/kT}. \quad (17.35)$$

Для визначення функції $f(N)$ звернемося до (17.32). У результаті, беручи до уваги, що $\varepsilon_h = |\mu - E_D|$, отримуємо

$$\sigma_h(T) = \sigma_{h0}^* e^{-1.73/(a_B N^{1/3})} e^{-|\mu - E_D|/kT}. \quad (17.36)$$

Перший канал – це звичайна провідність зонних електронів, обумовлена тепловим закиданням електронів із домішкових рівнів до зони провідності. Цей канал домінує за більш високих температур, коли опір виходить на загальну граничну криву (синій колір) у лівій частині графіку (рис. 17.9). Експоненційне зростання опору вздовж прямої

$$R = R_{b0} \exp(-\varepsilon_b/kT) \quad (17.34)$$

пов'язане із виморожуванням носіїв у цьому каналі за низьких температур. У (17.34) ε_b – середня енергія іонізації домішки. Зростання опору припиняється, коли основним механізмом формування провідності стають стрибки електронів з одного центра на інший без участі делокалізованих (зонних) носіїв.

Звідси маємо, що опір, який визначається стрибковим механізмом, пов'язаний із концентрацією домішкових центрів співвідношенням

$$\ln R \propto N^{-1/3}. \quad (17.37)$$

Залежність опору від концентрації локальних центрів ілюструється експериментальними даними, які подані на рис. 17.10. Значення числового коефіцієнту у показнику першої експоненти у (17.36) становлять: для p -Ge – 1.75; p -Si – 1.8; n -InP – 1.9; n -GaAs – 1.7. Ці коефіцієнти для різних матеріалів близькі до теоретичного значення теорії перколяції – 1.73. Експеримент підтверджує, що $\ln(R/R_0) \propto 1/T$, оскільки праві частини графіків на рисунку є прямими лініями.

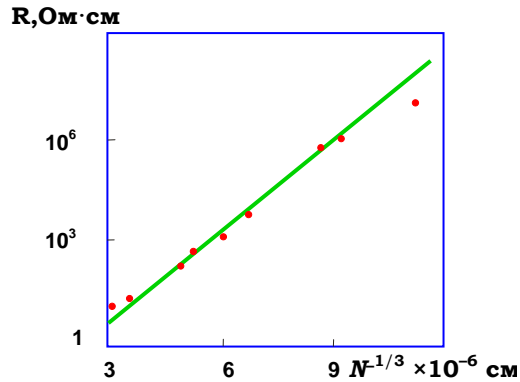


Рис. 17.10. Залежність опору в Ge, легovanому Ga від концентрації домішкових центрів

17.1.4.2. Переходи зі змінною довжиною стрибка. Закон Мотта

Умовою реалізації провідності через стрибки на найближчі центри є наявність великої кількості пар близьких сусідів, в яких один центр незайнятий. Якщо знизити температуру так, щоб $T \ll \mu - E_D$, то серед найближчих сусідів, більшість з яких характеризуються енергією E_D , виявиться дуже мало вільних центрів, і стрибки на найближчі центри вимерзнуть. Тоді провідність забезпечуватимуть стрибки між центрами з енергіями із деякого околу ε Фермі-рівня, серед яких, наперед можна стверджувати, є незайняті. При цьому залишиться з'ясувати наскільки близькими виявляться такі центри один до одного.

Розглянемо деякий окіл Фермі-рівня $\mu \pm \varepsilon$, в якому вважатимемо щільність станів сталою, наприклад $g = g_\mu$. Тоді число станів у цьому околі дорівнюватиме $N(\varepsilon) = g_\mu \varepsilon$, а середня відстань між центрами $\bar{r}_{ij} \approx [N(\varepsilon)]^{-1/3}$. Різниця енергій між такими локалізованими станами $\sim \varepsilon$. У цьому випадку основний статистичний фактор F у (17.28) визначається числом фононів, що забезпечують виконання закону збереження енергії. Оскільки фонон має у середньому компенсувати дефіцит енергії порядку ширини енергетичного околу поблизу Фермі-рівня, то

основним множником статистичного фактору F буде $\exp(-\varepsilon/kT)$. Для подальшого аналізу залишимо у випадковій мережі Абрахамса-Міллера лише ті вузли, енергія яких попадає до інтервалу $\mu \pm \varepsilon$. Сусіди у такій підмережі з'єднуються опорами

$$R_{ij} = R_0 \exp(\xi_{ij}), \quad (17.38)$$

де
$$\xi_{ij} = \frac{2}{a_B [N(\varepsilon)]^{1/3}} + \frac{\varepsilon}{T} = \frac{2}{g_\mu^{1/3} a_B \varepsilon^{1/3}}. \quad (17.39)$$

Щільність утвореної таким чином підмережі залежить від ширини енергетичної смуги із серединою на рівні Фермі, тобто від величини ε . Визначимо величину ε з умови мінімуму ξ_{ij} : $d\xi_{ij}(\varepsilon)/d\varepsilon = 0$. Підставляючи сюди (17.39), отримуємо

$$-\frac{1}{3} \frac{2}{g_\mu^{1/3} a_B} \varepsilon^{-4/3} + \frac{1}{T} = 0, \quad (17.40)$$

звідки
$$\varepsilon_{\min} = (T_M T^3)^{1/4}, \quad (17.41)$$

де
$$T_M = \left(\frac{2}{3}\right)^3 \frac{1}{g_\mu a_B^3}. \quad (17.42)$$

Таким чином, у середньому електрон втрачає (набуває) енергію, перестрибуючи із вузла на вузол розглянутої підмережі $\varepsilon \approx \varepsilon_{\min}$. При цьому середня довжина стрибка

$$\bar{r}_{ij} \cong \left(\frac{3 a_B}{2 g_\mu}\right)^{1/4} \frac{1}{T^{1/4}} \text{ або } \bar{r}_{ij} \cong \frac{3}{2} a_B \left(\frac{T_M}{T}\right)^{1/4}. \quad (17.43)$$

Зауважимо, що ймовірність стрибків на меншу відстань пропорційна фактору $\exp(-\Delta_{ij}/kT)$, тобто експоненційно зменшується порівняно зі стрибками на відстань \bar{r}_{ij} . Ймовірність стрибків на більшій відстані також зменшується пропорційно $\exp(-2r_{ij}/a_B)$. Таким чином, оскільки $r_{ij} \gg N^{-1/3}$, усередині кореляційного радіусу $r_C = \bar{r}_{ij}$ статистично вірогідно знайти центр з енергією із інтервалу ε . Це означає, що мережа таких центрів за будь-якої температури являє собою нескінченний кластер, стрибки між вузлами всередині якого визначають стрибкову провідність. Іншими словами, залежність опору неупорядкованої напівпровідникової структури за низьких температур, коли відбуваються стрибки зі змінною довжиною, виражається законом Мотта

$$R = R_0 \exp(T_M / T)^{1/4}. \quad (17.44)$$

Величина R_0 залежить від температури ступенєво, але цією залежністю зазвичай нехтують. Зазначимо, що ступінь $1/4$ у показнику експоненти отримують тільки у $3D$ -структурах. Для тонкої плівки можна записати

$$\bar{r}_{ij} \cong [N(\varepsilon)]^{-1/2}, \quad \xi_{ij} \frac{2}{a_B g_\mu^{1/2} \varepsilon^{1/2}}. \quad (17.45)$$

Умова мінімуму параметру ξ_{ij} дає рівняння на визначення критичного значення ширини енергетичної смуги

$$\frac{1}{g_\mu^{1/2} a_B} \frac{1}{\varepsilon^{3/2}} - \frac{1}{T} = 0, \quad (17.46)$$

звідки $\varepsilon_{\min} = (T_M T)^{1/3}, \quad (17.47)$

де критична температура $T_M = \frac{1}{g_\mu a_B^2}. \quad (17.48)$

Таким чином, залежність опору неупорядкованої напівпровідникової тонкої плівки за низьких температур, коли відбуваються стрибки зі змінною довжиною, виражається законом Мотта для $2D$ -систем

$$R = R_0 \exp(T_M / T)^{1/3}. \quad (17.49)$$

Отже, ввівши розмірність системи D , обидві формули – (17.44) і (17.49) – можна записати в універсальному вигляді

$$R = R_0 \exp(T_M / T)^{1/(D+1)}. \quad (17.50)$$

Кулонівська взаємодія електрона, яка перестрибує з одного центра на інший і дірки, що виникає на центрі, покинутому електроном, спричиняє підсилення локалізації і формування щілини на рівні Фермі. Завдяки електрон-електронній взаємодії щільність станів на рівні Фермі обертається у нуль за універсальним законом

$$g(\varepsilon) \sim \left| \frac{k\varepsilon}{e^2} \right|^{D-1}, \quad g(0) = 0, \quad (17.51)$$

де k – діелектрична проникність. Вигляд щільності станів в околі рівня Фермі не залежить ні від характеру взаємодії на малих відстанях, ні від затравочної щільності станів g_0 . Це явище називають *кулонівською*

щільною у щільності станів. У симетричному щодо рівня Фермі μ околі $\mu \pm \varepsilon$ кількість станів залежить від розмірності D системи

$$N(\varepsilon) \propto \left(\frac{\kappa \varepsilon}{e^2} \right)^D. \quad (17.52)$$

При цьому середня відстань між центрами, енергія яких належить інтервалу $\mu \pm \varepsilon$, не залежить від розмірності системи

$$\bar{r}_{ij} \cong [N(\varepsilon)]^{-1/D} \cong \frac{\kappa \varepsilon}{e^2}. \quad (17.53)$$

Аналогічно попередньому випадку сталої щільності станів показник експоненти для величини опору в підмережі Абрахамса-Міллера із величиною смуги розкиду енергії ε

$$\xi_{ij} = \frac{2}{a_B [N(\varepsilon)]^{1/2}} + \frac{\varepsilon}{kT} = \frac{2e^2}{\kappa a_B \varepsilon} + \frac{\varepsilon}{kT} \quad (17.54)$$

досягає мінімального значення за енергії

$$\varepsilon_{\min} = (\kappa^2 T_{ES} T)^{1/2}, \quad (17.55)$$

де

$$T_{ES} \approx \frac{e^2}{\kappa \kappa \varepsilon}.$$

При цьому, незалежно від розмірності системи, її опір змінюється за законом

$$R = R_0 \exp(T / T_{ES})^{1/2}, \quad (17.56)$$

що має назву закону Ефроса-Шкловського.

Зазначимо, що стрибковий механізм провідності реалізується не тільки у легованих напівпровідниках за дуже низьких температур, а також в органічних напівпровідниках.

■ 17.2. Аморфні напівпровідники

Аморфні напівпровідники – напівпровідникові матеріали, що характеризуються наявністю ближнього порядку та відсутністю далекого. Аморфні та склоподібні напівпровідники за складом та структурою належать до халькогенідних, оксидних, органічних і тетраедричних. Халькогенідні склоподібні напівпровідники отримують шляхом охолодження розплаву або випаровуванням у вакуумі. До таких матеріалів належать селен і телур, а також дво- та багатокомпонентні склоподібні сплави халькогенідів – сульфідів, селенідів і телурідів – із

різноманітними металами, наприклад As-S-Se, As-Ge-Se-Te, As-Sb-S-Se, Ge-Pb-Te тощо. Елементарні тетраедричні аморфні Ge та Si отримують найчастіше іонним розпиленням у різноманітних атмосферах, що містять водень, або через дисоціацію газів, зокрема SiH_4 і GeH_4 у високочастотному розряді.

Усі особливості аморфних і склоподібних напівпровідників визначаються особливостями енергетичного спектру електронів, зокрема завдяки ближньому порядку у таких матеріалах спостерігаються області із високою та низькою щільністю електронних станів. Тому можна умовно говорити про зонну структуру аморфних напівпровідників. При цьому розупорядкованість структури спричиняє появу додаткових дозволених електронних станів, щільність яких $g(E)$ спадає вглиб забороненої зони, утворюючи так звані хвости щільності станів (рис. 17.11). Як наслідок порушення далекого порядку різкий перехід між дозволеними та забороненими рівнями стає поступовим із розмитими границями. Деякі із локальних дефектів спричиняють появу всередині забороненої зони дозволених рівнів електронних станів і дуже часто із високою щільністю (рис. 17.11 б). Максимуми $N(E)$, обумовлені дефектами структури, можуть виникати всередині щілини та перекриватись один з одним, як і самі "хвости" (рис. 17.11 в). Унаслідок щільність станів електронів і дірок тільки знижується між зоною провідності та валентною зоною, але не досягає нульового значення, як зазвичай відбувається у забороненій зоні кристалічного напівпровідника.

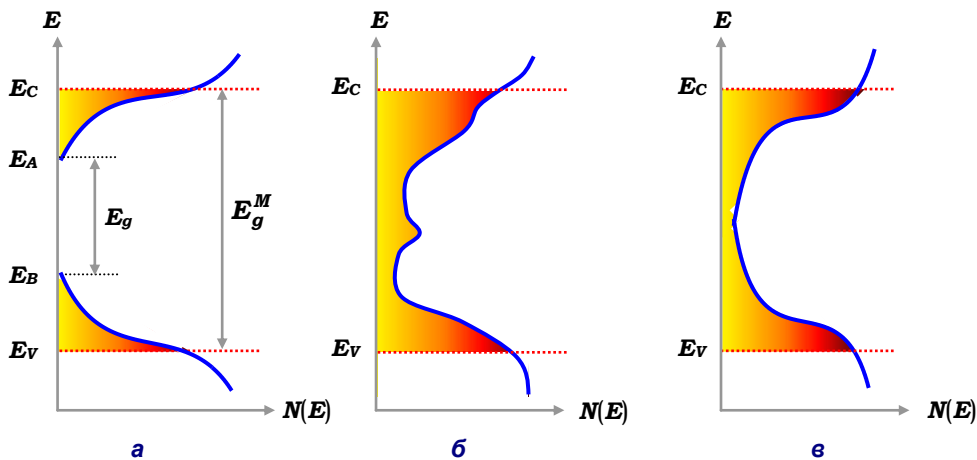


Рис. 17.11. Енергетична структура аморфного напівпровідника:

- а – "хвости" щільності станів у забороненій зоні;
- б – максимуми щільності станів всередині забороненої зони завдяки дефектам;
- в – перекривання "хвостів" щільності станів

Електронні стани у "хвостах" поділяються на локалізовані та делокалізовані. Останні дають внесок до процесів електроперенесення та часто називаються струмопровідними. Різкі границі між цими станами називаються краями рухливості (E_C та E_V). Відстань між ними називається забороненою зоною або щілиною рухливості (E_g).

Залежно від того, як формується густина станів аморфного матеріалу, виділяють три механізми провідності, які виявляються за різних температур. Перший – це перенесення по делокалізованих станах носіїв заряду, що збуджені за край рухливості. При цьому статична провідність σ у широкому інтервалі температур визначається формулою

$$\sigma = \sigma_0 \exp \left\{ -\frac{E_C - E_F}{k_B T} \right\}, \quad (17.57)$$

де E_F – енергія Фермі, а $\sigma_0 = 10^3 \div 10^4 \text{ Ом}^{-1}\text{см}^{-1}$. Іншим механізмом є стрибкове перенесення носіїв заряду, що збуджені у локалізовані стани поблизу країв рухливості (напр., у стани між E_A та E_C). У цьому випадку

$$\sigma = \sigma'_0 \exp \left\{ -\frac{E_A - E_F + W}{k_B T} \right\} \quad (17.58)$$

з енергією активації стрибка W і $\sigma_0 \leq 10 \text{ Ом}^{-1}\text{см}^{-1}$. Третім механізмом провідності є стрибкове перенесення по локалізованих станах поблизу E_F

$$\sigma = \sigma''_0 \exp \left\{ -k_B T^{-1/4} \right\}. \quad (17.59)$$

При цьому перші два механізми більш характерні для халькогенідних склоподібних напівпровідників, а третій – для елементарних тетраедричних напівпровідників. Рухливість носіїв заряду таких напівпровідників є незначною $\sim 10^{-5} - 10^{-8} \text{ см}^2\text{В}^{-1}\text{с}^{-1}$ і залежить від товщини зразка та прикладеного до нього електричного поля. Така залежність пов'язується або з багаторазовим захопленням носіїв на локалізовані стани, що розподілені за деякою функцією розподілу, або зі стрибковим перенесенням. Для більшості халькогенідних склоподібних напівпровідників провідність та енергія активації не залежать від природи та концентрації домішок, оскільки зазвичай домішкові атоми виявляють максимальну валентність, віддаючи всі свої валентні електрони на утворення ковалентних зв'язків з атомами основної речовини. Однак домішки перехідних металів, таких як Ni, Fe, W тощо спричиняють появу домішкової провідності, що викликає різке зростання провідності аморфного матеріалу, зокрема, аморфний кремній вда-

ється ефективно легувати фосфором і бором. У багатьох халькогенідних стеклах спостерігається ефект переключення, що полягає у швидкій ($\sim 10^{-10}$ с) зміні характеру провідності з високоомного до низькоомного стану під дією електричного поля із напруженістю $\geq 10^2$ В·см⁻¹. Цей ефект пов'язується з інжекцією носіїв заряду з контакту та делокалізацією захоплених носіїв. Поряд з цим збільшення провідності спричиняє зростання температури у шнурі струму. Зазвичай ефект переключення є відворотнім. Але в деяких халькогенідних стеклах низькоомний стан зразка зберігається досить довго, і для повернення у високоомний стан необхідно пропустити через зразок короткий імпульс струму. Такий ефект пам'яті може бути пояснений частковою кристалізацією речовини в області струмового шнура.

Що стосується оптичних властивостей аморфних і склоподібних напівпровідників, то варто зазначити, що край основного поглинання світла в них можна поділити на три ділянки. В області високих значень коефіцієнта поглинання ($\alpha > 10^4$ см⁻¹) його залежність від частоти можна описати формулою

$$\alpha = B \frac{\hbar\omega - E_g^{opt}}{\hbar\omega}, \quad (17.60)$$

де константа $B \sim 10^5 - 10^6$ см⁻¹еВ⁻¹, а E_g^{opt} – оптична ширина забороненої зони. В області середніх значень коефіцієнта поглинання

$$1 \text{ см}^{-1} < \alpha < 10^3 - 10^4 \text{ см}^{-1}$$

залежність α від частоти має вигляд

$$\alpha = \alpha_0 e^{A\hbar\omega}, \quad (17.61)$$

де $A = 15-20$ еВ⁻¹. Область слабкого оптичного поглинання ($\alpha < 1$ см⁻¹) визначається поглинанням світла на дефектах структури.

Фотопровідність в аморфних напівпровідниках залежить від інтенсивності світла за законом

$$\sigma_{ph} = A \cdot I^\gamma, \quad (17.62)$$

де A – константа, I – інтенсивність світла, що поглинається, а показник γ може змінюватись в інтервалі значень від 0,5 до 1,0. Спектральна залежність σ_{ph} характеризується максимумом і пологим довгохвильовим плечем. При цьому σ_{ph} досягає максимального значення у тій області температур, де $\sigma_{ph} \sim \sigma$ і за зниженням температури спадає експоненційно, а в області низьких температур спадання уповільнюється. Така поведінка σ_{ph} пояснюється прилипанням і рекомбінацією нерівноважних носіїв на локальних центрах, що неперервно розподілені у забороненій зоні.

Найкраще серед аморфних напівпровідників вивчено аморфний кремній, основні властивості якого визначаються характеристиками його структури: аморфна форма кремнію утворена групами, які складаються із чотирьох атомів та утворюють просторову структуру, що має форму тетраедра і, на відміну від кристалічного стану, тетраедри орієнтовані випадково (рис. 17.12).

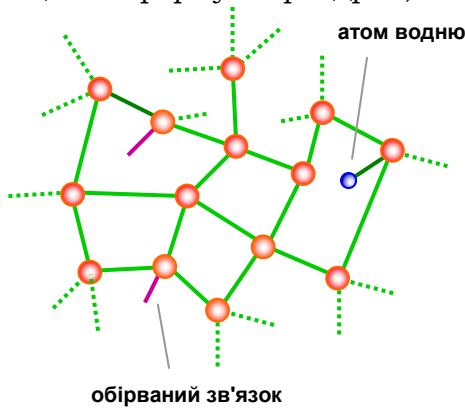


Рис. 17.12. Структура аморфного кремнію. Кожен із атомів кремнію пов'язаний з іншими чотирма зв'язками (подано зеленим). Обірвані зв'язки подано фіолетовим

ри орієнтовані випадково (рис. 17.12). Цей факт обумовлює наявність великої кількості обірваних зв'язків. Обірвані зв'язки з одного боку є "постачальниками" додаткових електронів, що можуть брати участь у формуванні електричного струму. З іншого боку, обірвані зв'язки є ефективними центрами рекомбінації, і їх наявність знижує провідність аморфного кремнію. Поверхня α -Si активно адсорбує з атмосфери водень, який "заліковує" обірвані зв'язки. Це знижує щільність локалізованих станів в забороненій зоні, що призводить до зміни структури, електричних та оптичних властивостей α -Si.

■ 17.3. Практичне застосування аморфних напівпровідників

Завдяки прозорості у довгохвильовій області спектра аморфні напівпровідники застосовуються в оптичному приладобудуванні, зокрема аморфний кремній – для виготовлення недорогих сонячних елементів. Поєднання високого електричного опору та значної фотопровідності використовуються й в електрофотографії, телевізійних передаючих трубках типу видикон і виготовлення фото-термопластичних перетворювачів зображення. Ефекти переключення та пам'яті дозволяють отримувати швидкодіючі перемикачі та матриці пам'яті. Фотолегування та оберненість фотостимульованої зміни оптичних властивостей аморфних напівпровідників використовується для виготовлення світлореєструючих середовищ у голографії. Стимульована зовнішніми впливами зміна розчинності аморфних напівпровідників дозволяє застосовувати їх для виготовлення фото-, електронно-, та

рентгенорезисторів, фотошаблонів, що використовуються у сучасній мікро- та нанотехнологіях.

Одним із напрямків практичного застосування аморфних халькогенідних напівпровідників, що широко обговорювався в 70-ті роки минулого століття, було створення перемикачів, що базувались на ефекті перемикання. Перехід аморфного напівпровідника із високоомного до низькоомного стану під дією зміщення V_t і подальше утримання зразка у стані високої провідності за прикладення невеликого зміщення V_H (рис. 17.13) служило базою для створення електронних перемикачів.

Іншим напрямком застосування аморфних напівпровідників є використання фотовольтаїчного ефекту в них для створення простих у виготовленні недорогих сонячних елементів. Оскільки аморфний кремій поглинає сонячне світло майже у сорок разів ефективніше, ніж монокристалічний кремій, то основним напрямком застосування аморфного кремію є побудова фотовольтаїчних приладів, зокрема сонячних елементів. Ще однією перевагою сонячних елементів на аморфному кремії є їхня радіаційна стійкість. Дійсно, оскільки аморфний кремій не є структурно однорідним, то навіть значні порушення структури, спричинені дією іонізуючого випромінювання, суттєво не змінять його властивості. Однак виникнення радіаційних дефектів спричинить більш активні рекомбінаційні процеси на цих дефектах, і провідність матеріалу зменшуватиметься. Зрозуміло, що застосування такого матеріалу в електронних приладах не є доцільним. Але якщо аморфний кремій отримується таким чином, що містить деяку кількість водню, то атоми водню насичуватимуть обірвані зв'язки і, таким чином, "заліковуватимуть" радіаційні пошкодження. Приклад можливої схеми сонячних елементів з аморфного кремію подано на рис. 17.14. Комірка складається із надтонкого (8 нм) шару аморфного кремію p -типу, середнього шару товщиною 0,5–1,0 мкм і тонкого (20 нм) нижнього шару n -типу. Верхній шар зроблено відносно прозорим, і майже все світло проходить крізь нього та генерує електрони у середньому шарі. p - та n -шари, виготовлені за допованого аморфного кремію, створюють електричне поле у чистому середньому шарі, таким чином створюючи рух електронів через нього, що спричиняє розділення зарядів і виникнення різниці потенціалів між верхнім і нижнім шарами.

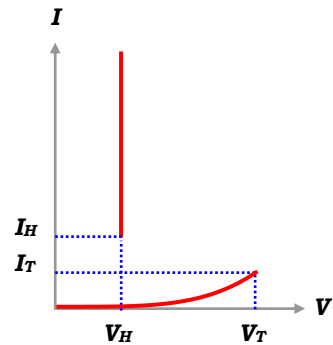


Рис. 17.13. Вольт-амперна характеристика аморфного халькогенідного напівпровідника у режимі перемикача

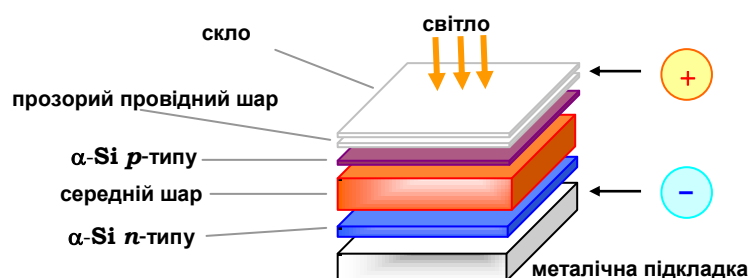


Рис. 17.14. Структура типової комірки сонячного елемента з аморфного кремнію

Останніми часами здійснюються інтенсивні роботи із розробки нових електронних приладів на базі аморфних матеріалів, наприклад, нещодавно (2005) на основі нового аморфного матеріалу IGZO (аморфна сполука In-Ga-Zn-O) було розроблено та виготовлено прозорий тонкоплівковий транзистор. Новий матеріал, на відміну від аморфного гідрогенізованого кремнію, де рухливість носіїв не перевищує $1 \text{ см}^2\text{В}^{-1}\text{с}^{-1}$, характеризується досить високою рухливістю, що сягає величини $10 \text{ см}^2\text{В}^{-1}\text{с}^{-1}$. Такі високі значення рухливості визначаються іншим, ніж у аморфного кремнію, механізмом провідності. Як відомо, електропровідність аморфного кремнію забезпечується стрибковим механізмом і тим самим значення рухливості сильно обмежені. Електропровідність аморфного IGZO забезпечується перенесенням носіїв по делокалізованим станам і в цьому сенсі нагадує перенесення носіїв заряду у звичайних кристалічних напівпровідниках.

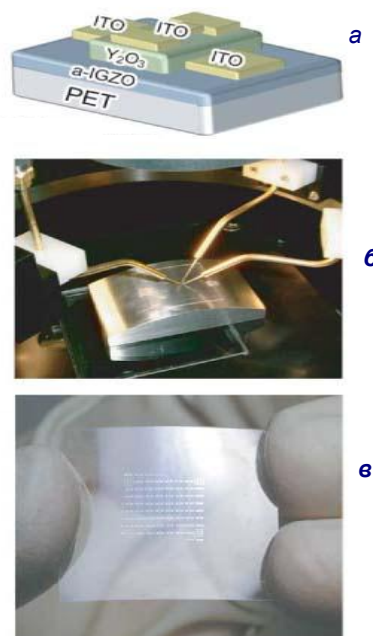


Рис. 17.15: а – схема тонкоплівкового транзистора на основі α -IGZO на полімерній плівці поліетилен-терефталату – PET; б–в – фотографії гнучкої плівки із прозорими транзисторами

Іншою суттєвою відмінністю аморфного IGZO від аморфного кремнію є велика заборонена зона, завдяки чому він прозорий у видимому діапазоні. Це дозволяє використовувати матеріал для виготовлення прозорих електронних пристроїв, наприклад прозорих дисплеїв, які можна наносити безпосередньо на лобове скло автомобілів або літаків. На рис. 17.15 подано структурну схему такого транзистора

та фотографії готових пристроїв. Масив гнучких прозорих тонкоплівкових польових транзисторів було створено на поверхні тонкої (200 нм) полімерної плівки PET. Нанесення шару α -IGZO на поверхню плівки відбувалось шляхом лазерного випаровування кристалічної мішені InGaZnO_4 за кімнатної температури в атмосфері кисню. Отримана плівка була напівпровідником n -типу. При цьому концентрацію носіїв можна було варіювати у широких межах від 10^{14} см^{-3} до 10^{20} см^{-3} , змінюючи тиск кисню.

У наступному етапі за допомогою стандартної літографії було сформовано масив польових транзисторів із верхнім затвором. Тонкий шар α -IGZO відіграв роль каналу. Електроди та діелектричні шари також виготовлялись із прозорих матеріалів ITO та Y_2O_3 . Вивчення характеристик транзисторів показали, що нова технологія дає перспективний матеріал для створення гнучких прозорих тонкоплівкових електронних пристроїв, зокрема було виявлено, що вигинання плівки практично не впливає на робочі характеристики транзисторів (рис. 17.16), які здатні функціонувати за температур до 120° C .

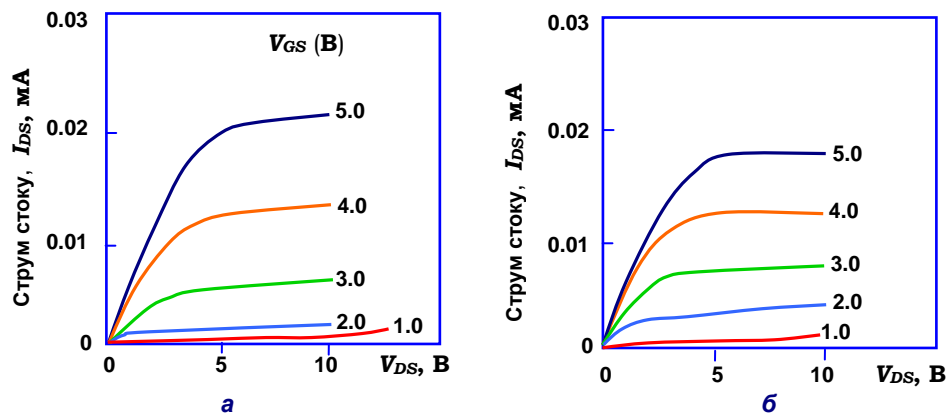


Рис. 17.16. Залежність струму витік–стік I_{DS} від напруги між виток і стоком V_{DS} за різних напруг на затворі V_{GS} у плівці: а – ненапружений, б –напружений (вигнутій)

Аморфні напівпровідники також використовуються при побудові потужних напівпровідникових лазерів інфрачервоного діапазону, зокрема відомо, що надійність потужних напівпровідникових лазерів визначається, крім всього, багатошаровим інтерференційним дзеркалом, що формується на торці структури (див. розд. 15). Матеріали, що використовуються у таких оптичних покриттях, мають бути стійкими до дії інтенсивних випромінювань (аж до випромінювання з інтенсивністю $\sim 10^6 \text{ Вт}\cdot\text{см}^{-2}$). Використання аморфного кремнію як речовини із високим значенням коефіцієнту заломлення у шарі, що

формує інтерференційне дзеркало лазерної структури на основі твердих розчинів $A^{III}B^V$, має багато переваг порівняно із суто діелектричними дзеркалами. По-перше, таке покриття може бути достатньо малої товщини, що зумовлена високими значеннями відбиваючої здатності на межах розподілу **аморфний кремній-діелектрик**, яка визначається великою різницею коефіцієнтів заломлення. По-друге, оскільки зазвичай напівпровідникові лазери на основі матеріалів $A^{III}B^V$ випромінюють в області ~ 1 мкм, тобто поблизу області прозорості аморфного кремнію, втрати через поглинання, отже й нагрівання вихідних дзеркал лазерів будуть незначними. Крім того, оптичні властивості аморфного кремнію значно залежать від умов його отримання. Це означає, що ретельний підбір умов отримання α -Si дозволяє створювати покриття із мінімальними втратами й тим самим оптимізувати лазерну структуру у цілому.

■ 17.7. Завдання

1. Пояснити, чому у некристалічних речовинах формуються енергетичні зони.
2. Визначте переваги та недоліки використання аморфного кремнію у сонячній енергетиці

Список літератури

1. *Смит Р.* Полупроводники. – М.: Мир, 1982.
2. *Мотт Н.*, *Электронны в неупорядоченных структурах*. – М.: Мир, 1969.
3. *Мотт Н., Дэвис Э.* *Электронные процессы в некристаллических веществах: В 2 т.* – М.: Мир, 1982
4. *Pollak M.*, *Noncrystalline Semiconductors: v.1–3.* – CRS Press, 1987.
5. *Kasap S.O.* *Principles of Electronic Materials and Devices*, McGraw-Hill, 2005.

Розділ 18

МОЛЕКУЛЯРНІ НАПІВПРОВІДНИКИ

Сучасна мікро- та наноелектроніка базується на твердотільних технологіях і за багатьма параметрами класична¹ твердотільна електроніка наблизилась до межі подальшого вдосконалення. Як реальна альтернатива їй розглядається молекулярна електроніка.

Молекулярна електроніка – це наука про електронні властивості молекулярних матеріалів та можливості їх застосування в електроніці, як елементів та приладів для запису, зберігання, обробки та передачі інформації.

Тому є кілька причин. За мільйони років еволюції природа створила велике розмаїття молекул, що виконують всі необхідні для живих організмів функції – сенсорні, логічно-аналітичні, пам'яті, моторні тощо. Ці молекули характеризуються оптимальною конфігурацією та нанометровими розмірами. Крім того, електроніка, що базується на молекулярних сполуках, за необхідності є одноелектронною, тобто характерними електричними струмами в молекулярних структурах є одноелектронні збудження. Це знімає, наприклад проблему тепловиділення, що стоїть перед сучасною твердотільною мікроелектронікою. Тому молекулярну електроніку можна вважати найефективнішою з погляду енергетичних витрат на один біт обробленої інформації. Основні переваги молекулярної електроніки перед традиційною твердотільною:

- ♦ лінійний розмір активного елемента принаймні на три порядки менший;
- ♦ розташування молекулярних фрагментів точно детерміновано (напр., як у молекули ДНК), на відміну від випадкового розподілу домішок у неорганічних матеріалах;
- ♦ можливість тривимірної організації і самоорганізації, на відміну від переважно планарних сучасних напівпровідникових технологій;
- ♦ відсутність обірваних хімічних зв'язків;
- ♦ зменшення розмірів елементів молекулярної електроніки до 10^{-7} см, що означає зростання швидкодії молекулярних пристроїв. Оцінки показують, що максимальна швидкодія (за швидкості передачі заряду 10^8 см/с) становитиме 10^{-15} с. Зазначимо, що у низці випадків необхідно застосовувати жорсткіші обмеження швидкодії, пов'язані з обмеженістю частоти молекулярних коливань, які зменшують час перетворення молекули до 10^{-13} – 10^{-14} с;

¹ Під терміном "класична твердотільна електроніка" ми маємо на увазі електроніку, що базується на неорганічних кристалічних (або неупорядкованих) матеріалах – Si, Ge, GaAs тощо.

♦ можливість застосування багатозначної логіки (кубіти) та, відповідно, ефективніших способів обробки інформації (як це відбувається, наприклад у мозку людини).

Визначимо молекулярні напівпровідники як такі, основні властивості яких пов'язані з хімічною будовою молекул. Міжмолекулярна взаємодія, що в основному визначається ван-дер-ваальсовими силами, у цьому випадку виявляється набагато слабкішою, ніж внутрішньо-молекулярна взаємодія. Таким чином, електрони виявляються в основному локалізованими на молекулах, а не "розмазаними" по кристалічній ґратці. Як приклад можна навести бензол та антрацен, що кристалізуються в однаковій кристалічній групі зі схожими параметрами ґратки. Однак електронні параметри, наприклад такі, як ширина забороненої зони у цих молекулярних кристалах, виявляються різними. Справа в тому, що в таких речовинах енергія забороненої зони E_g суттєво залежить від числа π -електронів (див. далі), яке є пропорційним кількості n бензольних кілець у молекулі. У бензолі $n = 1$, в антрацені $n = 3$, а ширина забороненої зони виражається емпіричною формулою

$$E_g = a + \frac{b}{k^n}, \quad (18.1)$$

де a , b , k – константи.

Можна виділити дві основні властивості молекулярних напівпровідників. По-перше, це локалізація електрона на молекулі; по-друге, визначене розташування атомів у молекулі, на відміну від розташування домішок в кристалічному напівпровіднику. До молекулярних напівпровідників можуть бути зараховані як органічні, так і неорганічні сполуки. Однак, оскільки органічні матеріали існують у набагато більшій кількості, ніж неорганічні, та, крім того, характеризуються багатьма цінними властивостями, ми, коли йтиметься про молекулярні напівпровідники, в основному матимемо на увазі органічні сполуки.

Треба зауважити, що молекулярні напівпровідники являють собою системи різної розмірності. Молекули в розчині є нічим іншим, як системою квантових точок; довгі полімерні молекули можна розглядати як квантові нитки; молекулярні плівки та шари типу графіту є прикладом двовимірних електронних систем; молекулярні кристали являють собою тривимірні системи.

■ 18.1. Фізичні обмеження твердотільної мікроелектроніки та переваги молекулярних систем

Розвиток твердотільної електроніки, що відбувався під знаком несподіваного зменшення характерних розмірів елементів приблизно вдвічі кожні півтора роки, у 90-ті рр. XX ст. почав сповільнюватись і на початок XXI ст., діставши показник ~ 0.2 мкм, практично наблизився до фізичної межі, що обмежує подальшу мініатюризацію напівпровідникових приладів. Розглянемо детальніше ці фізичні обмеження. Перша полягає в наявності *флуктуацій концентрацій домішок*. У розд. 7 ми детально обговорювали проблему легування напівпровідників і з'ясували, що сучасні напівпровідникові технології значною мірою визначаються процесами легування напівпровідникових матеріалів. При цьому розподіл домішок усередині зразку значною мірою залишається випадковим. Середньоквадратична флуктуація ΔN числа домішок в об'ємі V

$$\Delta N \cong N^{1/2} = (nV)^{1/2}. \quad (18.2)$$

Для стабільної роботи електронних приладів необхідно, щоб відносна флуктуація концентрації домішки не перевищувала 1 %, тобто

$$\frac{\Delta N}{N} \leq 10^{-2}, \quad (18.3)$$

звідки маємо $nV \geq 10^4$. Це означає, що за концентрації легуючої домішки $nV \approx 10^{18} \cdot \text{см}^{-3}$ об'єм системи не має бути меншим за $V \approx 10^{-14} \cdot \text{см}^3$, тобто граничне значення лінійних розмірів системи становить $l \approx 0,2$ мкм. Розвиток сучасних технологій модуляційного легування дає можливість значно зсунути цей розмір у бік менших значень, але це зменшення не може перевищувати порядок l , тобто, навіть зважаючи на майбутнє значне вдосконалення технологій легування, навряд чи можливе обмеження, пов'язане з наявністю флуктуацій концентрації домішки до показника, меншого за 0,01 мкм. Наступною перешкодою до подальшого зменшення елементів електронних приладів є *дифузія домішок на поверхні*. Поверхневий коефіцієнт дифузії зазвичай виявляється істотно більшим за його значення в об'ємі. Ця обставина обмежує мінімальні значення латеральних розмірів елементів величиною ~ 0.1 мкм.

При зменшенні лінійних розмірів елементів електронних пристроїв виникає проблема зміни *характеру електронного транспорту*. Дійсно, у розділі 16 цієї книги ми бачили, що якщо розмір пристрою

менший за довжину вільного пробігу електрона l , то характер транспорту визначається специфічними квантовими ефектами, наприклад такими, як квантування кондактансу. Крім того, рух електронів у цьому випадку є балістичним, що підвищує швидкодію пристроїв, оскільки час перебування електрона в пристрої набагато менший за час процесів, що визначаються дифузійним переносом. Але при цьому різко збільшуються вимоги до реалізації пристрою, адже процеси розсіювання на границях пристрою набувають визначального характеру.

Іншим важливим фактором є ефекти, пов'язані з квантовою когерентністю, що виявляються в ситуації, коли транспорт у межах пристрою є фазово-когерентним. Це відбувається, якщо розмір пристрою є меншим за довжину дифузії збою фази $\sqrt{lv\tau_\phi}$ (v – швидкість електрона, τ_ϕ – час збою фази, що в найпростішому випадку дорівнює часу релаксації електронів на фононах). У випадку високих температур час релаксації фази є дуже малим, і квантово-інтерференційні ефекти не виявляються. Але для пристроїв, що працюють в умовах низьких температур, довжина фазової когерентності може досягати 1 мкм, та електронний транспорт можуть визначати квантово-інтерференційні ефекти. Ці обставини, що по суті пов'язані з розмірними ефектами, також призводять до принципових обмежень створення пристроїв малих розмірів (але не варто виключати можливість побудови приладів на інших засадах).

Ще одним фактором, що принципово обмежує можливість зменшення розмірів напівпровідникових пристроїв, є нагрівання елементів електронних пристроїв при проходженні через них електричного струму. Вважається, що для нормальної роботи електронних пристроїв їхній нагрів не має перевищувати величини 10^2 Вт/см². Ця вимога приводить до визначення величини мінімальних лінійних розмірів електронних пристроїв значеннями ~ 0.25 мкм. Але перехід до мідних технологій і надслабких струмів уже в наш час зсунув цю прогнозовану величину до значень ~ 0.1 мкм.

Але головна проблема, що постає перед твердотільною електронікою, полягає у філософії побудови електронних пристроїв як набору великої кількості елементарних об'єктів, які поєднані величезною кількістю зв'язків. Це призводить до надзвичайної складності сучасних приладів (мікросхем) і пов'язаної з цим проблеми низького виходу приладів у процесі їх виробництва та великих труднощів при їх тестуванні. З іншого боку, вивчення біологічних систем привело до розгляду окремих біомолекул, які за своєї складності та, головне, за функціональними можливостями порівняні із сучасними інтегральними мікросхемами. Крім того, окрема біомолекула є одиничним елементом молекулярної

електроніки. Саме ця перевага й змушує все більшу кількість учених і наукових груп переходити до вивчення молекулярних структур як елементів електроніки майбутнього.

■ 18.2. Елементи молекулярної електроніки

Наскільки ж взагалі є реальною молекулярна електроніка? При аналізі цього питання стриманого оптимізму надає факт існування людського мозку як неперевершеного приладу обробки інформації та органів відчуття. Найкращими сенсорами є очі та ніс. Людське око може реєструвати світлові потоки на рівні одного фотона, а ніс детектує хімічні речовини концентрацією у кілька молекул у кубічному метрі.

Дослідження з молекулярної електроніки почались із синтезу молекул, що здатні виконувати логічні функції, зокрема бістабільних молекул, які мають два стабільних стани та змінюють свій стан під дією зовнішніх факторів (електричного поля, тиску, світла тощо), тобто працюють у режимі тригерів. Проблема передача інформації безпосередньо до самих молекул після відкриття електропровідних полімерів – молекулярних сполук зі спряженими зв'язками – фактично переведена в практичну площину пошуку відповідних полімерів і контрольованої побудови молекулярних структур. Виготовлення та маніпулювання пристроями на молекулярному рівні може базуватись на добре розвинених методах нанолітографії, коли структури з органічного матеріалу виготовляються шляхом молекулярної адсорбції на відповідно виготоввану поверхню. Існують також безпосередні методи молекулярного дизайну, коли потрібні молекули висаджуються на поверхню за допомогою зонду скануючого мікроскопа атомних сил. Останнім часом велику увагу приділяють вивченню процесів самоорганізації органічних молекул при влаштуванні відповідних умов. Наприклад, отримання мономолекулярних шарів за допомогою метода Ленгмюра-Блоджетт (див. додаток К) може поєднуватись з організацією молекул у досить добре контрольовані молекулярні комплекси. Про такі комплекси, а саме молекулярні пакети фталоціанінів металів, ми вже згадували, коли обговорювали можливість утворення одновимірних кристалічних структур.

Таким чином, молекулярні органічні сполуки можуть уже тепер складати конкуренцію твердотільній напівпровідниковій електроніці. Які ж характерні риси електронних станів є причиною незвичайних властивостей органічних сполук?

■ 18.3. sp^2 -гібридизація у вуглеці

Незвичайні властивості органічних напівпровідників значною мірою зумовлені структурою хімічних зв'язків, що утворюють атоми вуглецю, з'єднуючись у молекули. Серед різноманітних конфігурацій атом вуглецю може утворювати електронні стани, що формуються з одного s -рівня та двох рівнів p -симетрії. У результаті виникають три стани, які ще називають sp^2 -гібридизованими орбіталями. Оскільки ці орбітали мають ковалентну природу, вони мають визначений напрямок у просторі: sp^2 -орбітали направлені вздовж променів правильної трикутної зірки та розташовуються в одній площині. Перпендикулярно до цієї площини розташовується p^z -орбіталь. Коли два атоми вуглецю взаємодіють, дві sp^2 -орбітали перекриваються та утворюють σ -зв'язок (рис.18.1). Відстань між основним і першим збудженим станами σ -зв'язку є досить великою та відповідає частотному діапазону ультрафіолету (рис.18.2). Однак при з'єднанні атомів вуглецю можуть виникати додаткові зв'язки між p^z -орбіталями, які називаються π -зв'язками. Ці зв'язки мають набагато менший енергетичний зазор між основним і збудженим станами, який відповідає оптичному діапазону частот і забезпечує напівпровідникові властивості органічних сполук. Атоми вуглецю можуть з'єднуватися у довгі ланцюжки. Відповідно, довгі ланцюжки зв'язків вуглецю характеризуватимуться деякою енергетичною щільною між найвищим заповненим рівнем (НОМО – *highest occupied molecular orbital*) і найнижчим, незайнятим молекулярним рівнем (LUMO – *lowest unoccupied molecular orbital*).

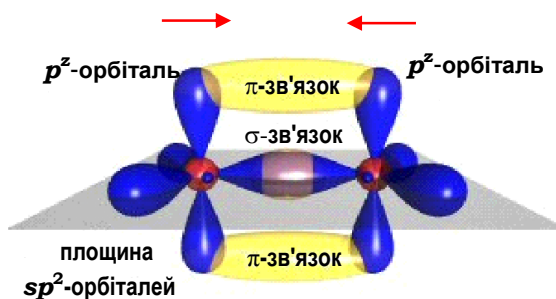


Рис.18.1. Утворення зв'язків між атомами вуглецю

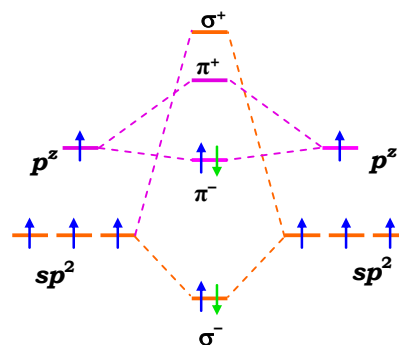


Рис.18.2. Енергетична діаграма хімічних зв'язків вуглецю із врахуванням sp^2 -гібридизації

На прикладі формування бензольного кільця розглянемо, як такі стани виникають при формуванні органічних молекул з невеликою

кількістю атомів вуглецю. Коли атоми вуглецю формують бензольне кільце (що часто є елементом побудови складніших органічних молекул), електрони на π -зв'язках колективізуються та стають спільними для всієї молекули (рис. 18.3). У результаті колективізації π -електронів щілина між НОМО та LUMO зменшується, виникають енергетичні стани, що дуже нагадують енергетичні зони у напівпровіднику (рис. 18.4).

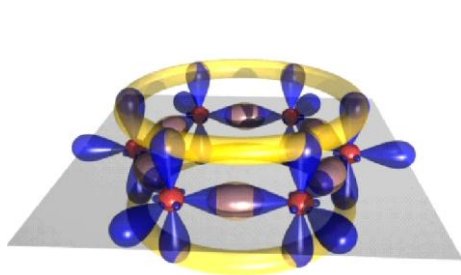


Рис.18.3. Формування бензольного кільця атомами вуглецю.
Жовтими торами показано просторовий розподіл колективізованих π -електронів

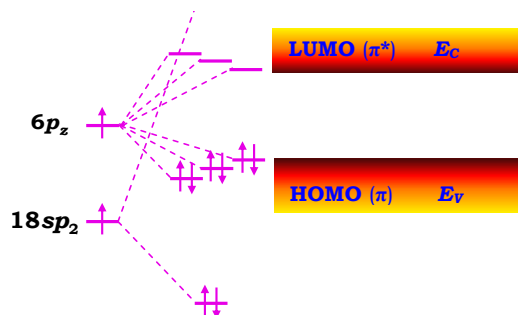



Рис.18.4. Енергетична діаграма електронних станів бензольного кільця

Варто зважати на те, що широко використовуване в молекулярній електроніці поняття "дірка" має дещо інше значення, ніж у кристалічному напівпровіднику. Дійсно, згадаємо, що поняття дірки ми ввели, активно використовуючи трансляційну інваріантність кристалу. У молекулярних системах найчастіше так робити не можна, що не можна сказати про формування електронних станів у бензольному кільці. Тут є трансляційна інваріантність, можна сформулювати граничні умови, аналогічні умовам Борна-Кармана. Але все одно, через значну просторову локалізацію електронних станів у молекулі та їхню відносно невелику кількість у кільці (шість) "класичне" визначення дірки зіштовхується з великими труднощами. Тому, використовуючи поняття "дірка", найчастіше мають на увазі незайнятий стан у смугі зайнятих станів. Таким чином, за молекулярної формули C_6H_6 структурна формула бензолу матиме вигляд, показаний на рис. 18.5. Ця структурна формула відбиває той факт, що в молекулі бензолу всі атоми вуглецю розташовані в одній площині, і вуглецевий скелет молекули має форму шестикутника, у вершинах якого розташовані атоми вуглецю, формально зв'язані між собою впереміжку трьома одинарними та трьома подвійними зв'язками. Четвертий зв'язок кожного з атомів вуглецю насичений воднем. Оскільки довжини одинарних і подвійних зв'язків не однакові, то, якщо вуглецевий каркас має такий вигляд, як на рис. 18.5 а, відстані між атомами вуглецю в такому каркасі

мають бути попарно різними. Таку структурну модель бензолу запропонував німецький хімік Ф.А. Кекуле в 1865 р. Проте рентгеноструктурні дослідження показали, що всі атоми вуглецю у бензолі розташовані один від одного на однаковій віддалі (0,14 нм), що більша за характерну довжину подвійного зв'язку та менша за характерну довжину одинарного зв'язку. Таким чином, можна стверджувати, що порівняно зі структурою Кекуле, молекулярна структура, що відтворює еквівалентність усіх шести вуглець-вуглецевих зв'язків, являє собою резонуючу структуру між двома структурами Кекуле (рис. 18.5 б). Дійсно, суперпозиція двох структур Кекуле, де довжина одинарного зв'язку дорівнює 0,154 нм, а довжина подвійного зв'язку – 0,133 нм, дає резонуючий зв'язок із довжиною $\sim 0,14$ нм. У зв'язку з цим у хімії використовують позначення для структурних формул бензольних кілець (рис.18.5). У подальшому для спрощення рисунків використовуватимемо для позначення структури бензольного кільця форму, аналогічну до поданої на рис. 18.5 б, в, зважаючи на резонуючий характер зв'язків між атомами вуглецю в бензольному кільці.

Аналогічно спрощено подаватимемо структури інших органічних речовин. Наприклад, структурна формула елемента ланцюжка ацетилену $[-CH=CH]$ подаватиметься як . Отже, замість структурної формули на рис. 18.6 а зображуватимемо спрощену формулу як на рис. 18.6 б.

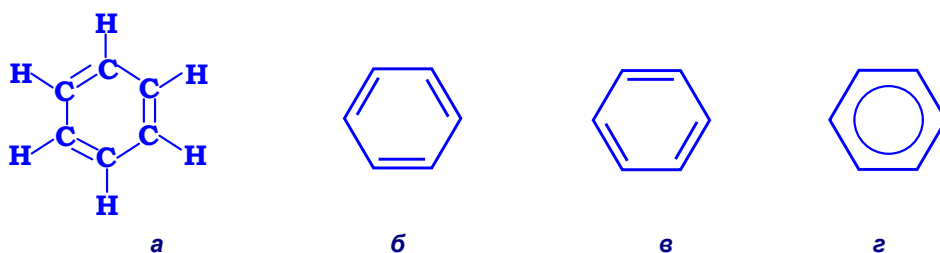


Рис. 18.5. Структурна формула бензолу – а, спрощенні позначення бензольного кільця – б і в у формі Кекуле; г – форма, що показує еквівалентність вуглець-вуглецевих зв'язків за рахунок резонансу між структурами б і в

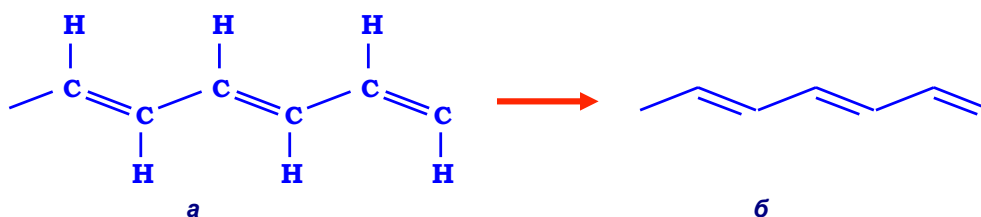


Рис. 18.6. Фрагмент полімерного ланцюжка поліацетилену – а та його спрощене зображення – б

Бензолні кільця можуть бути структурними елементами, із яких формується молекулярний кристал. Завдяки щільному зв'язку π -систем молекул у цьому кристалі, вони у чистій формі демонструють вражаючі транспортні властивості. Наприклад, за кімнатної температури такі молекулярні кристали можуть характеризуватися рухливістю $\sim 1\text{--}10\text{ см}^2/\text{В}\cdot\text{с}$. Більшість молекул, що мають в основі бензолне кільце, можуть також легко випаровуватись у полікристалічну форму, де реалізується стрибковий транспорт із типовими значеннями рухливості $\sim 10^{-3}\text{ см}^2/\text{В}\cdot\text{с}$, або ж в аморфну форму зі стрибковою провідністю та рухливістю $\sim 10^{-5}\text{ см}^2/\text{В}\cdot\text{с}$ при 300 К.

■ 18.4. Напівпровідникові плівки композицій органічних матеріалів

Одним із напрямків у створення нових функціональних матеріалів для використання в молекулярній електроніці є синтез компонент та дослідження напівпровідникових властивостей плівок композицій органічних матеріалів (ПКОМ). Ці композитні матеріали повинні відповідати таким вимогам:

- ♦ можливість інжекції носіїв заряду з низькоомних електричних контактів до ПКОМ;
- ♦ можливість виходу нерівноважних носіїв заряду з ПКОМ у низькоомний контакт;
- ♦ достатньо висока рухливість носіїв заряду;
- ♦ невелика ймовірність захоплення рухливих носіїв заряду на енергетичні пастки.

У загальному випадку електронні властивості ПКОМ можна описувати в моделі твердого розчину молекул двох типів у нейтральному плівкоутворюючому матеріалі: з акцепторними властивостями (A) та з донорними властивостями (D). При цьому прийнято вважати, що носіями з негативним електричним зарядом є неспарені електрони на LUMO акцепторних молекул, а носіями з позитивним електричним зарядом є дірка в сенсі відсутності одного із двох електронів на HOMO донорних молекул. Наслідком останнього визначення є уточнення терміну "дірка" для ПКОМ: транспорт позитивного заряду як рух дірки є не чим іншим, як переходом електрона з повністю заповненої HOMO молекули на молекулу, де HOMO частково (наполовину) заповнена. Простіше кажучи, при наближенні молекул A та D одна до одної електрон з донорної молекули переноситься на акцепторну, лишаючи позитивно заряджений катіон. Акцепторна молекула набуває електрон і стає аніоном, тобто ві-

дбується реакція з перенесенням заряду $A + D \leftrightarrow A^- + D^+$. Ці заряди вже не беруть участь в молекулярних зв'язках і можуть пересуватись у ПКОМ, утворюючи транспортний струм.

Як півкоутворюючу основу (матрицю) часто-густо використовують полімери або олігомери.

Полімери (від грец. *polymeres* – що складається з багатьох частин) – це хімічні сполуки з високою молекулярною масою (від кількох тисяч до кількох мільйонів), молекули яких складаються з великої кількості однакових груп (мономірних ланцюжків).

Олігомери (від грец. *oligos* – малий, і *meros* – частина), сполуки, що займають проміжний стан за розмірами молекул між мономерами та полімерами. Верхня межа молекулярної маси олігомерів залежить від їх хімічного складу та за порядком величини збігається з молекулярною масою мономірного ланцюжка полімеру. На відміну від полімерів, властивості олігомерів значно залежать від кількості повторюваних ланцюжків і природи кінцевих груп.

Таким чином, молекули речовин зі складу ПКОМ бувають або розчинені в півкоутворюючій основі, або з'єднані хімічним зв'язком зі структурними одиницями полімерів чи олігомерів основи. Як π -донори найчастіше виступають ароматичні сполуки з електродонорними замісниками (алкіламіно-, алкоксигрупа тощо) та гетероциклічні сполуки.

У цих сполуках, завдяки участі неподієних електронних пар замісників (гетероатомів) у спряженні, виникає частковий негативний заряд. До числа таких сполук належать, наприклад карбазол, епоксіпропілкарбазол (ЕПК) (рис. 18.7) і похідні алкосізаміщених антрацена. Ці сполуки є головними структурними фрагментами фотопровідних полімерів та олігомерів на основі полі-N-вінілкарбазолу (ПВК), полі-N-епоксіпропілкарбазолу (ПЕПК) та антраценіагліцинового ефіру (ОГАЕ), що використовуються в електрографії, голографії, фотоелектричних перетворювачах сонячної енергії, електролюмінісцентних і фоторефрактивних середовищах. На рис. 18.8 подано структурні формули цих речовин. Величина n у цих формулах може набувати значення від одиниць (як для олігомерів) до десятків і сотень тисяч (що характерно для полімерів).

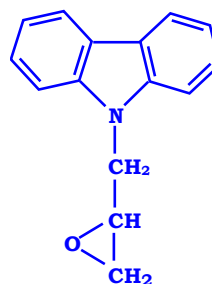


Рис. 18.7. Структурна формула епоксіпропілкарбазолу (ЕПК)

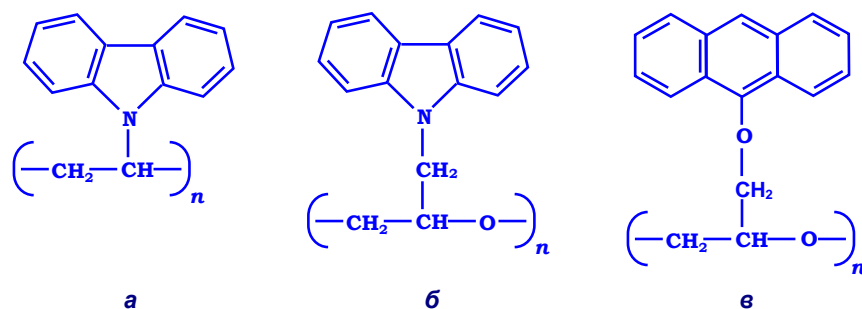


Рис.18.8. Структурні формули головних фрагментів фотопровідних полімерів та олігомерів, що використовуються як плівкоутворюючі основи:
 а – полі-N-вінілкарбазол (ПВК); б – полі-N-епоксіпропілкарбазолу (ПЕПК);
 в – антраценілгліцинового ефіру (ОГАЕ)

До найбільш сильних молекул-донорів належать тетрагіофульвалени, характерним представником яких власне є тетрагіофульвален (ТТФ), структурну формулу якого подано на рис. 18.9. Класичними π -акцепторами є органічні сполуки, π -зв'язки яких спряжені із сильними електроакцепторними групами, наприклад тетраціанетилен і поліароматичні сполуки. Введення таких груп утворює в цих молекулах частковий позитивний заряд, величина якого зростає за мірою збільшення їх кількості. Із цієї причини ефективними акцепторами виступають, наприклад такі сполуки, як 2,4,7-тринітро-9-флуоренон (ТНФ) і 2,4,5,7-тетранітро-9-флуоренон (ТеНФ). До сильних акцепторів належать також фулерени та їхні хімічні димери та тримери. Карбазоліл- та антраценілмісткі полімери та олігомери поглинають світло в ультрафіолетовій частині спектру, їхні спектри поглинання майже повторюють спектр поглинання мономерних структурних одиниць. З іншого боку, ПКМ без спеціально введених домішок не поглинають в оптичному та ультрафіолетовому діапазонах.

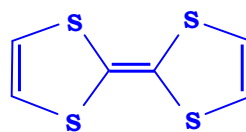


Рис. 18.9. Тетрагіофульвален (ТТФ)

Електронні рівні донорних та акцепторних молекул у композиті зазвичай не розщеплені та залишаються локальними. Це означає: щільність ймовірності знаходження неспареного електрона на відповідній орбіталі визначається квадратом модуля хвильової функції, що експоненційно спадає зі збільшенням відстані від молекули. Переходи неспарених електронів з молекули на молекули відбуваються за тунельним механізмом, ймовірність цих переходів залежить від відстані між молекулами

$$W_{n,p} \propto R_{n,p}^2 \exp(-2R_{n,p}/\alpha_{n,p}). \quad (18.4)$$

де $R_{n,p} = N_{n,p}^{-1/3}$, $N_{n,p}$ – концентрація акцепторних (n) і донорних молекул (p) у композиті. При цьому радіуси локалізації електронів на акцепторних молекулах α_n і дірок на донорах α_p можуть відрізнятись, але не перевищують 1 нм. Таким чином, завдяки слабкому перекриттю хвильових функцій неспарених електронів на сусідніх молекулах транспорт нерівноважних носіїв заряду відбувається у результаті тунельних переходів електронів між локальними електронними рівнями молекул. Для більшості ПКОМ характерною є емпірично встановлена залежність рухливості електронів (μ_n) і дірок (μ_p) від напруженості зовнішнього електричного поля E та температури T

$$\mu_{n,p} \sim R_{n,p}^2 \exp(-2R_{n,p}/\alpha_{n,p}) \exp\left\{-\frac{U_{n,p} - \beta E^{1/2}}{kT} \cdot \frac{T_0 - T}{T_0}\right\}, \quad (18.5)$$

де $U_{n,p}$ – енергія активації рухливості електронів і дірок, відповідно; β – коефіцієнт, що кількісно збігається зі сталою Пула-Френкеля, T_0 – характеристична температура, яка для більшості ПКОМ перебуває в діапазоні 400–600 К. У таких ПКОМ рухливість носіїв заряду змінюється в широкому інтервалі від 10^{-7} до 10^{-3} см²/(В·с). Ці вирази дуже схожі на залежності, що використовуються в моделі розділення двох заряджених часток (моделі Пула-Френкеля та Онзагера). У цих моделях зменшення енергетичного бар'єру для розділення двох зарядів відбувається завдяки можливості збільшення енергії рухливих носіїв заряду у зовнішньому електричному полі. Оскільки перший множник у (18.5) подібний до виразу ймовірності тунельного переходу між локалізованими станами, а другий множник – ймовірності подолання електроном потенціального бар'єру, що створений протилежно зарядженим центром, то транспорт нерівноважних носіїв заряду можна представити таким, що складається з дифузії електронів усередині молекул та електронного тунелювання – між сусідніми молекулами. Звідси можна зробити висновок, що зменшення енергії активації рухливості нерівноважних носіїв заряду треба очікувати за збільшення лінійних розмірів π -спряженої системи молекул акцепторів і донорів. Саме такий ефект досягається у випадку спряжених полімерів, про що йтиметься далі. Але моделі Пула-Френкеля та Онзагера не можуть повністю бути адекватними для пояснення залежностей (18.5), оскільки для цих моделей передбачується наявність у ПКОМ неконтрольованих кулонівських центрів, а експерименти показали, що залежності (18.5) малочутливі до електрично заряджених домішок у складі ПКОМ. Також для всіх ві-

домих ПКОМ величина енергії активації $U_{n,p}$ лежить у межах 0,3–0,8 еВ та є не дуже чутливою до введення домішок. Тому було запропоновано розглядати транспорт нерівноважних носіїв заряду в ПКОМ таким, що відбувається не за ізоенергетичними рівнями, а за енергетичними рівнями з дисперсією по енергії і геометричним відстаням між молекулами. Розрахунки ймовірності електронних переходів із врахуванням цих обставин дали такі самі аналітичні залежності, як (18.5).

Як же з'являються нерівноважні носії заряду в ПКОМ? Це відбувається так само, як і в класичних напівпровідниках: у результаті термопольової емісії з електричних контактів, зовнішнього та внутрішнього фотоефекту. Донорні молекули характеризуються потенціалом іонізації I_{gd} , якому пропорційна енергія НОМО, а акцепторні молекули характеризуються енергією електронної спорідненості A_{en} (якій пропорційна енергія ЛУМО). Для того, щоб відбувався процес термопольової інжекції заряду з контактів у ПКОМ і відбору електронів на інший електричний контакт, матеріали для контактів вибирають таким, щоб були виконані умови $|I_{gy}| \ll |F|$ та $|A_{en}| \ll |F|$, де F – потенціал іонізації (рівень Фермі) електричного контакту. Але використовуючи метал як контакт з ПКОМ, ці умови дуже важко виконати, оскільки величина A_{en} для органічних речовин не часто перевищує 3 еВ і величина енергетичного бар'єру контакту на межі метал–ПКОМ залишається досить великою для електронних переходів через бар'єр за кімнатної температури ($kT = 0,025$ еВ). Тому перспективним вважається використання як електричних контактів у пристроях молекулярної електроніки полімерних матеріалів з металічним типом провідності, в яких енергія електронної спорідненості може змінюватись у широкому енергетичному інтервалі.

Оскільки більшість полімерів та олігомерів, що є плівкоутворюючою основою ПКОМ, не мають забарвлення у видимій частині спектру, то для забезпечення внутрішнього фотоефекту до складу ПКОМ вводять спеціальні домішки. Молекули цих домішок поглинають світло та сприяють утворенню нерівноважних носіїв заряду при освітленні ПКОМ, тобто вони є центрами фотогенерації. Таким чином, механізм фотогенерації носіїв заряду в ПКОМ складається із двох стадій. На першій стадії у результаті поглинання кванта світла центр фотогенерації переходить з основного у збуджений стан, із якого відбуваються міжмолекулярні переходи електронів. При цьому утворюється кулонівськи зв'язана електронно-діркова пара. На другій стадії фотогенерації під впливом зовнішнього електричного поля за ненульової температури носії заряду з електронно-діркової пари розходяться по молекулам А та D через стрибковий механізм провідності, створюючи тим самим

струм фотопровідності. Молекули центрів фото генерації підбирають таким чином, щоб виконувались умови $|I_{gc}| > |I_{gd}|$ та $|A_{ec}| < |A_{en}|$, де I_{gc} та A_{ec} є відповідно потенціалом іонізації та енергії електронної спорідненості молекули центру фотогенерації. При створенні фоточутливих ПКМ особливу увагу привертають до вибору центрів фотогенерації носіїв заряду. Найбільш ефективними є органічні барвники, яким притаманні інтенсивні смуги поглинання та широкий діапазон енергій електронних станів. Але введення органічних барвників до складу ПКМ хоч і зміщує спектральну область фоточутливості зазначених матеріалів, але не забезпечує підвищення рухливості нерівноважних носіїв заряду в них. Останню вимогу можна виконати при переході від полімерів з ароматичними замісниками до π -спряжених полімерів.

■ 18.5. Енергетична структура та елементарні збудження у полімерах

Якщо вуглецеві молекули вибудовуються в довгі полімерні ланцюжки, то π -зв'язки стають делокалізованими вздовж всього ланцюжка та формують одновимірну електронну систему. У результаті одновимірні електронні стани, що мають походження від НОМО, утворюють деяку зону визначеної ширини. Виявляється, що ця зона відділена від зони незайнятих станів (які походять від LUMO) енергетичним зазором так, що електронна структура молекулярного ланцюжка є енергетичною структурою одновимірного напівпровідника (рис. 18.10–18.11). Транспортні властивості таких полімерів зазвичай визначаються дефектами в одновимірному ланцюжку або процесами стрибкової провідності між полімерними ланцюжками. Таким чином, полімерні системи здебільшого демонструють не провідність зонних електронів, а термічно активовану стрибкову провідність.

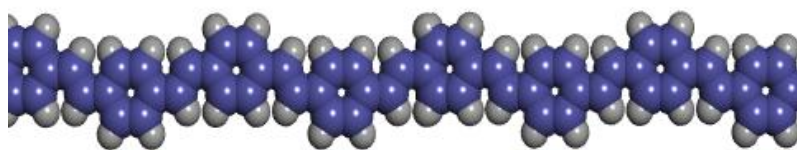


Рис.18.10 Полімерний ланцюжок, що утворюється з'єднанням бензольних кілець і виникненням системою спряжених зв'язків

Розглянемо детальніше причини виникнення провідності у спряжених полімерах. Спряженими полімерами називають молекули, в яких чергуються прості, подвійні та (або) потрійні зв'язки. Узагалі

кажучи, полімери зазвичай є діелектриками, але за деяких умов спряжені полімери можуть характеризуватись провідністю, що наближається до металічної. Типовим представником спряжених полімерів є поліацетилен $(\text{CH})_n$, частину ланцюжка якого подано на рис. 18.6. Виявляється, якщо поліацетилен допувати, тобто окислювати у парах йоду, бромю або хлору, його провідність зростає на 7–9 порядків. Таке суттєве підвищення провідності за допущення матеріалу дуже подібне до зростання провідності кристалічних напівпровідників за їх легування. Високі значення провідності легованих полімерів зі спряженими зв'язками пояснюються не рухом убудованих між полімерними ланцюжками іонів, а рухом у полімерному ланцюжку деяких заряджених частинок, що утворені зарядами, наведеними легуючою домішкою – солітонів і поляронів, які вільно переміщуються вздовж ланцюжка.

Делокалізований подвійний зв'язок у полімері означає, що електронна густина не локалізована між двома сусідніми p -орбіталями, а розподілена по всій одновимірній π -орбітальній системі. Якщо в спряженій системі існує дефект (напр., типу повороту зв'язку), то він приведе до повної локалізації електрона. Але й в ідеальному ланцюжку спряжених зв'язків провідність не є великою. Емпірично доведено, що ширина енергетичної щілини в спряжених полімерах обернено пропорційно залежить від числа N елементів, що повторюються. Наразі можна записати формулу, аналогічну (18.1)

$$E_g = E_0 + \frac{B}{N}, \quad (18.6)$$

де E_0 та B – константи. Залежність ширини енергетичної щілини від N типу (18.6) спостерігається тільки для полімерів з обмеженою кількістю ланок. Такі полімери називаються, як вже зазначалось, олігомерами. Виявляється, що існує така кількість ланок N_c , що при перевищенні цього числа енергетична щілина змінюється дуже незначно. Таке число називається числом спряження. Помноживши N_c , на розмір елементарної комірки a , отримаємо довжину спряження. Для застосування полімерів в оптичних пристроях достатньо мати довжину молекули, що дорівнює або дещо перевищує довжину спряження. Такі полімери матимуть спектр поглинання у видимому діапазоні. Однак для транспортних властивостей цього виявляється не достатньо, оскільки провідність в основному визначатиметься не внутрішньо-молекулярним перенесенням, а стрибковою провідністю між полімерними ланцюжками. Отже, можна стверджувати, що полімери з довгими ланцюгами характеризуватимуться більшою провідністю, оскільки в цьому разі транспорт реалізується меншою кількістю міжмолекулярних стрибків.

Крім того, у полімерах із довгими ланцюгами зменшується концентрація кінцевих груп, які можуть виступати як пастки для нерівноважних носіїв заряду, що також спричиняє підвищення провідності.

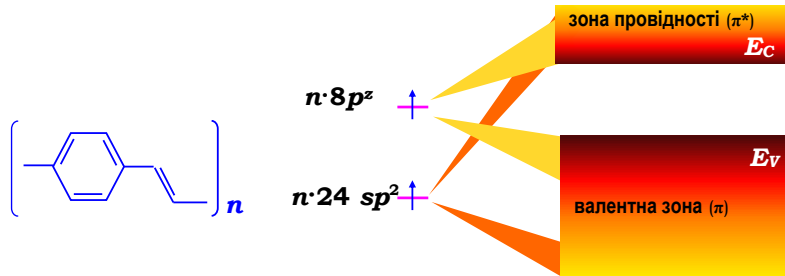


Рис.18.11. Формування енергетичної структури при утворенні полімерного ланцюжка зі спряженими зв'язками на прикладі поліфенілену-винілену

Вираз (18.6) для енергетичної щільності був отриманий емпіричним шляхом, виходячи з аналізу існуючих експериментальних даних. Для обчислення закону дисперсії електронів в одновимірному ланцюжку зі спряженими зв'язками можна, використовуючи метод молекулярних орбіталей, знайти розв'язки рівняння Шредингера, у припущенні, що в одновимірному ланцюжку відстані між атомами вуглецю є сталими та дорівнюють a .

Покажемо на прикладі поліацетилену, яким чином можна обчислити енергетичні стани електронів у спряжених полімерах. У методі молекулярних орбіталей хвильова функція електрона у полімерному ланцюгу має вигляд лінійної комбінації атомних орбіталей (АКАО)

$$\Phi_g = \frac{1}{\sqrt{N}} \sum_n e^{ikan} \phi_n, \quad (18.7)$$

де a – відстань між атомами вуглецю в полімері, k – хвильовий вектор, ϕ_n – атомна хвильова функція, N – повне число атомних хвильових функцій. Підсумовування ведеться за позиціями атомів вуглецю n . Зрозуміло, що повне число доданків у сумі дорівнює N . На рис. 18.12 подано графічно методику проведення таке підсумовування. Ситуація, що показана на рисунку, аналогічна випадку, коли в методі сильного зв'язку обчислюються енергетична структура атомного ланцюжка, де за рахунок перекриття атомних хвильових функцій виникають енергетичні зони, що розширюються зі збільшенням кількості атомів у ланцюжку. Граничний випадок нескінченного ланцюжка спричиняє суцільне перекривання хвильових функцій і появу зонної структури металічного типу. Іншими словами, відбувається повна делокалізація π -електрону та "розмазка" його хвильової функції по всьому ланцюжку.

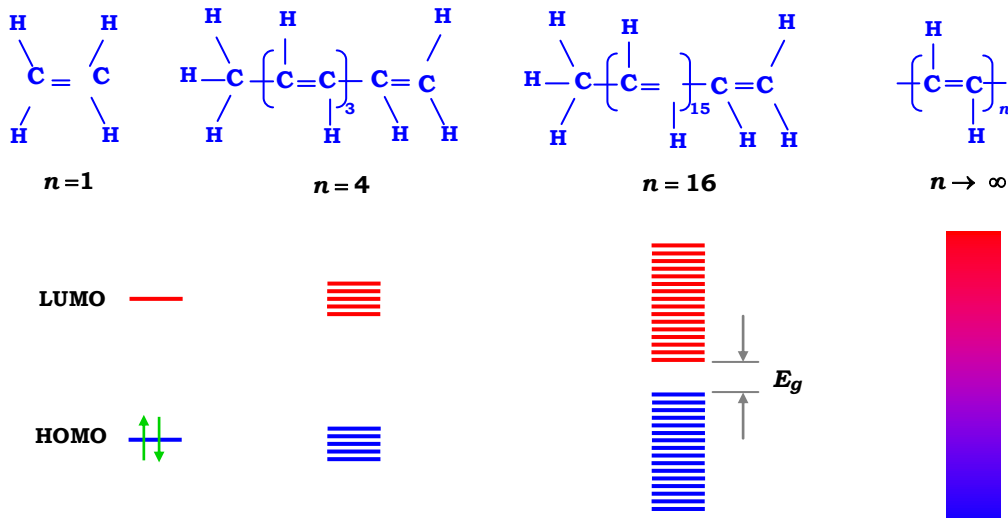


Рис.18.12. Залежність енергетичної структури полімеру від кількості молекулярних одиниць, що повторюються

Розв'язок рівняння Шредингера із хвильовою функцією типу (18.7) показує, що залежність енергії від хвильового вектора матиме дві гілки

$$E^V(k) = -2t \cos ka, \quad (18.8)$$

$$E^C(k) = 2t \cos ka, \quad (18.9)$$

де t – інтеграл перекривання p хвильових функцій між сусідніми атомами вуглецю, що характеризує ступінь делокалізації електронів. Оскільки кожен атом вуглецю віддає один π -електрон, то згідно із принципом Паулі на кожному енергетичному рівні знаходиться два (із протилежно направленими спінами) електрона. Відповідну схему зони провідності та щільність станів подано на рис. 18.13.

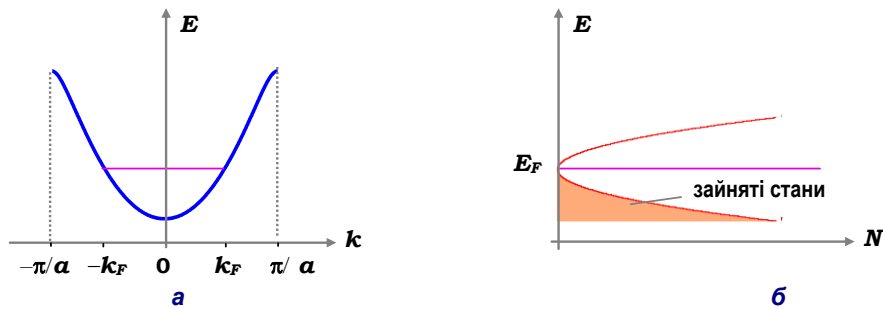


Рис. 18.13. Закон дисперсії (а) та щільність станів (б) полімерного ланцюжка поліацетилену за припущення, що довжина одинарних і подвійних зв'язків однакова

Як видно, щільність станів має вигляд, типовий для одновимірного руху, коли її максимальні значення досягаються на краю зон

$$N(E) = \frac{1}{2\pi t a} \sqrt{\frac{t}{E - E_0}}. \quad (18.10)$$

Повне число станів у зоні дорівнює числу атомів у ланцюжку, що утворюють спряжені зв'язки. З огляду на те, що кожний стан характеризується визначеним спіном \uparrow або \downarrow , число станів, що відповідають заданій енергії, збільшується вдвічі. А оскільки кожен атом постачає один електрон, то виявиться, що зона заповнена наполовину. Таким чином, розглянута система демонструватиме металічні властивості.

Але ж експеримент свідчить про те, що поліацетилен є напівпровідником. Це означає, що запропонована модель зонної структури є непридатною. Як бути? Прискіпливіший розгляд системи дає відповідь на поставлене питання. Виявляється, що довжини одинарного та подвійного зв'язків хоч і не набагато, але розрізняються. Подвійний зв'язок виявляється коротшим за одинарний на $0,03\text{--}0,04 \text{ \AA}$. Однак така невелика різниця в довжинах зв'язків спричиняє помітну зміну електронних властивостей поліацетилену. Тепер, у димеризованому ланцюжку поліацетилену, довжина елементарної комірки збільшується у два рази та становить $2a$ (рис. 18.14). Для визначення електронних властивостей таких ланцюжків використовують гамільтоніан Си-Шриффера-Хеєгера (SSH), що містить два доданки

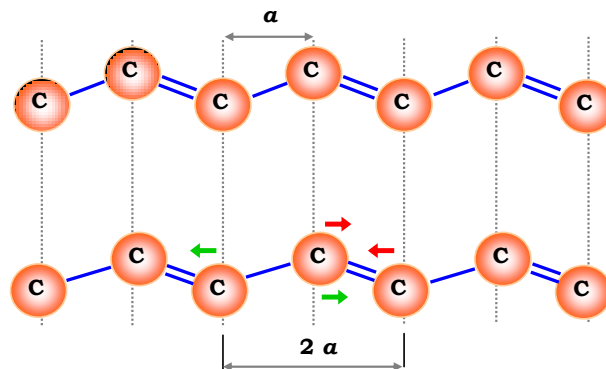


Рис. 18.14. Перехід Паєйрлса в поліацетилені

Угорі – ланцюжок з одинарними та подвійними зв'язками однакової довжини (нестійка структура).

Унизу – утворення ланцюжка з димеризованих блоків

(червоним подано дві можливості зсуву атомів вуглецю в межах однієї комірки

– атоми зсуваються назустріч один одному;

зеленими – атоми зсуваються у протилежні боки – стрілочками)

$$H = H_{electr} + H_{elast}, \quad (18.11)$$

де електронна частина гамільтоніану має вигляд

$$H_{electr} = -\sum_{n,s} t_{n+1,n} (C_{n+1,s}^+ C_{n,s}^+ C_{n+1,s}). \quad (18.12)$$

Цей доданок описує π -електрон у поліацетилені з інтегралом передачі $t_{n+1,n}$, що визначається перекриванням хвильових функцій π -електронів двох сусідніх атомів – $n + 1$ та n . Оператори народження $C_{n,s}^+$ і знищення $C_{n,s}$ описують появу або знищення π -електрона, що характеризується спіном $s = \pm 1/2$ в околі n -го атому вуглецю. Такий вигляд електронної частини гамільтоніану дозволяє описувати електронні стрибки між сусідніми вузлами (n та $n + 1$) ланцюжка. Пружна частина гамільтоніана, що має вигляд

$$H_{elast} = \frac{1}{2} \sum_n K(u_{n+1} - u_n)^2, \quad (18.13)$$

описує σ -електрони, локалізовані на суміжних атомах вуглецю. Тут u_n є зміщенням атому вуглецю, що перебуває у позиції n полімерного ланцюжка, K – пружна константа. Пружний доданок не дає безпосереднього внеску до формування електронних властивостей полімеру, але непрямо він значно впливає на них. Наприклад, якщо два сусідніх атоми вуглецю зміщуються назустріч один одному та розташовуються ближче один до одного, то перекриття хвильових функцій π -електронів на сусідніх вузлах збільшується, що спричинить збільшення константи $t_{n+1,n}$, і навпаки, якщо відстань між атомами вуглецю збільшується,

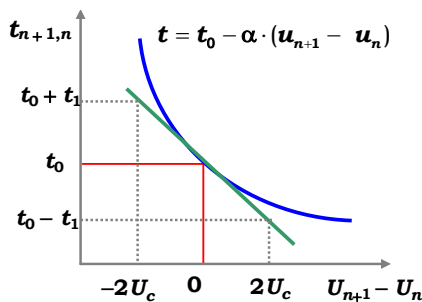


Рис. 18.15. До визначення залежності інтегралу передачі від зміщення атомів вуглецю

Походження цієї формули можна проілюструвати графічно (рис. 18.15). Тут t_0 – інтеграл передачі між двома однорідно зміщеними атомами вуглецю, α – константа електрон-фононної² взаємодії. Іншими словами, (18.14) можна розглядати як розкладення в ряд Тейлора інтеграла передачі за малим відносним зміщенням атомів u . Тепер в SSH-гамільтоніані можна покласти, що всі зміщення атомів вуглецю однакові та припустити, що власні функції гамільтоніану можна розкласти за атомними орбіталями ϕ , тобто

! Зауважимо, що термін "електрон-фононна взаємодія" в полімерах має дещо інший смисл, ніж у кристалах. А саме, у першу чергу мається на увазі вплив зміни величини зміщення між атомами вуглецю на константи, що визначають енергетичну структуру електронних станів полімеру.

перекриття хвильових функцій становиться меншим, тим самим зменшується константа $t_{n+1,n}$. У випадку невеликих зміщень залежність інтегралу передачі від зміщень атомів вуглецю можна подати у вигляді

$$t_{n+1,n} = t_0 - \alpha(u_{n+1,n} - u_n). \quad (18.14)$$

$$\Phi_n = \sum_n a_K(n) \phi_n. \quad (18.13)$$

Розв'язок рівняння Шредингера при цьому, має вигляд

$$E(k) = E_0 - 2t \cos ka. \quad (18.16)$$

Таким чином, ми прийшли до попереднього результату (формули (18.8–18.9)). Для того, щоб отримати правильну енергетичну структуру поліацетилену, треба взяти до уваги, що зміщення сусідніх атомів вуглецю у бік подвійного зв'язку є більшим, ніж у бік одинарного зв'язку. Тоді можна записати:

$$u_n = (-1)^n u = u e^{\pm 2ink_F a}, \quad (18.17)$$

де $k_F = \pi/2a$ має сенс хвильового вектора електронів на рівні Фермі, тобто найбільшої енергії, що мають електрони в основному стані. Розв'язок рівняння Шредингера тепер матиме вигляд

$$E(k) = \sqrt{4t_0^2 \cos^2 ka + \Delta^2 \sin^2 ka}, \quad (18.18)$$

де параметр енергетичної щілини $\Delta = 4\alpha u$. Звідси видно, що на грані зони Брилюена (тобто при $k = \pi/2a$) відкривається зонна щілина. При цьому виявляється, що валентна зона повністю заповнена, а зона провідності – порожня (рис. 18.16). Іншими словами, ми бачимо, що за невеликої зміни кристалічної структури відбувається докорінна перебудова енергетичного спектру системи. Отже, у системі відбувається фазовий перехід. На такий структурний фазовий перехід уперше вказав Пайерлс. Він показав, що одновимірний метал є нестабільним щодо дисторсії ґратки, яка є причиною появи зонної щілини на рівні Фермі. Така дисторсія в поліацетилені пов'язана зі зміщенням атомів вуглецю, що належать подвійному зв'язку назустріч одне одному (див. рис. 18.14). Розрахунки показують, що ширина забороненої зони поліацетилену становить $\sim 1,5$ еВ, що збігається з експериментально отриманими даними з поглинання світла за механізмом зона–зона.

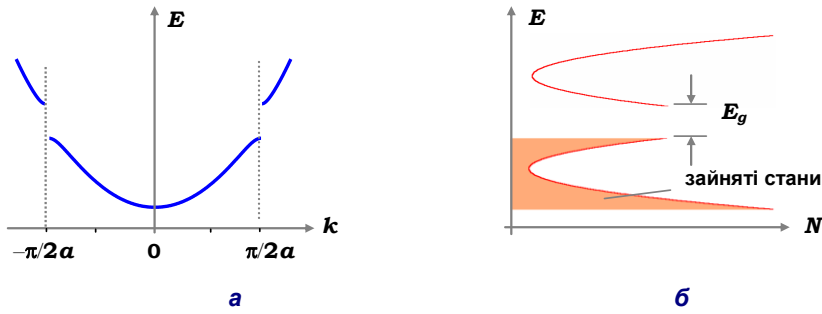


Рис.18.16. Енергія (а) та щільність (б) електронних станів у димеризованому поліацетилені

Якщо побудувати графік залежності абсолютної енергії, що припадає на один атом вуглецю в полімерному ланцюжку від зміщення, то отримаємо криву (рис. 18.17). Видно, що при $u = 0$ енергія системи має локальний максимум, що підтверджує нестабільність системи відносно переходу Пайерлса. Два локальні мінімуми відповідають двом можливостям утворення нової комірки полімеру – A або B . При цьому отримуються ті самі димеризовані полімери. Це означає, що основний стан поліацетилену є виродженим. Дефект, що виникає на межі розподілу двох фаз A та B , називається солітоном.

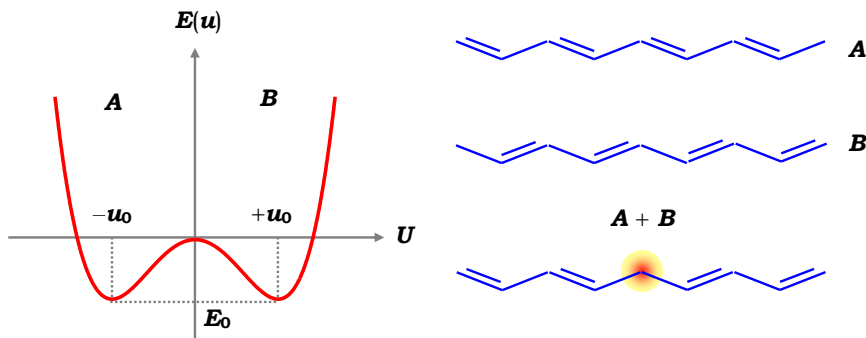


Рис.18.17. Енергія, що припадає на один атом вуглецю в поліацетилені (а); дві можливості (A та B) утворення димеризованого поліацетилену; виникнення солітона ($A + B$)

Припускаючи, що полімерний ланцюжок поліацетилену є нескінченним (тобто нехтуємо ефектами, що пов'язані з наявністю кінців полімерного ланцюга), за аналогією з магнетизмом солітон можна розглядати як доменну стінку між фазами A та B полімеру (рис. 18.18). Тоді хвильову функцію такого утворення можна записати у вигляді:

$$\phi_n = U_0 \operatorname{th} \left(\frac{na}{\xi} \right), \quad (18.19)$$

де ξ – підгоночний параметр, що визначає довжину "доменної стінки", тобто розмір солітону. Розрахунки показали, що у статичному наближенні енергія солітону виражається формулою

$$U_n = U_0 \operatorname{th} \left(\frac{(n - n_0) \cdot a}{\xi} \right) \cos(n\pi/2) \quad (18.20)$$

з n_0 позицією, що займає центр солітону, і нульовою енергією зв'язаного стану для електрону, що описується хвильовою функцією

$$\phi_n(n) = \left(\frac{a}{\xi} \right) \operatorname{sech} \left[(n - n_0) \cdot a/\xi \right] \cos(n\pi/2). \quad (18.21)$$

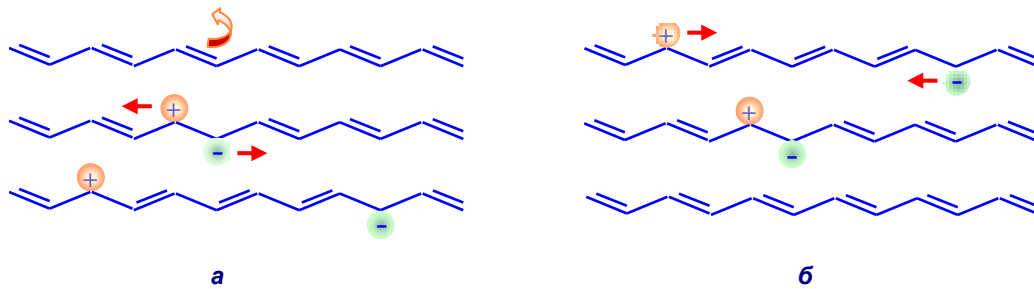


Рис. 18.18. Схематичне представлення нейтрального солітона (а) та антисолітона (б)

Мінімізація цих формул дає розмір солітону та найнижчий енергетичний рівень. Зокрема, можна знайти, що

$$\xi = \frac{aW}{\Delta} \approx 7a. \quad (18.22)$$

У цій формулі W – ширина зони π -електронів, а Δ є параметром щілини, за якої енергія солітонного стану набуває значення $2\Delta/\pi$. Таким чином отримуємо, що солітон є стабільнішим утворенням, ніж електрон-діркова пара в області ланцюжка, що містить 14 атомів вуглецю. Гамільтоніан системи характеризується симетрією зарядового спряження, тобто знак заряду трансформується зі зміною позиції атома вуглецю, на якому локалізований солітон, згідно із формулою

$$C_n \rightarrow (-1)^n C_n, \quad (18.22)$$

тобто енергетичний спектр системи є симетричним відносно нуля. Для солітону повний заряд і спин визначаються таким чином:

$$Q = e(N_+ + N_- - 1), \quad s = (1/2)(N_+ - N_-). \quad (18.23)$$

При цьому, якщо на межі розподілу локалізований один неспарений електрон, то загальний заряд такого утворення буде нульовим і такий

солітон називатиметься нейтральним. У цьому випадку $Q = 0$, $s = 1/2$. При видаленні (або додаванні) електрона в полімерний ланцюжок за допомогою легування утворюються заряджені солітони та нейтральні антисолітони, які за руху вздовж полімерного ланцюжка можуть анігілювати (рис. 18.19).

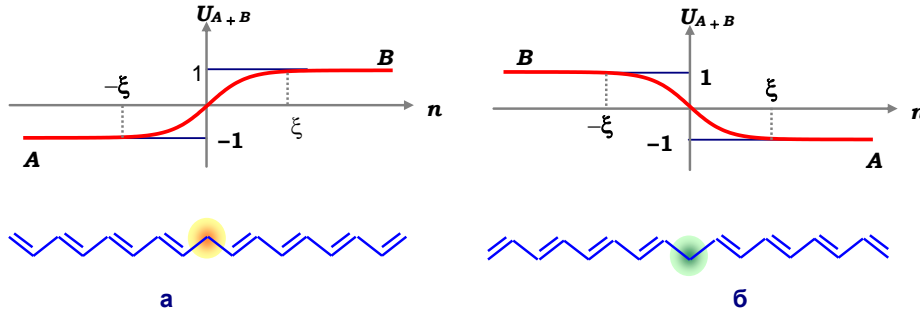


Рис. 18.19. Пара солітон–нейтральний антисолітон, що утворилась видаленням одного електрона із подвійного зв'язку (а). Анігіляція зарядженого солітону та антисолітону, що рухаються назустріч один одному (б)

Заряджені солітони несуть на собі один позитивний або негативний заряд і, оскільки зв'язують або два електрона з протилежними спінами, або область, де електрон відсутній, – нульовий спін. У забороненій зоні полімера солітон створює один енергетичний стан точно посередині зони (рис. 18.20). Солітони характеризуються всіма параметрами квазічастинки, тобто їх можна вважати вільними частинками, які характеризуються ефективною масою $\sim 6m_e$, радіусом локалізації (наближено становить 14 сталей ґратки) і дуже високою середньою швидкістю $\sim 10^6$ см/с.

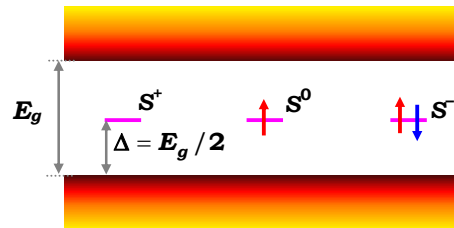


Рис. 18.20. Енергетична діаграма і спіновий стан позитивного (S^+), нейтрального (S^0) та негативного (S^-) солітону

Якщо в полімері наявні два стани, що відповідають структурам A та B , мають різну енергію (рис. 18.21), то полімери є невиродженими. У невироджених полімерах також може існувати дефект у порядку альтернування зв'язків, тобто виникне границя між структурами типу A та B . Але такий дефект має іншу фізичну природу. Різниця в енергіях $\Delta E = |E_B - E_A|$ основного стану структур типів A та B , що припадає на кожен повторюваний

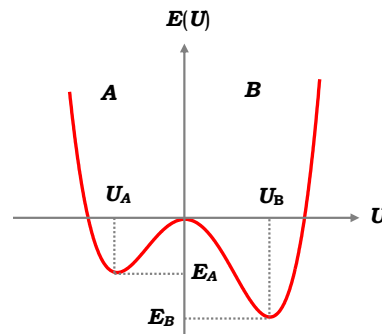


Рис. 18.21. Залежність енергії від зміщення атома у невиродженому полімері

елемент полімерного ланцюжка, призводить до інших ефектів, ніж ті, що спостерігаються у вироджених полімерах. Уявимо, що два напівнескінченних ланцюжка (ліворуч – типу A , праворуч – типу B) з'єднуються з утворенням дефекту типу солітона. За рахунок того, що енергія, яка розділяє стани двох ланцюжків, буде досить великою, стан солітона не буде стійким. Такий дефект не існуватиме. Якщо ланцюжки будуть короткі, то дефект рухатиметься до того боку ланцюжка, де розташовується високоенергетичний стан, тому в кінці високоенергетична структура зникне, і система перейде в стан із нижчою енергією. Таким чином, ми доходимо висновку, що існування дефекту типу солітону в невироджених полімерних структурах не є ефективним.

Для релаксації збуджень у невироджених полімерах реалізується інший, поляронний, механізм. Завдяки великій константі електрон-фононного зв'язку заряд, інжектований у полімерний ланцюжок, завжди самозахоплюється на вузлі та утворює полярон. У цьому випадку полярон є утворенням, аналогічним до полярону в твердому тілі, тобто являє собою електрон, зв'язаний із фононом завдяки тому, що електрон викликає локальну поляризацію ґратки й тим самим утворює потенціальну яму, до якої і потрапляє. При цьому утворення потенціальної ями є швидкоплинним процесом із часом $\sim 10^{-13}$ с. Полярон локалізований у досить великій області полімерної ґратки (6–10 елементарних комірок), із розмірами $\sim 2,5 - 4$ нм. Енергія зв'язку полярону відносно краю зони становить величину 0,012–0,03 еВ. Ефективна маса полярону дорівнює кільком масам електрона. Спін полярону дорівнює $1/2$. Полярон є досить рухливою частинкою. Він може переміщуватись полімерним ланцюжком, доки не буде захоплений на дефекті.

Однак полярони не є основними агентами електронного транспорту в полімерах. Електричні та магнітні вимірювання на легованих невідроджених полімерах показали, що основними носіями заряду в таких структурах є безспінові збудження, що несуть заряд $2e$. Найвірогіднішим кандидатом на такі утворення є або негативно заряджений солітон, або біполярон³ – спарений полярон, енергія якого є суттєво меншою за енергію двох незв'язаних поляронів. Таким чином, енергетичний вигравш від формування біполярна має бути достатнім, щоб подолати кулонівське відштовхування двох однакових зарядів. Розрахунки підтверджують цей висновок: енергія біполярона є нижчою на 0,4 еВ, ніж енергія двох поляронів. Рис. 18.22 ілюструє біполярони в поліфенілені та політіофені.

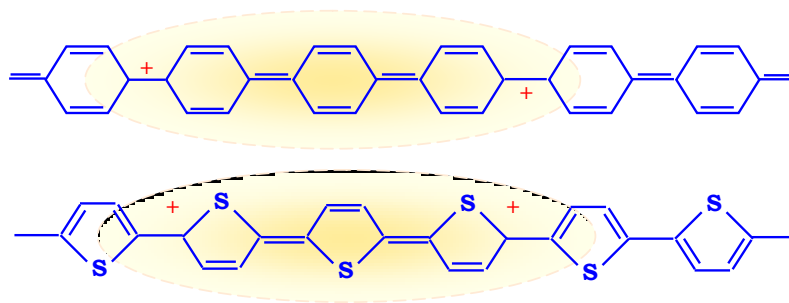


Рис. 18.22. Біполярон у поліфенілені (зверху) та політіофені (знизу)

На відміну від поляронів, які формуються майже миттєво за інжекції у полімер заряду, біполярони формуються із уже утворених поляронів. Розрахунки енергетичних рівнів поляронів і біполяронів, що виникають у забороненій зоні, показали: структура цих рівнів є такою, що полярон створює три можливих смуги поглинання, а біполярон – дві. Інтенсивність відповідних смуг залежить від перерізу захоплення фотона або від сили осцилятора. Цей параметр є малим для високоенергетичних переходів. Таким чином, експериментально можна спостерігати одну смугу поглинання з енергією, дещо меншою за половину забороненої зони (переходи за участю біполярна), та іншу смугу в інфрачервоній області – переходи за участю полярона. На рис. 18.23 подано залежність коефіцієнта поглинання нелегованого та легованого в парах йоду поліацетилену від енергії світла, що поглинається. Зауважимо, що рівень легування для полімерів визначається величиною провідності, оскільки нелегований полімер є непровідним. Як видно, легування спричиняє різкі зміни у спектрах поглинання: додатково до основного піку із краєм 1,4 еВ

³ Питання створення електричного струму в полімерах за механізмом біполярна не є загальновизнаним на цей момент.

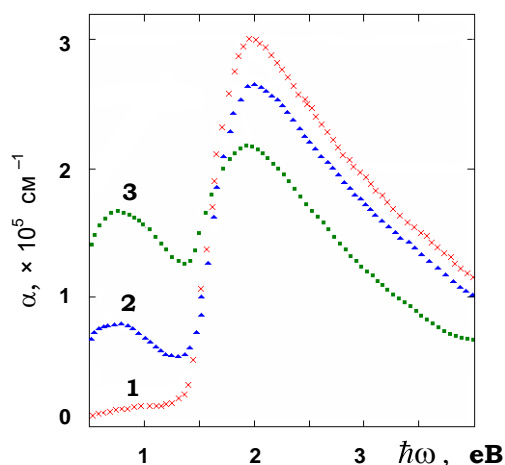


Рис. 18.23. Спектральна залежність коефіцієнта поглинання полі ацетилену за різними рівнями легування:
 1 – нелегований полімер;
 2 – $\sigma = 20$ Снм/см;
 3 – $\sigma = 50$ Снм/см

і максимумом 1,9 еВ, що відповідає прямим міжзонним переходам, у легованих полімерах додається смуга поглинання в області 0,8 еВ, що виникає внаслідок появи солітонних станів усередині забороненої зони. За високих рівнів легування міжзонний пік поглинання зменшується, а солітонна лінія стає інтенсивнішою. У спектрах поглинання невироджених полімерів спостерігаються дві або три смуги поглинання, які зростають зі збільшенням рівня легування та відповідають переходам за участю поляронів і біполяронів.

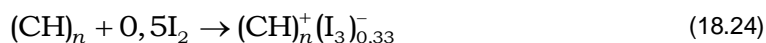
При легування також спостерігаються додаткові фононні моди в інфрачервоній області спектру, які зазвичай пов'язують з дисторсією одновимірної ґратки в результаті утворення солітону

■ 18.6. Електронна провідність спряжених полімерів

Розглядаючи елементарні збудження в полімерних напівпровідниках, ми частково торкалися проблеми електронного транспорту. Тепер детальніше обговоримо особливості електронного транспорту в спряжених полімерах. Провідні полімерні матеріали поділяються на дві великі групи – полімери з іонною провідністю (тверді полімерні електроліти), і полімери з електронною провідністю. Оскільки останні в основному виявляють напівпровідникові властивості, розглядатимемо тільки такі полімерні матеріали. Полімери з електронною провідністю, у свою чергу, поділяються на провідні полімери (органічні метали) та REDOX-полімери, в яких перенесення електронної густини відбувається, головним чином, за рахунок окислювально-відтворювальних реакцій між сусідніми фрагментами полімерного ланцюжка. Почнемо із провідних полімерів.

18.6.1. Провідність полімерів з хімічного погляду

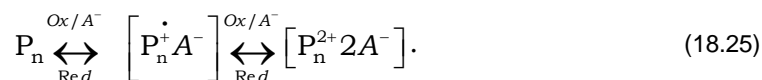
Першим провідним полімером, з якого почались систематичні дослідження провідності полімерних матеріалів, був, як зазначалось, поліацетилен. У вигляді твердих сріблястих плівок його вперше було синтезовано в 1974 році Хідекі Ширакава з ацетилену. Пізніше було встановлено, що часткове окислення поліацетилену молекулярним йодом



або іншими реагентами збільшує його провідність на дев'ять порядків.

У частково окисленій формі полімер можна розглядати як сіль, що містить позитивно заряджені іони полімеру та рівномірно розподілені вздовж його структури протиіони, що підтримують загальну електро-нейтральність системи.

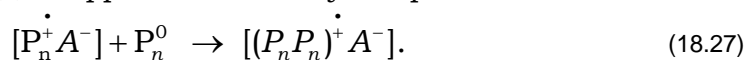
Визначимо P_n як фрагмент ланцюжка полімеру, що періодично повторюється та здатний до обміну повним зарядом з REDOX-партнером (або з електродом електрохімічної системи) і внаслідок цього переходить від одного REDOX-стану до іншого, тобто окислюється або відтворюється. Зазначимо, що структурна одиниця ланцюжку полімеру P , яка періодично повторюється, і фрагмент полімеру P_n не завжди ідентичні. Останній може утримувати в собі n структурних фрагментів P , при чому n може бути й нецілим числом. У результаті REDOX-реакції за участі P_n утворюється катіон (аніон) радикал, а потім, можливо й дікатіон (діаніон):



Катіон-радикальний фрагмент полімеру, може вступати в реакцію диспропорціювання



Реакція диспропорціювання залежить від температури та інших параметрів, у тому числі від природи та характеру взаємодії із протиіонами. Таким чином, в одному й тому самому полімері за рахунок реакцій диспропорціювання можуть зустрічатись різні REDOX-стани фрагментів. Іон-радикали та дійони, що утворюються в сполуках з розгалуженою системою π -зв'язків, здатні до утворення комплексів із перенесенням заряду з вихідними незарядженими сполуками (якщо цьому не заважають просторово-структурні фактори). Утворення подібного комплексу із двох фрагментів описують рівнянням:



Важливим є те, що електронні властивості полімерного матеріалу визначаються взаємодією як окремих фрагментів одного ланцюжка, так і між фрагментами різних ланцюжків. Коли електрон видаляється з валентної зони полімеру з системою спряжених зв'язків, вакансія, що виникає при цьому (дірка або катіон-радикал) не стає повністю делокалізованою, як це можна було б очікувати, базуючись на класичній зонній теорії. У полімерах відбувається лише часткова делокалізація, що розповсюджується на фрагмент P_n і викликає структурну деформацію цієї ділянки полімера. Рівень енергії, що відповідає катіон-радикалу, по суті являє собою рівень дестабілізованої зв'язуючої орбіталі, тобто розташовується вище за стелю валентної зони. Це підвищення енергії аналогічне збільшенню енергії зв'язуючої молекулярної орбіталі після видалення з неї електрона. З хімічного погляду поляроном є катіон-радикал, частково делокалізований по фрагменту полімеру. Цей катіон-радикал є нестабільним утворенням; для своєї стабілізації він поляризує оточуюче середовище. Якщо наступний електрон видаляється із вже окисленого полімерного ланцюжка, що містить полярон, то може відбуватись наступне: електрон видаляється з іншої ділянки полімерного ланцюжка та виникає новий, слабко взаємодіючий із першим полярон. Або електрон видаляється з тієї самої ділянки полімерного ланцюжка, де вже існує полярон. При цьому утворюється подвійний катіон, який є ні чим іншим, як біполяроном. Таким чином, спочатку процесу окислення в основному утворюються полярони. За подальшого збільшення окисленості полімеру відбувається утворення біполяронів. Виникнення біполяру також спричиняє структурну деформацію середовища. Як вже зазначалось, полярони та біполярони мобільні, і під дією електричного поля

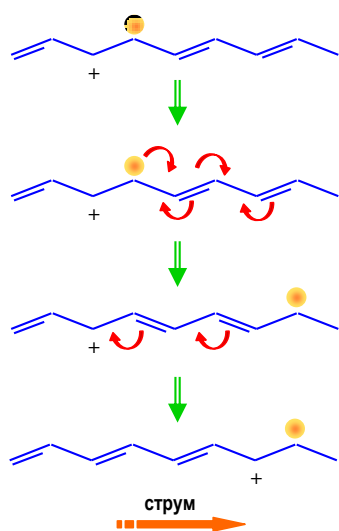


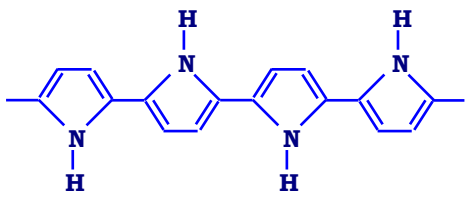
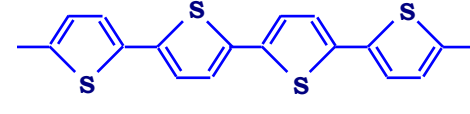
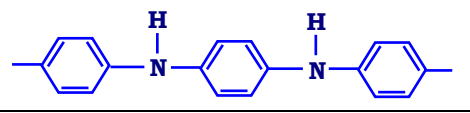
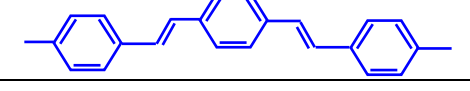
Рис. 18.24. Утворення діркового струму в полімерному ланцюжку шляхом реорганізації одиничних і подвійних зв'язків

вони здатні рухатись уздовж полімерного ланцюжка шляхом реорганізації подвійних та одиничних зв'язків у спряженій системі. На рис. 18.24 схематично показано, як шляхом реорганізації подвійних та одиничних зв'язків у поліацетилені відбувається рух полярона під дією електричного поля, тобто реалізується електронний транспорт. Розгляд електронного транспорту з хімічного погляду дає наглядну уяву процесу формування в окислених полімерах діркової провідності. З іншого боку стає більш зрозумілим, що поняття "дірка" у фізиці полімерів є тільки аналогом цього поняття в кристалах та описує рухливу область локалізації позитивного заряду (відсутність електрона, що в процесі окислення був зі зв'язку в полімерному ланцюжку переведений до окислювача).

У табл. 18.1 подано структурні формули деяких розповсюджених полімерів і значення їхньої провідності. Зважаючи на те, що провідність міді – 10^6 Сим/см, а власного кремнію – 10^{-2} Сим/см, ми бачимо, що провідність полімерів займає проміжну нішу між провідністю металів і діелектриків. Поляронна концепція провідності нині досить широко використовується в поясненні переносу заряду в спряжених полімерах. Ідеалізована поляронна модель дуже далека від ситуації, що існує в реальних полімерних матеріалах. Але вона дає основу для розуміння процесів, що відбуваються в існуючих матеріалах і, головне – може бути основою для побудови теорії електронного транспорту в полімерах, що дає можливість передбачення багатьох властивостей полімерів.

Таблиця 18.1

Назва полімеру	Структурна формула	Провідність (Сим/см)
поліацетилен		10^4
поліфенілен		10^3

поліпірол		10^3
політіофен		10^3
поліанілін		10^2
поліфенілен-венілен		10^4

18.6.2. Провідність REDOX-полімерів

Більшість відомих REDOX⁴-полімерів являють собою металокомплексні сполуки, які, в основному, отримують шляхом електрохімічної полімеризації вихідних мономерних комплексних сполук. Як приклад такого REDOX-полімера можна навести поліпірідінові комплекси, що мають склад $\text{poly}[\text{Me}(\nu\text{-bpy})_x(\text{L})_y]$. У цих комплексах зазвичай використовують метали: $\text{Me} = \text{Co}, \text{Fe}, \text{Ru}, \text{Os}$. Лігандами L можуть бути такі: $\nu\text{-bpy}$ (4-вініл-4'-метил-2,2'-біпірідін), фенатролін-5,6-діон, 4-метилфенантролін тощо ($x + y = 3$). Зшивання вихідних молекул відбувається за трьома просторовими напрямками відповідно до їх геометрії. У результаті такого синтезу полімер набуває тривимірної структури. REDOX-центрами (атомами, що беруть участь у окислювально-відновлювальних реакціях у полімері) є іони металу, що можуть перебувати в різних зарядових станах, тому комплекси металів, що мають тільки один можливий зарядовий стан (напр., цинк, кадмій), не утворюють REDOX-полімерів. Необхідною умовою провідності REDOX-полімерів є наявність у лігандному оточенні комплексів розгалуженої системи спряжених π -зв'язків, що служать провідними місцями між REDOX-центрами.

⁴ Детальніше про окислювально-відновлювальні реакції (REDOX) див. у додатку L.

Якщо REDOX-полімер повністю окислений або відновлений, тобто всі його REDOX-центри перебувають в одному зарядовому стані, перенесення заряду вздовж ланцюжка полімеру неможливе та його провідність близька до нуля. Якщо REDOX-центри мають різний зарядовий стан, між ними можливий обмін електроном. Таким чином, провідність REDOX-полімерів пропорційна константі швидкості самообміну електроном між REDOX-центрами (k_{∞}) та добутку концентрацій окислених n_{OX} і відновлених n_{RED} центрів у полімері

$$\rho_{REDOX} \sim k_{\infty} n_{RED} n_{OX}. \quad (18.28)$$

У такому випадку провідність буде максимальною, коли концентрації окислених і відновлених REDOX-центрів є однаковими, оскільки в цьому випадку перенесення зарядів відбувається найефективніше. Перехід молекул REDOX-полімерів з окисленого до відновленого стану супроводжується виходом зарядкомпенсуючих протиіонів із полімеру в розчин електроліту, де відбувається REDOX-процес, і навпаки. У цілому, провідність REDOX-полімерів суттєво менша від провідності органічних металів і становить $\sim 10^{-3} - 10^{-2}$ Сим/см. Це пов'язано з багатьма причинами, однією з яких є обмеженість швидкостей REDOX-реакцій. Однак, поряд із цими недоліками, REDOX-полімери мають істотні переваги. По-перше, металокомплекси у цілому порівняно з органічними сполуками, мають значно ширші можливості керування властивостями полімерів за рахунок варіювання природи (електронної структури) метала, що відіграє роль REDOX-центру. Як один із чисельних прикладів можна навести суттєве розширення числа оптичних електронних переходів у металокомплексах порівняно з органічними сполуками за рахунок появи переходів ліганд-метал та інтервалентних переходів. Це відкриває нові можливості з керування оптичними властивостями полімерів за допомогою зовнішнього електричного поля. По-друге, деякі із металокомплексів є аналогами найважливіших природних ферментів (біокатализаторів), зокрема оксигенази та пероксидази. Це дозволяє прогнозувати створення на їхній основі нових полімерних електрокаталітичних систем з високою ефективністю роботи (напр., для біокатализаторів).

Таким чином, короткий огляд властивостей органічних напівпровідників свідчить про розгортання широких досліджень органічних напівпровідників як у царині синтезу нових матеріалів і вивченні їх фундаментальних властивостей, так і в напрямку їх практичного застосування.

■ 18.7. Застосування

молекулярних напівпровідників

У 2002 році корпорація Hewlett-Packard продемонструвала лабораторний зразок 64-бітового запам'ятовуючого пристрою (ЗП) площиною 1 мкм^2 , в якому активними елементами виступають молекулярні ключі. Структура ЗП подібна до структури магніторезистивної пам'яті та складається із подвійної сітки перехрещуваних провідників, між якими магніторезистивний шар замінений шаром органічних молекул. За прикладення напруги до точок перетину провідників опір матеріалу змінюється. Для зчитування цієї зміни достатньо невеликої напруги. На тестовому чипі ЗП величина зміни опору становила 10^4 , що суттєво перевищує зміну опору в магніторезистивному пристрої, де вона становить 40 %. Зразок чіпу методом оптичної та електронно-променевої літографії було виготовлено протягом одного дня. Робота пристрою продемонструвала, що органічні молекули можуть працювати як діоди, ключі та конденсатори й базові молекулярні компоненти можна зібрати в ЗП і логічні схеми з високою щільністю елементів. Однак прорив у молекулярній електроніці відбудеться лише тоді, коли переконливо буде продемонстровано здатність органічних матеріалів задовольняти мінімальному набору стандартів, що необхідні для функціонування електронних приладів – стійкість до дії високотемпературних технологічних процесів, здатність працювати при температурах до 140°C і після 10^{12} циклів запису–зчитування.

Однак використання процесів самозбирання, що притаманні для органічних матеріалів, можуть суттєво послабити ці вимоги, тому перспектива застосування молекулярних напівпровідників стає близькою та актуальною. Для елементів пам'яті комп'ютерів розробляються запам'ятовуючі середовища на основі REDOX-молекул, що наносяться на поверхню кремнію й таким чином можуть успішно спрягатись з існуючими сучасними системами електроніки. Основною ідеєю зберігання інформації у середовищах на основі REDOX-молекул є фіксація визначеного REDOX-стану окремих елементів речовини. Як активний елемент пам'яті вибирається, зокрема порфірин, оскільки він характеризується стабільним дво- та чотирибітовим REDOX-станами (залежно від просторової будови порфірина). Крім того, ці стани можуть зчитуватись за відносно низьких потенціалів (менших за 1,5 В). Час зберігання заряду в елементах ЗП на основі порфірина становить хвилини, що є набагато більшим порівняно з мілісекундами твердотільних напівпровідникових динамічних ЗП із довільною вибіркою – DRAM. Усі ці властивості порфіринів дозволяють створювати ЗП із великою щільністю пам'яті та низьким рівнем енергоспоживання, що

досягається за рахунок низьких потенціалів і тривалим часом зберігання зарядів.

Органічні світловипромінюючі діоди (OLED) та плоскочасельні дисплеї (FPD) є прикладами широкого практичного застосування органічних напівпровідників у наш час. Принципи роботи OLED є простими. Тонка плівка органічної речовини контактує з благородним металом – з одного боку та звичайним металом – з іншого. Коли до структури прикладається електрична напруга, із благородного металу (Me1) в органічний напівпровідник інжектуються дірки, а зі звичайного металу – електрони. Електрони та дірки, рухаючись в органічній речовині, зустрічаються та, рекомбінуючи, випромінюють фотон заданої частоти (рис. 18.25). Частіше використовують ефективнішу схему OLED, в якій між двома металічними плівками розташовано двошарову структуру органічного напівпровідника (рис. 18.26). Така структура за рахунок наявності енергетичного бар'єру для електронів і дірок усередині органічного шару різко підвищує ймовірність одночасного перебування в тонкому шарі поблизу бар'єру електронів і дірок, що спричиняє підвищення інтенсивності рекомбінаційних процесів і, тим самим – підвищення ефективності роботи пристрою. Як матеріал в OLED і дисплеях використовуються як невеликі органічні молекули, так і полімери. Технологія виготовлення світловипромінюючих пристроїв на основі органічних напівпровідників є досить дешевою. У результаті такі пристрої поступово завойовують світ. Такі дисплеї характеризуються високою якістю передачі кольорів і високою роздільною здатністю.

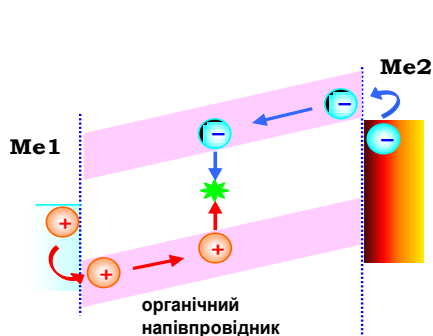


Рис. 18.25. Енергетична схема OLED

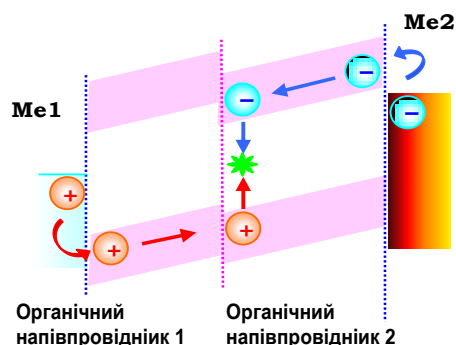


Рис. 18.26. Енергетична схема OLED із двошаровою структурою органічного напівпровідника

■ 18.6. Завдання

1. Сформулюйте циклічні граничні умови для полімерних ланцюжків.
2. У якого полімеру провідність вища за однакових умов – з довгими чи короткими ланцюжками.
3. Чому за окиснення полімеру у нього виникає помітна провідність?

Список літератури

1. *Богуславский Л.И., Ванников А.В.* Органические полупроводники и биополимеры. – М.: Наука, 1968.
2. *Гутман Ф., Лайонс Л.* Органические полупроводники. – М.: Мир, 1970.
3. *Hadis Morkos.* Advanced Semiconductor and Organic Nano-Techniques. Part I: Nanoscale Electronics for Computers and Optoelectronics for Telecommunications.– N.Y.: Elsevier, 2003.
4. *Farchioni R., Grosso G.* (Eds.) Organic Electronic Materials: Conjugated Polymers and Low Molecular Weight Organic Solids.– Berlin, Heidelberg, N.Y.: Springer, 2001.
5. *Adachi Ch., Oyamada T.* Data book on HOMO Levels of Organic Thin Films in Organic Semiconductor Devices. – CMC Publish., 2006.

Розділ 19

ВЛАСТИВОСТІ ЕКЗОТИЧНИХ СИСТЕМ

Розглядаючи властивості напівпровідників, ми бачили, що розвиток науки про напівпровідники йде двома шляхами. Це вивчення властивостей:

- ♦ матеріалів;
- ♦ структур і систем.

При цьому ми підкреслювали, що останнім часом процес створення структур перетворився на процес створення нових штучних матеріалів, оскільки перехід до мікронних і наномасштабів спричинив виникнення в напівпровідникових і металічних структурах нових властивостей, що зіставні з основними властивостями напівпровідникових матеріалів. Таким чином, у наш час по суті створюються нові штучні матеріали як на основі твердотільних наноструктур, так і на основі молекулярних напівпровідників і металів. Однією з цікавих особливостей подібних структур є вплив на характеристики структури стану одного електрона.

■ 19.1. Одноелектронні ефекти в напівпровідникових структурах

Вивчаючи особливості квантового транспорту в квазіодновимірних структурах, ми бачили, що важливу роль у формуванні електропровідності в таких системах відіграють закони статистики електронів. Зрозуміло, що використання ефектів квантового транспорту можуть спричинити не тільки різке зменшення енерговитрат електронних пристроїв, але завдяки специфічним ефектам, що пов'язані з квантовим рухом електронів – появу нових властивостей електронних пристроїв.

Перехід до одноелектронних ефектів, крім всього іншого, викликає цілу низку питань, що не можуть бути розв'язані в межах стандартних підходів. Зокрема, на перший план виходять проблеми флуктуацій струму і просторового квантування. Звичайний макроскопічний струм флуктує випадково так, що поява у електричному ланцюгу нового елементарного заряду не скорельована з появою попереднього. Такі флуктуації є причиною так званого дробового шуму. Ці флуктуації описуються пуассонівською статистикою. Якщо в електричному ланцюгу джерело підтримує сталий електричний струм, коли через одиничний переріз проходить n_0 зарядів у секунду, то у середньому за час t ланцюгом проходить $N = n_0 t$ зарядів. При цьому значення цієї

величини буде флукутувати із середньоквадратичним відхиленням $\Delta N \sim (n_0 t)^{1/2}$. Абсолютна потужність дробового шуму зростає зі зростанням потужності самого сигналу, однак відносна потужність $\Delta N/N$ спадатиме. Із цієї причини в макрофізиці квантуванням заряду зазвичай нехтують, оскільки за великих значень струмів відносні флукутації дуже незначні. Якщо сигнал представити кількістю електронів у зарядовому пакеті, то кількість класичної інформації, що передається струмом за час t , із врахуванням дробового шуму матиме порядок $\log_2(1 + \Delta N/N) = \log_2(1 + (n_0 t)^{1/2})$.

Логічні елементи, що спрацьовують на визначену величину зарядового пакету, як, наприклад у випадку КМОН-схем, часто припускаються помилки, якщо пакет виявиться недостатньо великим. Наприклад, при кодуванні логічної одиниці пакетом із десяти електронів із порогом спрацювання у п'ять зарядів, логічний елемент буде неправильно спрацьовувати у $\sim 3\%$ випадків. Іншими словами, згідно з пуассонівською статистикою, у трьох випадках зі ста у зарядовому пакеті буде знайдено менш ніж п'ять електронів. При цьому суттєво збільшити надлишковий заряд, що зберігається в структурі манометрових розмірів, неможливо. Наприклад, на сферичному кластері радіусом 2–3 нм можна розташувати лише кілька зайвих електронів.

Інший аспект проблеми одноелектронної фізики полягає у наявності деяких енергетичних обмежень. Для розуміння базових принципів одноелектроніки розглянемо рис. 19.1. Нехай маленька частинка є електронейтральною, тобто кількість електронів у частинці (m) компенсується точно такою самою кількістю протонів у кристалічній ґратці. У цьому випадку слабка зовнішня сила \mathbf{F} здатна піднести до частинки та розмістити на ній додатковий електрон ззовні. Іншими словами, відбудеться інжекція носія через енергетичний бар'єр на поверхні частинки. Частинка зарядиться до величини $-e$, та електростатична сила $-e\mathbf{E}$ протидіятиме зовнішній силі помістити в частинку ще один електрон. Незважаючи на те, що заряд електрона дуже малий, електростатична сила, що виникає, може досягати досить значної величини, оскільки на масштабах порядку розміру частинки електричне поле буде достатньо великим. Наприклад, поблизу поверхні сферичної частинки радіусу 10 нм електричне поле досягатиме значної величини $\sim 1.4 \cdot 10^5$ В/см.

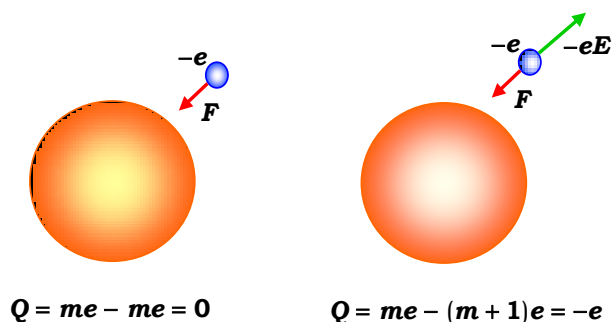


Рис. 19.1. Взаємодія одиничного електрона:
а – із незарядженою, б із зарядженою наночастинкою

У теорії одноелектронних явищ використовується більш адекватний параметр, ніж величина електростатичного поля – додаткова енергія, пов'язана із зарядом частинки

$$E_C = e^2 / C, \quad (19.1)$$

де C – ємність частинки. У випадку, коли розміри частинки зіставні із довжиною хвилі де Бройля електрона, провідну роль починають відігравати квантування енергії. У цьому випадку зарядові ефекти даються більш загальною величиною додаткової енергії електрона E_a . У більшості випадків ця енергія може бути апроксимована простою сумою

$$E_a = E_C + E_k, \quad (19.2)$$

де E_k – квантова кінетична енергія додаткового електрона. Для виродженого електронного газу цю енергію можна оцінити як $E_k = 1/g(E_F)V$, де V – об'єм частинки, $g(E_F)$ – щільність станів на ізоенергетичній поверхні $E = E_F$ (поверхні Фермі). Для частинок із лінійними розмірами ~ 100 нм енергія E_a в основному визначається енергією E_C і становить за порядком величини 1 меВ, або 10 К в одиницях температури. Оскільки термічні флуктуації подавляють одноелектронні ефекти для енергій E_a , що не перевищують 10 кТ, ці ефекти можна спостерігати лише в області субтемператур. З іншого боку, для частинок із лінійними розмірами, меншими 10 нм, додаткова енергія E_a збільшується на порядок і становить ~ 100 меВ; деякі одноелектронні ефекти можуть спостерігатись за кімнатної температури. Однак для побудови одноелектронних пристроїв часто необхідно виконання нерівності $E_a \geq 100$ кТ. Це є необхідною умовою, наприклад для виключення ефектів термічно індукованого хаотичного тунелювання. Такі вимоги потребують для спостереження одноелектронних ефектів за кімнатних температур величин E_a порядку кількох електрон-вольт, що означає вимогу зменшення лінійних розмірів частинки до величини

1 нм і менше. Крім низьких температур, іншою проблемою є ємність – вона складається із ємності не тільки самого конденсатора, але всіх контактів та об'єктів, що утворюють та оточують установку. Для запобігання негативному впливу додаткових ємностей було запропоновано створювати систему, що складається не з одного, а з кількох

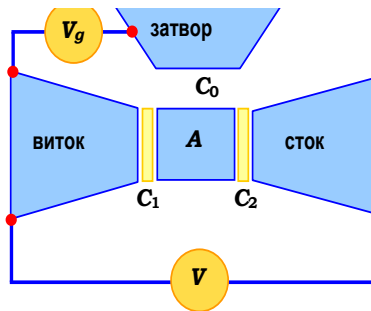


Рис. 19.2. Схема одноелектронного транзистора

подачу керуючої напруги на затвор. При цьому струм у приладі формується потоком одиничних електронів. Цей прилад можна розглядати як граничний випадок польового транзистора з ізольованим затвором. Але в такому пристрої не спостерігатиметься підсилення потужності, тобто описаний прилад скоріше є одноелектронним ключем.

Таким чином, для спостереження одноелектронних струмів необхідно, по-перше, мати справу з частинками дуже невеликих розмірів, а по-друге, – проводити вимірювання в умовах низьких температур. Для спостереження ефектів одноелектронного транспорту запропоновано використовувати систему квантових точок у матриці. Наприклад, це може бути така система: виготовляється підкладка (травленням верхнього шару кремнію в SIMOX-структурі до величин ~ 30 нм), на якій вирощується методом газофазної епітаксії тришарова структура Si/SiGe/Si. Далі методом електронної літографії нарізуються нано-дротинки шириною 20–50 нм і довжиною 50–120 нм. Потім витравлюється шар SiO₂, розташований нижче, а "місточки" SiGe повністю окислюються. При цьому германій збирається у квантові точки діаметром кілька нанометрів за щільності $\sim 10^{11}$ см⁻² та утворюється структура, яку схематично подано на рис. 19.3 а. Ця технологія була застосована в одній з лабораторій на Тайвані. На такій структурі було виміряно вольт-амперні характеристики (за $T = 300$ К) залежно від затворної напруги. Як видно із рис. 19.3 б, ці характеристики демонструють яскраво виражені кулонівські осциляції, що означає наявність одноелектронного струму в досліджуваних структурах.

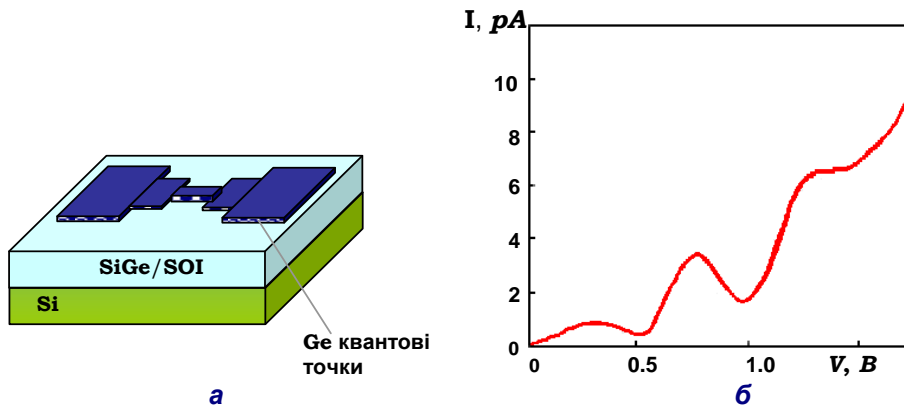


Рис. 19.3. Залежність струму стоку від напруги на затворі (б) у системі квантових точок структури на панелі (а). Напруга стоку $V_{dr} = 0.05$ В, температура $T = 300$ К

■ 19.2. Вуглець – майбутнє електроніки?

Вуглець є одним з найбільш цікавих елементів періодичної системи. Він може конденсуватись у тверде тіло з великою кількістю алотропних форм (рис. 19.4). Наприклад, тривимірна алотропія вуглецю є найтвердішим матеріалом у природі – алмазом. Алмаз є широкозонним напівпровідником із шириною забороненої зони $E_g = 5$ еВ. З іншого боку, вуглець конденсується в іншу тривимірну алотропну форму – графіт – властивості якого суттєво відрізняються від алмазу. Широко відомою квазі-нульвимірною формою вуглецю є фулерен. Останнім часом інтенсивно досліджується одновимірна алотропна форма вуглецю – нанотрубки. Але найбільш екзотичні властивості як з погляду фундаментальної науки, так і щодо можливих прикладних застосувань демонструє двовимірна алотропна форма вуглецю – графен.

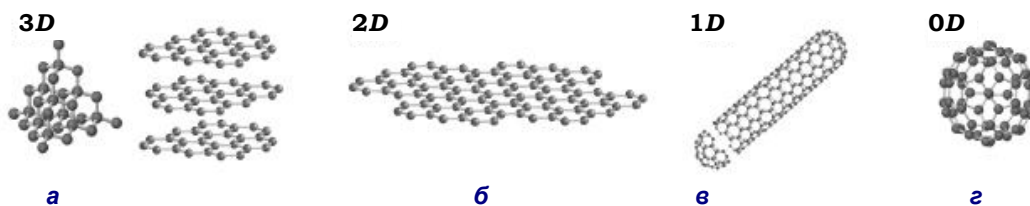


Рис. 19.4. Алотропні форми вуглецю: а – тривимірні, алмаз і графіт; б – двовимірна, графен; в – доновимірна, нанотрубка; г – нульвимірна, фулерен.

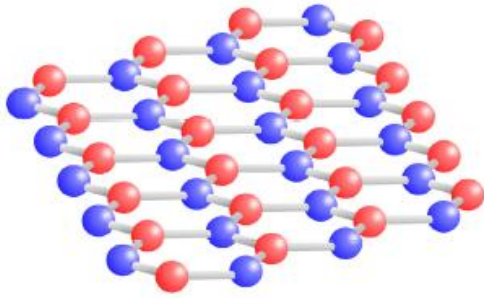


Рис. 19.5. Кристалографічна структура графену.
Сині та помаранчеві кульки належать до різних підґраток

структуру, яка складається із двох підґраток тригональної симетрії. Отриманий таким чином матеріал – графен – є стабільним, має добрі механічні властивості (гнучкий і надзвичайно міцний), досить високу електропровідність. Завдяки унікальним властивостям вуглецю в просторовій ґратці графену електронні стани характеризуються високою мобільністю електронів. Крім того, зонна структура графену характеризується наявністю у законі дисперсії точок K і K' (рис. 19.6), що відповідають електронам з нульовою ефективною масою. Поблизу цих точок безмасові електрони поведуться не як звичайні фермі-частинки, а як ферміони Дірака, що описуються релятивістськими рівняннями та характеризуються лінійною залежністю енергії від швидкості. З погляду електронних властивостей графен є напівпровідником із нульовою забороненою зоною, в якому низькоенергетичні збудження у кожній долині формально описуються гамільтоніаном діраківського типу

$$\hat{H} = \hbar v_F \begin{pmatrix} 0 & k_x - ik_y \\ k_x + ik_y & 0 \end{pmatrix} = \hbar v_F \boldsymbol{\sigma} \cdot \mathbf{k}, \quad (19.3)$$

де \mathbf{k} – імпульс частинки, $\boldsymbol{\sigma}$ – двовимірний матриця Паулі, а незалежна від \mathbf{k} швидкість Фермі v_F відіграє роль швидкості світла в квантовій електродинаміці. Рівняння Дірака є безпосереднім наслідком кристалічної симетрії графену, зокрема подвійної стільниковидної ґратки, що складається з двох еквівалентних підґраток A та B . Енергетичні зони

Якщо відділити один атомарний шар від кристалу графіту, можна отримати двовимірну структуру, товщина якої становить один атом (рис. 19.5). При цьому в площині шару зв'язки між атомами утворюють таку саму гексагональну упаковку, як і в шарі графіту. Як відомо (див. розд. 1), гексагональна ґратка не є ґраткою Браве. Це означає, що вуглець утворює кристалографічну

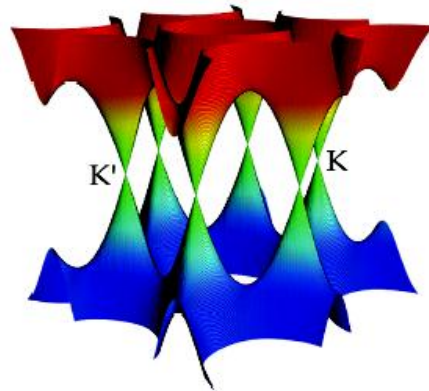


Рис. 19.6. Зонна структура графену.
Зона провідності та валентна зона торкаються у точках K та K'

пов'язуються з перетинами підґраток за нульової енергії поблизу країв зони Бриллюена. На тривимірному графіку закону дисперсії (рис. 19.6) – це конусоподібні області поблизу точок K і K' . Варто підкреслити, що лінійний закон дисперсії $E = \hbar v_F k$ не єдина суттєва риса зонної структури. Дійсно, електронні стани поблизу нуля E (де зони перетинаються) компонується зі станів, що належать до різних підґраток, і їх відносні внески до створення квазічастинок мають братись до уваги через застосування, наприклад двокомпонентних хвильових функцій (спінорів). Компоненти спінорів розрізняються індексами, що відповідають підґраткам A та B . Ці індекси відповідні спіновим індексам у квантовій електродинаміці, і називаються псевдоспінами. Відповідно, за формальним розв'язком рівняння Дірака для визначення станів квазічастинок графену, у гамільтоніані графена (19.3) величина σ відповідає псевдоспіну. Аналогічно до того, як це робиться у квантовій електродинаміці, при розгляді електронних властивостей графену можна ввести поняття хіральності, що формально є проекцією σ на напрямок руху \mathbf{k} , позитивною для електронів і негативною для дірок. Важливою рисою хіральності в графені є те, що стани k -електронів і k -дірок внутрішньо пов'язані, що є віддзеркаленням їхнього походження зі станів еквівалентних вуглецевих підґраток. Зауважимо, що концепція псевдоспіна та хіральності відіграють важливу роль у визначенні електронних властивостей графену – ці величини підкоряються фундаментальним законам збереження.

Головні зусилля з експериментального вивчення електронних властивостей графену було спрямовано на розуміння ефектів, що визначаються особливостями спектрів квазічастинок, і схожі на спектри релятивістських електронів у квантовій електродинаміці. Такими ефектами є хіральний квантовий ефект Холла та мінімум квантової провідності за зменшення до нуля концентрації носіїв заряду та сильного придушення ефектів квантової інтерференції.

На рис. 19.7–19.8 подано різну поведінку квантового ефекту Холла в графені. Перший тип (рис. 19.7) є релятивістським аналогом цілочисельного квантового ефекту Холла – він спостерігається у одношаровому графені. Зокрема, тут показано що недиагональна провідність являє собою еквідистантні сходинки, що проходять через нейтральну – Діраковську - точку зміни знаку носіїв заряду. Висота сходинок визначається формулою

$$\sigma_{xy} = \pm(4e^2/h)(N + 1/2), \quad (19.4)$$

де N – номер рівня Ландау. Фактор 4 з'являється через подвійне виродження долин і спінове виродження електронних станів. Таким чином, квантовий ефект Холла у графені є квантовим ефектом Холла

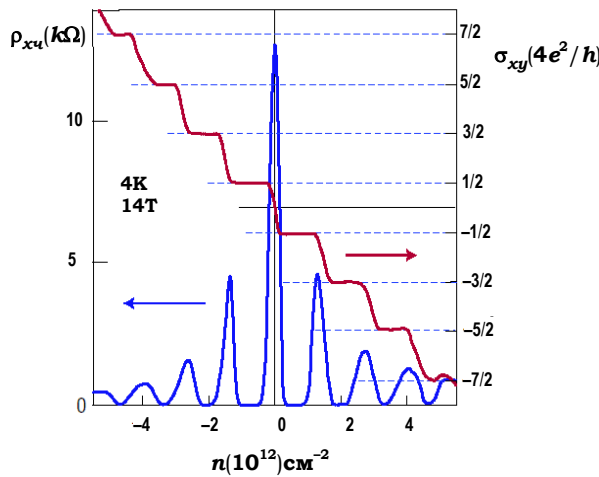


Рис. 19.7. Хіральний квантовий ефект Холла.
Виявом квантового ефекта Холла безмасових ферміонів Дірака є наявність плато в недиагональній провідності σ_{xy} за напівцілих значень величини $4e^2/h$

чиняє появу геометричного фазового множника в імпульсі вздовж напрямку циклотронних траєкторій. Інші два типи поведінки квантового ефекта Холла масивних дираківських ферміонів демонструють подвійні шари графену.

На рис. 19.8 подано поведінку σ_{xy} у графені за температури 4К під дією сильного (12 Тесла) магнітного поля. Ця поведінка є стандартною з плато, що відповідають цілочисельним значенням $N (4e^2/h)$, за винятком $N = 0$ (помічено червоною стрілкою). Хімічне допування графену спричиняє зсув нейтральної точки до області більших значень напруги на затворі V_g . У результаті виникає асиметрична щілина ($\sim 0,1$ еВ) у спектрі електронних станів, що відкривається завдяки ефекту поля. Наявність цієї щілини є причиною появи плато за $N = 0$.

порядку $1/2$ і в цьому сенсі не є звичайним. Його специфіка визначається тим, що за наявності магнітного поля закон дисперсії носіїв визначається формулою

$$E = \pm v_F \hbar 2eB N, \quad (19.5)$$

де плюс відповідає електронам, а мінус – діркам. Існування аномального квантового ефекта Холла у графені можна пояснити квантуванням рівнів за нульової енергії. Альтернативним поясненням напівцілого квантового ефекта Холла у графені є наявність зв'язку між псевдоспіном та орбітальним рухом носіїв, що спри-

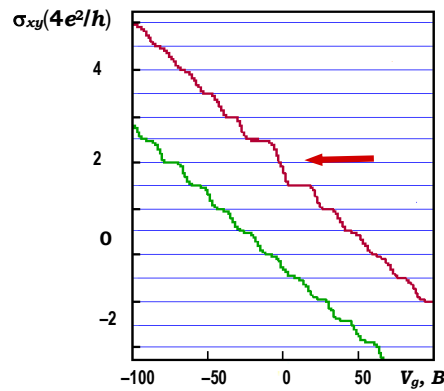


Рис. 19.8. Аномальний КЕХ у подвійному шарі недопованого (червоний) та допованого (зелений) графену

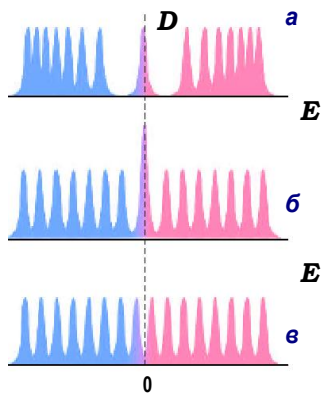


Рис. 19.9. Щільності станів за різних типів квантування Ландау

Для масивних ферміонів Дірака, що існують у подвійному шарі графену в точках K та K' , $E_N \propto [N(N-1)]^{1/2}$. Хімічне легування спричиняє появу енергетичної щілини у спектрі ферміонів Дірака, а залежність енергії від квантового числа набуває стандартного вигляду $E_N \propto N+1/2$.

Різні типи квантування Ландау у графені визначаються особливістю щільності станів (рис. 19.9). Для безмасових ферміонів Дірака, де залежність енергії за квантування Ландау $E_N \propto (N)^{1/2}$ (рис. 19.9 а) квантовий ефект Холла є особливим, і плато в поведінці недіагональної компоненти провідності спостерігаються за напівцілих N . Такі ферміони, як зазначалось вище, спостерігаються в одношаровому графені. Двошаровий графен демонструє інший тип поведінки квантового ефекту Холла, що характеризується відсутністю плато за $N=0$. На рисунку подано рівні Ландау у такому випадку.

■ 19.3. Ефекти кулоновської блокади у структурі метал-фулерен-метал

Раніше ми зазначили: якщо частинка велика, то тунельний струм протікає за будь-якої різниці потенціалів. Якщо ж частинка мала, стає важливим той факт, що струм переноситься окремими електронами. Якщо молекула (ємність якої C є порядку її радіусу r) перебуває у діелектричному середовищі між двома металевими контактами, то перенесення одного електрона потребує енергії $W_e = e^2/2C$ або $W_e = e^2/2r$. У цьому випадку порогове поле для перенесення одного електрона (і появи одноелектронного струму) відповідатиме різниці потенціалів $V_e = e/2C$, для перенесення двох електронів – $V_{2e} = e/C$, трьох електронів – $V_{3e} = 3e/2C$. Вольт-амперна характеристика виглядатиме так: зона мовчання за $V < V_e$, і сходинка струму за $V \sim V_e$, $V \sim V_{2e}$, $V \sim V_{3e}$ і так далі. Як зазначалось раніше, у макрофізиці такі ефекти можуть спостерігатися за досить низьких температур, тобто $T_C \ll W_e$. Зрозуміло, що для спостереження одноелектронних ефектів кулоновської блокади за кімнатних температур потрібно, щоб розміри частинок були невеликими становили $\sim 1-5$ нм. На рис. 19.10 подано вольт-амперну харак-

теристичну одноелектронного пристрою, що являє собою молекулярний кластер розміром 2 нм між металічними контактами. Видно чітко виражені особливості типу кулоновських сідців із характерним масштабом за напругою $\sim 0,1$ В.

Зменшення розміру частинок, на яких може формуватись одноелектронний транспорт, наводить на думку про використання як наночастинки однієї органічної молекули або супрамолекулярного комплексу, наприклад C_{60} . Для вивчення транспорту крізь одну молекулу C_{60} зручно використовувати метод тунельної спектроскопії, застосовуючи скануючий тунельний мікроскоп (СТМ). Як відомо, основним процесом у СТМ є квантово-механічне тунелювання електронів з голки зонда у підкладку (як таку варто використовувати провідний матеріал із тонким молекулярним шаром на його поверхні). Тунельний струм, що протікає через молекулу та кінчик голки, визначається перекриттям атомних орбіталей металічної голки та молекулярних орбіталей молекулярного шару. Для одновимірного випадку струм через контакт молекула-голка зазором s визначається формулою:

$$I \approx \exp(-2ks), \quad (19.6)$$

де константа k визначає затухання хвильових функцій голки та зразку. Для матеріалів із великою роботою виходу цей коефіцієнт має $\sim 0,1 \text{ нм}^{-1}$, тобто тунельний струм значно зменшується за збільшення зазору між голкою та молекулою на один ангстрем. Спеціальним п'єзопристроєм положення голки контролюється із точністю 0,01 нм як у площині сканування, так й у вертикальному напрямку (перпендикулярному до площини поверхні зразку). Таким чином, використовуючи

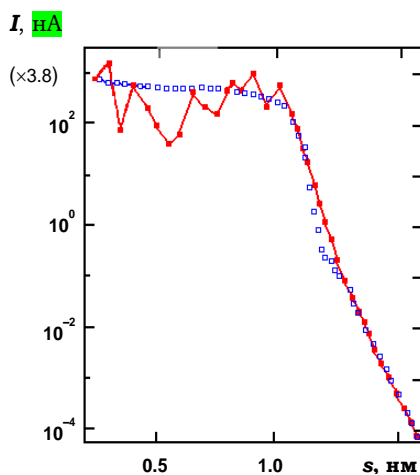


Рис. 19.10. Залежність струму у структурі Au- C_{60} -СТМ від зазору s

СТМ, можна вивчати транспорт крізь окрему молекулу. На рис. 19.10 подано залежність струму, що проходить через структуру металічна (Au) підкладка-молекула фулерену-голка СТМ від величини зазору s . За великих значень s струм має тунельний характер і залежить від зазору s згідно із (19.6). Коли відстань між голкою та підкладкою досягає значення $s = 1,24$ нм, установлюється контакт між молекулою та голкою. Подальше зменшення відстані s спричиняє деформацію молекули та зсув молекулярних рівнів до рівня Фермі металу. При цьому тунелювання стає резонансним та опір наближується до квантової границі $h/2e^2 \approx 12$ кОм.

Таким чином, транспорт у розглянутих системах визначається характером розподілу енергетичних рівнів молекули. Зміна положення молекулярних енергетичних рівнів за рахунок тих чи інших факторів приводить до різких змін характеру транспорту, що лежать в основі побудови одноелектронних приладів, приклади яких наведено далі. Принцип зміни опору молекули фулерену під дією слабкого тиску голки СТМ покладено до принципу будови електромеханічного транзистора. Зміна провідності вуглецевої нанотрубки під дією зовнішнього поля покладено в основу іншого типу молекулярного транзистора. На рис. 19.11 подано схеми молекулярного транзистора на основі молекули фулерену та нанотрубки.

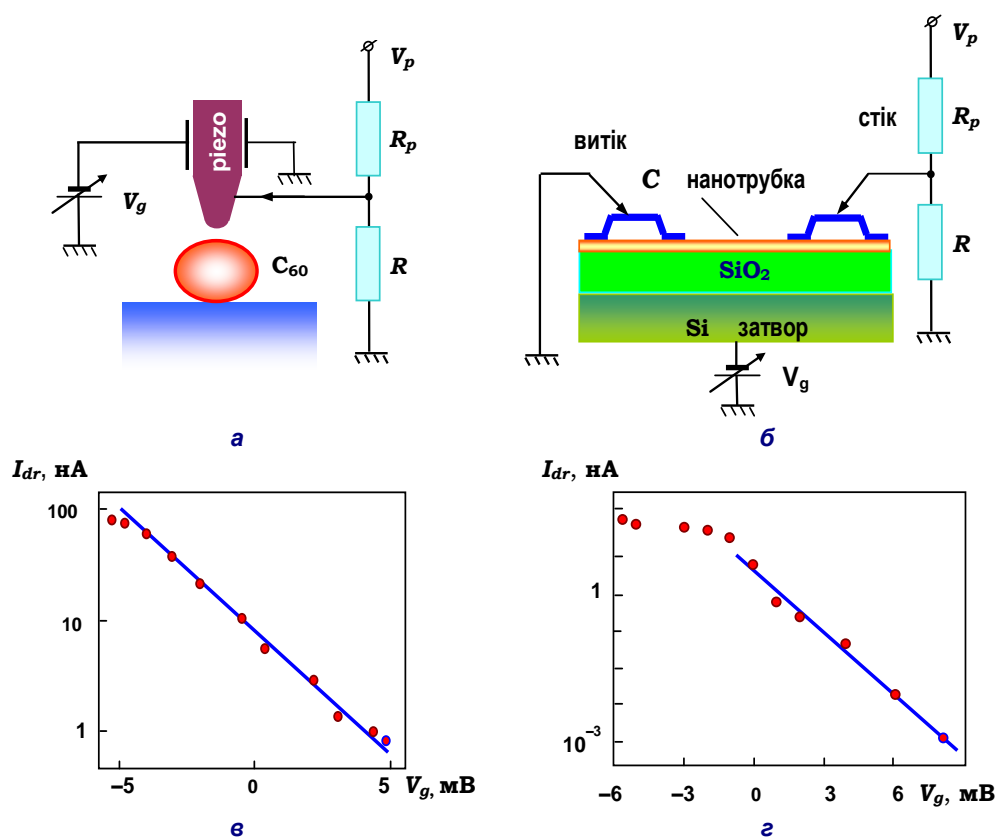


Рис. 19.11. Електромеханічний транзистор на основі молекули фулерену (а) та залежність струму стоку цього транзистора від напруги на затворі (с). Польовий транзистор на основі одностінкової вуглецевої нанотрубки (б) та залежність струму стоку від напруги на затворі (д)

Управляючим затвором електромеханічного транзистора, що побудований на основі молекули фулерену (рис. 19.11 а), є напруга на п'є-

зоелементі, що з'єднаний із голкою зонда СТМ. Зміна цієї напруги (V_g) приводить до зміни струму стоку (рис. 19.11 в).

Інший транзистор побудований на основі одностінкової напівпровідникової вуглецевої нанотрубки, діаметром 1 нм, що розташована на кремнієвій підкладці. Провідність нанотрубок суттєво залежить від їх структурних особливостей, зокрема від кута закручування хімічних зв'язків між атомами вуглецю відносно осі нанотрубки. Основою цього транзистора є напівпровідникова нанотрубка, розташована поперек металічних платинових контактів, що нанесені на поверхню оксиду кремнію. Напруга затвора прикладена до провідної підкладки, звідки відбувається інжекція носіїв (дірок) у нанотрубку. Такий молекулярний транзистор є аналогом польового транзистора p -типу. Він характеризується значною (кілька порядків) зміною струму витoku залежно від напруги на затворі. Важливою характеристикою цього приладу є те, що він працює за кімнатних температур. Робоча частота приладу визначається опором на контактах метал-молекула та за величини останніх ~ 1 МОм становить 0,1 ТГц. У випадку, якщо вдасться зменшити величину опору контактів до квантової межі ~ 10 Ком, частота приладу може зрости до 10 ТГц.

Наведені приклади незвичних електронних властивостей різних алотропних форм вуглецю дають підстави вважати, що вуглець у майбутньому може використовуватись як базовий матеріал для нової напівпровідникової електроніки.

■ 19.4. Спінтроніка

Як зазначалось на початку цього розділу, одним із можливих напрямків подальшого розвитку електроніки є використання одноелектронних ефектів і перехід до одноелектроніки. При цьому ми в основному розглядали транспорту заряду одного електрона та виникнення додаткових ефектів, цікавих як з погляду фундаментальної фізики, так і практичних застосувань. Але поза увагою залишився ще один аспект електронних властивостей напівпровідника, що викликаний наявністю спіну електрона.

19.4.1. А як же спін?

Ви, шановний читачу, помітили, що розглядаючи властивості напівпровідників, ми в основному використовували такі параметри електронів у напівпровіднику, як заряд, ефективна маса та майже не

говорили про спин. Чи дійсно така важлива характеристика електрона, як спин та ефекти, пов'язані з наявністю спін $1/2$ в електрона не є причиною фізичних явищ у напівпровідниках, цікавих з погляду як фундаментальної науки, так і практичних застосувань? Зрозуміло, що це не так. По-перше, ми деякою мірою брали до уваги наявність спін, коли розглядали спін-залежні явища у процесах рекомбінації електрон-діркових пар у напівпровідниках або обговорювали вплив спин-орбітальної взаємодії на формування зонної структури напівпровідника. По-друге, існує ціла область фізичних ефектів та явищ, що пов'язані з електронним спіном. Розглянемо тепер деякі ефекти у напівпровідниках, що визначаються електронним спіном та обговоримо можливе їх застосування у світі нової галузі сучасної науки, що має назву "спінтроніка".

"Спінтроніка" – це загальна назва напрямку прикладної фізики та електроніки, що вивчає та використовує спін-залежні явища у напівпровідниках, металах і їх структурах. Хоча на час написання цієї книги пристрої спінтроніки, що тією чи іншою мірою використовуються на практиці, побудовані на основі металів, досягнення у дослідженні спін-залежних явищ у напівпровідниках і розвиток відповідних технологій дозволяють з оптимізмом дивитись на майбутнє напівпровідникових структур у спінтроніці. Зокрема, на основі магнітних напівпровідників можна створювати мікросхеми нового типу, що одночасно реалізують функції пам'яті та обробки інформації. Крім того, деякі спін-залежні явища, особливо в органічних напівпровідниках, можна розглядати як одноелектронні. Створенню спінтронних приладів на основі твердотільних напівпровідників для практичних застосувань заважає те, що явище напівпровідникового магнетизму спостерігається за низьких температур. Але виготовлення магнітного напівпровідника із фосфіду галію з домішками марганцю, що зберігає магнітні властивості за кімнатних температур, та інтенсивний розвиток технологій магнітних напівпровідникових структур, особливо органічних, дає надію на створення комерційних пристроїв напівпровідникової спінтроніки. Крім того, в органічних напівпровідниках спінова поляризація електронів зберігається і за кімнатних температур.

19.4.2. Спін-залежні явища у напівпровідниках

Для вивчення ефектів (а тим більше для їх застосування), що пов'язані з наявністю спін електронів у напівпровідниках, необхідно вміти утворювати потоки сильно поляризованих за спіном вільних електронів. Тому потрібно спочатку визначити, яким чином можна створювати електронні струми, що характеризуються певним напря-

мком спіну. Можна визначити кілька методів створення спіно-поляризованих електронних (або діркових) потоків:

- ♦ орієнтація власних магнітних моментів носіїв заряду під дією зовнішнього магнітного поля;
- ♦ оптична орієнтація спінів;
- ♦ інжекція спіно-поляризованих носіїв з магнітного напівпровідника чи ферромагнітного металу у напівпровідник.

Розглянемо детальніше фізику цих методів. Як відомо з курсу квантової механіки, наявність спіна \mathbf{s} обумовлює магнітний момент частинки

$$\mathbf{m}_s = -g\mu_B \mathbf{s}, \quad (19.6)$$

де μ_B – магнетрон Бора, а електронний g -фактор є числом, у більшості випадків близьким до двійки. Енергією взаємодії спіну із магнітним полем є¹

$$W = -\mu \mathbf{H}. \quad (19.7)$$

Тоді стає зрозумілим, що найбільш очевидний спосіб утворення спінової поляризації системи вільних електронів – це *поляризація під дією зовнішнього магнітного поля*. Але цей метод виявляється недостатньо ефективним за кімнатних (або не дуже низьких) температур. Наприклад, за кімнатної температури поле в 3000 Е викликає поляризацію електронів напівпровідника величиною $\sim 0,1\%$.

Іншим, ефективнішим, методом створення системи спіно-поляризованих електронів є *оптична орієнтація спінів*. Цей метод базується на тому, що у результаті міжзонних переходів, викликаних циркулярно-поляризованим світлом, генеруються нерівноважні електрони з ненульовою сумарною спіновою проекцією, знак якої визначається напрямком поляризації (права чи ліва) світла. Експериментально отримати оптично орієнтовані за спіном електрони можна, освітлюючи зразок циркулярно-поляризованим світлом. У зразку генеруються електрони та дірки з переважно орієнтованим спіном за чи проти напрямку розповсюдження світлової хвилі. За деякий час вони рекомбінують із випромінюванням світла. Якщо за час життя носіїв заряду їх спіни не встигнуть повністю релаксувати, то світлове випромінювання (фотолюмінесценція) буде частково поляризоване, що й реєструється на експерименті.

Розглянемо виникнення оптичної поляризації носіїв заряду на прикладі прямозонного напівпровідника – GaAs (арсеніду галію), ширина забороненої зони якого дорівнює 1,43 еВ. Квантові стани електронів

¹ Строго кажучи взаємодія спіну та магнітного поля в гамільтоніані системи описуватиметься доданком $\hat{H} = g\mu_B \hat{s}_j H_j$, і для визначення впливу магнітного поля на спіновий стан електронів системи треба проводити послідовний квантовомеханічний розрахунок.

провідності поблизу дна зони провідності (при $k=0$) описуються хвильовими функціями s -типу $S_j(m_j)$, що відповідають стану електрона зі спіновим квантовим числом $s=1/2$ та орбітальним квантовим числом $l=0$. Стани зони провідності дворазово вироджені за проекцією спіну ($s_z=m_j=\pm 1/2$). Електронні стани валентної зони описуються хвильовими функціями p -типу – $P_j(m_j)$, де орбітальний момент $l=1$. Завдяки спін-орбітальній взаємодії валентна зона розщеплюється на дві підзони $P_{1/2}$ та $P_{3/2}$, які відрізняються величиною повного кутового моменту електрона, $j=1/2$ та $j=3/2$, відповідно. Дірки, що належать до цих підзон, мають різну ефективну масу. Підзона легких дірок $P_{1/2}$ лежить на 0,34 еВ нижче підзони важких дірок $P_{3/2}$. Підзона $P_{1/2}$ дворазово вироджена ($m_j=\pm 1/2$). Кратність виродження підзони $P_{3/2}$ дорівнює чотирьом ($m_j=\pm 3/2, \pm 1/2$) (рис. 19.12). Для світла з правою циркулярною поляризацією проекція спіну фотону на напрямок розповсюдження хвилі дорівнює $+1$. Якщо енергія світлового кванта більша за ширину забороненої зони E_g та менша за $E_g+0,34$ еВ, то, відповідно до правил відбору за квантовими числами $\Delta m_j=1, \Delta l=\pm 1$, відбуваються електронні переходи $P_{3/2}(-3/2) \rightarrow S_{1/2}(-1/2)$ та $P_{3/2}(-1/2) \rightarrow S_{1/2}(1/2)$. Детальний розрахунок матричних елементів оператора взаємодії електромагнітної хвилі з електроном для цих переходів дає

$$\begin{aligned} b_1 &= \langle P_{3/2}, -3/2 | x + iy | S_{1/2}, -1/2 \rangle = -\sqrt{2/3} R_3, \\ b_2 &= \langle P_{1/2}, -1/2 | x + iy | S_{1/2}, 1/2 \rangle = -2,3 R_1, \\ b_3 &= \langle P_{3/2}, -1/2 | x + iy | S_{1/2}, 1/2 \rangle = -\sqrt{2/3} R_3. \end{aligned} \quad (19.8)$$

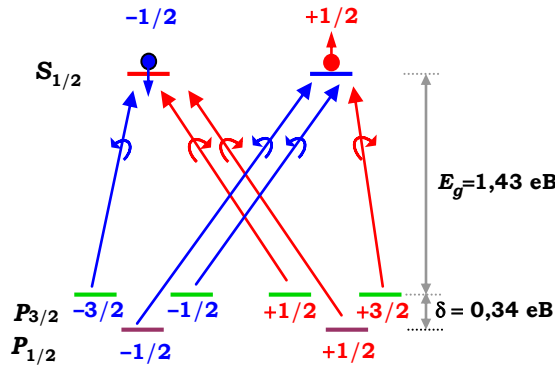


Рис. 19.12. Квантові стани електронів поблизу дна зони провідності та стелі валентної зони GaAs; електронні переходи, спричинені циркулярним правополяризованим (сині стрілки) та лівополяризованим (червоні стрілки) світлом

Тут вважалось, що оператор електродипольної взаємодії циркулярно поляризованого світла з електроном є пропорційним до оператора "правої" циркулярної координати $x + iy$. Отримані вирази свідчать про те, що за генерації нерівноважних носіїв циркулярно поляризованим світлом ймовірність генерації електронів провідності у стані $|S_{1/2}, -1/2\rangle$ у $b_1^2/b_3^2 = 3$ рази вища, ніж у стані $|S_{1/2}, 1/2\rangle$, тобто в зоні провідності електронів із проекцією спіну $-1/2$ більше, ніж електронів із проекцією спіну $1/2$. У першому наближенні концентрація носіїв із різною проекцією спіну пропорційна до відповідної ймовірності генерації, тобто до квадрата матричного елемента. Таким чином, очікувана спінова поляризація електронів у зоні провідності:

$$P = \frac{b_3^2 - b_1^2}{b_3^2 + b_1^2} = -0,5. \quad (19.9)$$

Знак мінус указує, що при збудженні електронів правополяризованим світлом їх спіни орієнтуються у напрямку, протилежному напрямку розповсюдження світла. Таким чином, теоретична оцінка дає максимально можливу спінову поляризацію за оптичної інжекції спін-поляризованих електронів 0,5, що є величиною, набагато більшою, ніж та, яку можна отримати за не дуже низьких температур у магнітних полях $\sim 10^3$ Е. На практиці величина спінової поляризації за оптичної інжекції носіїв не перевищує 0,4 завдяки процесам спінової релаксації. У ненульовому зовнішньому магнітному полі відмінність реального середнього значення спінової поляризації від його максимальної величини зумовлена не лише процесами спінової релаксації, а й прецесією електронних спінів навколо напрямку магнітного поля.

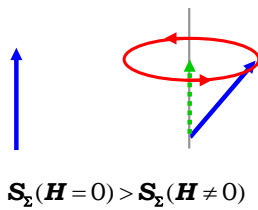


Рис. 19.13. Прецесія ефективного сумарного спіну поляризованих електронів у магнітному полі

Деполаризацію електронів під дією магнітного поля називають ефектом Ханле. Грубо це явище можна уявити як зменшення проекції ефективного сумарного спіну за рахунок його прецесії (рис. 19.13). Для коректнішого опису ефекту Ханле необхідно записати рівняння руху електронних спінів у магнітному полі із врахуванням процесів генерації електронів, їх рекомбінації та релаксації їх спінів:

$$\frac{\partial w_{++}}{\partial t} = \frac{i\gamma H_x}{2} (w_{-+} - w_{+-}) + g_+ - \frac{w_{++}}{\tau} - \frac{w_{++} - w_{--}}{2\tau_s},$$

$$\begin{aligned}\frac{\partial w_{-}}{\partial t} &= \frac{i\gamma H_x}{2}(w_{+} - w_{-}) + g_{-} - \frac{w_{-}}{\tau} - \frac{w_{-} - w_{++}}{2\tau_s}, \\ \frac{\partial w_{+}}{\partial t} &= \frac{i\gamma H_x}{2}(w_{-} - w_{++}) - \frac{1}{T}w_{+}.\end{aligned}\quad (19.10)$$

У цих формулах g_{+} та g_{-} – швидкості генерації електронів зі спіном $+1/2$ та $-1/2$, відповідно, τ та τ_s – час життя електрону та час його спінової релаксації, $1/T = 1/\tau + 1/\tau_s$. Дорівнявши до нуля похідні від елементів матриці густини w_{ij} за часом, отримуємо рівняння для стаціонарних кількостей електронів:

$$\begin{aligned}w_{++} + w_{--} &= (g_{+} + g_{-})\tau, \\ w_{++} - w_{--} &= -\frac{2\text{Im}(w_{+-})}{\gamma H_x T},\end{aligned}\quad (19.11)$$

$$w_{++} - w_{--} = (g_{+} - g_{-})T - 2\gamma H_x T \text{Im}(w_{+-}).$$

Розв'язуючи ці рівняння і враховуючи, що швидкість генерації електронів пропорційна квадрату відповідного матричного елемента, знайдемо величину стаціонарної спінової поляризації електронів

$$P_e = \frac{b_3^2 - b_1^2}{b_3^2 + b_1^2} \cdot \frac{1}{1 + \tau/\tau_s} \cdot \frac{1}{1 + \gamma^2 H_x^2 T^2}.\quad (19.12)$$

Бачимо, що ненульове магнітне поле спричиняє зменшення спінової поляризації в $(1 + \tau/\tau_s)(1 + \gamma^2 H_x^2 T^2)$ разів. За збільшення магнітного поля спінова поляризація зменшується.

І наостанок розглянемо процес утворення спін-поляризованого струму шляхом *інжектування носіїв з переважно визначеним напрямком орієнтації спінів* з магнітовпорядкованого матеріалу через бар'єр у напівпровідник. Оскільки магнітовпорядковані (ферромагнітні) матеріали характеризуються "вбудованим" магнітним полем, то виникає можливість спінової поляризації носіїв заряду в таких матеріалах. Дійсно, наявність магнітного поля в ферромагнетиках спричинить збільшення концентрації електронів зі спіном, який протилежно направлений щодо внутрішнього поля. І тепер необхідно інjektувати ці носії з ферромагнетика у напівпровідник. Таким чином, на перший погляд природним розв'язком проблеми інжекції спін-поляризованих носіїв у напівпровідник здається використання магнітних інжекторних контактів, що формуються з ферромагнітних металів, таких як Fe, Co або Ni. Фізичною основою спінової поляризації електричного струму та ферромагнетика є різниця у щільності станів електронів зі спінами вниз та вгору і, як наслідок, різниця в провідності електронів із різною спіновою поляризацією. Спінова інжекція такого типу успішно реалізується для системи "ферромагнітний метал–надпровідник". Але неодно-

разові спроби реалізувати цю ідею для системи "ферромагнітний метал-напівпровідник" виявилися дивовижно невдалими. Ефективність спінової інжекції у цьому випадку не перевищує одного відсотка, тобто на кожні 101 електронів зі спіном униз припадає 100 електронів зі спіном уверх. Зрозуміло, що така величина спінової поляризації не дає можливість будувати будь-які прилади. Однією із причин цього є велика різниця в провідності напівпровідника та інжектуючого металічного контакту. Альтернативою інжекції з металічних контактів може стати використання як інжектора спінів у немагнітний напівпровідник із магнітного напівпровідника. Зонна структура магнітних напівпровідників відрізняється від двозонної структури звичайних напівпровідників наявністю ще однієї зони, що утворюється d - і f -електронними оболонками перехідних або рідкоземельних елементів. Аналогічно у перехідних ферромагнітних металах, таких як Fe, Ni, або Co

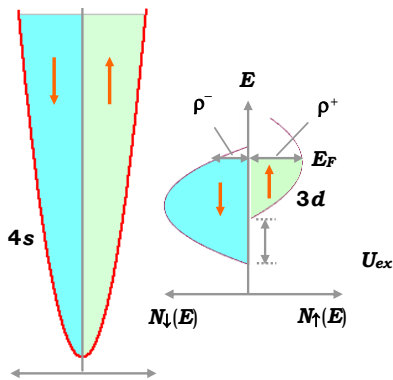


Рис. 19.14. Щільність станів ферромагнітного матеріалу

Завдяки розщепленню $3d$ -зони, кількість електронів зі спіном угору (щільність таких станів ρ^+ подано зеленим кольором) не збігається з кількістю електронів зі спіном униз (щільність електронних станів зі спіном вниз ρ^- подано блакитним кольором). У матеріалі виникає нескомпенсований магнітний момент

рівень Фермі E_F перетинає $4s$ - та $3d$ -зони. Як схематично показано на рис. 19.14, у ферромагнітному металі вузька $4s$ -зона поділяється на дві підзони з однаковою кількістю електронів з електронними станами "спін угору" (*spin-up*) та "спін униз" (*spin-down*). З іншого боку $3d$ -зона розщеплюється та характеризується "спін угору"- та "спін униз"-підзонами, що зсунуті за енергією на величину U_{ex} . Як результат, кількість заповнених станів зі спіном униз і вгору не збігається. Таким чином, спонтанна намагніченість речовини виникає завдяки тому, що заповнених станів зі спіном униз більше, ніж зі спіном угору. У цьому сенсі у магнітних напівпро-

відниках зона з електронними станами зі спіном униз називається зоною основних носіїв (*majority band*), а зона з електронними станами зі спіном угору – зоною неосновних носіїв (*minority band*).

Концепція про те, що в ферромагнітних перехідних металах електричний струм може формуватись спін-поляризованими електронами з'явилась близько 70 років тому. Обчислюючи інтенсивність розсіяння $3d$ -електронів зі спіном вгору та вниз, Мотт з'ясував, що електрони

можуть характеризуватись деяким сумарним спіном. Сумарна поляризація струму в феромагнетиках може бути виражена феноменологічно через спінові провідності в підзонах: σ^+ , що відповідає провідності електронів зі спіном вгору, і σ^- – зі спіном униз

$$P = \frac{(\mathbf{j}^+ - \mathbf{j}^-)}{(\mathbf{j}^+ + \mathbf{j}^-)} = \frac{\sigma^+ - \sigma^-}{\sigma^+ + \sigma^-} \neq 0. \quad (19.13)$$

У цій формулі \mathbf{j}^+ та \mathbf{j}^- – парціальні струми в спінових підзонах. Експериментально підтвердження ідеї Мотта було отримано на планарній тунельній структурі, що являє собою тонку надпровідну (S) алюмінієву

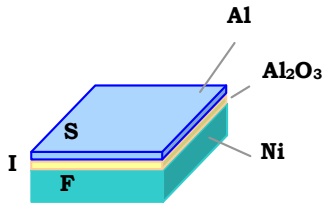


Рис. 19.15. Планарна тунельна структура для створення спін-поляризованого струму

плівку, тунельний бар'єр з оксиду алюмінію (I), і феромагнітний метал (F) як протилежний електрод (рис. 19.15). На такій структурі були проведені низькотемпературні тунельно-спектроскопічні вимірювання. Інтерпретація тунельної провідності дозволила визначити результуючу спінову поляризацію

P тунельного струму через F-I-S-структуру. Для структур, де як F використовувались Ni, Fe або Co, величина спінової поляризації P зі зміною товщини бар'єру та температури варіювалась в інтервалі 10–45 %. У цих експериментах надпровідна алюмінієва плівка використовувалась як детектор поляризованих спінів. Пізніше було з'ясовано, що замість надпровідної плівки можна використовувати іншу феромагнітну плівку, що може виконувати ті самі функції. Зрозуміло, що ідеї Мотта можуть бути застосовані й для магнітних напівпровідників, а величина спінової поляризації в них може також описуватись (19.12).

Найбільш ефективно спінова інжекція з напівпровідника у напівпровідник може реалізуватись, якщо спіновими інжекторами виступають напівмагнітні напівпровідники II–VI груп, наприклад такі матеріали, як CdMnTe, BeMnZnSe, ZnMnSe. Їх провідність порівняна з провідністю немагнітних напівпровідників, до яких відбувається інжекція спін-поляризованих електронів – CdTe або AlGaAs. З іншого боку, спінова поляризація електронів у напівмагнітних напівпровідниках під дією зовнішнього магнітного поля може дорівнювати ста відсотків. Таке велике значення спінової поляризації досягається в результаті $sp-d$ -обмінної взаємодії вільних електронів з $S = 5/2$ спінами локалізованих $3d^5$ -електронів іонів марганцю. Відносно невеликі зовнішні поля (менші за 0,5 Тесла), що можуть бути вироблені сталими мікромагнітами, спричиняють у твердих розчинах із невеликим вміс-

том Mn (2–10 %) гігантське ефективне зєсманівське розщеплення Δ -станів зони провідності (величина цього розщеплення досягає кількох міліелектронвольт). Як результат, спіни електронів будуть направлені вздовж напрямку зовнішнього магнітного поля. Експериментально таку інжекцію можна спостерігати в діодних структурах, де інжекція відбувається з напівмагнітного напівпровідника групи II–VI у світловипромінюючий діод з активною областю на основі AlGaAs/GaAs-квантової ями. Рекombінація спін-поляризованих носіїв заряду спричиняє випромінювання світла з циркулярною поляризацією, згідно із правилами відбору. Ступінь циркулярної поляризації дозволяє оцінити ефективність спінової інжекції. Отримані в таких експериментах значення ступеня поляризації становлять 43 % за температури 2,5 K та 50 % – за температури 4,2 K. Це дало змогу оцінити ефективність спінової інжекції у таких структурах, як 85 % та 50 %, відповідно. Варто зауважити, що серйозною проблемою використання структур із напівмагнітними твердими розчинами напівпровідників групи II–VI як інжекторів спін-поляризованих електронів є необхідність прикладення зовнішніх магнітних полів за низьких температур. Справа в тому, що ці матеріали є парамагнетиками, де підвищення температури приводить до швидкого зменшення їхньої намагніченості навіть під дією фіксованого магнітного поля. Для подолання цієї проблеми найбільш прийнятним є використання спінових інжекторів, що являють собою феромагнітні напівпровідники з високою температурою Кюрі (T_c – температура, за якої феромагнетик утрачає спонтанну намагніченість). До таких напівпровідників належать магнітні напівпровідники групи III–V, в яких T_c залежить від концентрації вільних носіїв заряду. Найбільша температура Кюрі (110 K) спостерігалась у твердому розчині GaMnAs з p -типом провідності. Цей матеріал було використано як спіновий інжектор у нульовому магнітному полі в електролюмінесцентному діоді з немагнітною квантовою ямою InGaAs/GaAs. Ефективність спінової інжекції такої структури оцінювалась у кілька відсотків за $T = 6$ K. Зрозуміло, що до впровадження таких інжекторів у практику ще далеко. Але інтенсивний пошук нових матеріалів та отримані результати дозволяють зі стриманим оптимізмом сподіватись на розв'язання проблеми спінових інжекторів, що працювали б за кімнатних температур і нульових магнітних полів. Наприклад, у магнітному напівпровіднику зі структурою халькопїриту $\text{Cd}_{1-x}\text{Mn}_x\text{GeP}_2$ феромагнітний стан спостерігався аж до кімнатних температур ($T_c = 320$ K). Нещодавно був синтезований новий магнітний напівпровідник, що являє собою оксид індію, легований хромом. Цей матеріал є спіновим інжектором для кремнію, причому спінова інжекція відбувається за кімнатних температур.

19.4.3. Гігантський магніторезистивний ефект. Спіновий клапан

Можливість практичного застосування спітроніки стала очевидною після відкриття у 1988 році гігантського магніторезистивного ефекту (ГМР)². У штучних матеріалах, що являють собою тонкі плівки магнітного та немагнітного матеріалу, які нашаровуються по чергово один на одного, спостерігається мінімальний опір, якщо поля у феромагнітних шарах направлені паралельно. Якщо магнітні моменти у феромагнітних шарах антипаралельні, опір системи зростає. Основним пристроєм, що використовує ефект ГМР, є так званий спіновий клапан. Він містить два шари феромагнітного матеріалу (зазвичай це залізо або кобальт), розділених тонким шаром немагнітного металу (мідь для Co та хром для Fe – вибір немагнітного матеріалу визначається необхідністю формування необхідного потенціального бар'єру на межі розподілу феромагнетик–неферомагнітний метал). Наприклад, у структурі, що складається з тонких плівок кобальту, розділених міддю, опір спінового клапана за антипаралельних магнітних моментів у феромагнітних шарах (Co) на 5–10 % перевищує опір системи у випадку, якщо напрямки магнітних моментів в обох феромагнітних шарах збігається.

Використання магнітних моментів ядер дає додаткові можливості для побудови систем зберігання та обробки інформації. Але завдяки дуже малому значенню ядерних магнітних моментів використання цієї ідеї є проблематичним. Цю проблему можна розв'язати, якщо керувати станами ядерних магнітних моментів за допомогою електронних спінів, використовуючи при цьому електрони, локалізовані в малих областях квантових ям усередині магнітних напівпровідників. Ці ями утворюються всередині магнітного напівпровідника за допомогою літографічних методів. У кожній такій вузькій квантовій ямі рух електронів практично двовимірний. Попадаючи в таку яму, електрони з визначеним напрямком спіну (поляризовані електрони) утворюють тонкий диск радіусом у ~ 20 мкм і товщиною 20 нм. За допомогою електричного поля цей диск можна рухати в будь-якому напрямку вздовж напівпровідника, подібно до того, як магнітна головка рухається вздовж поверхні магнітного накопичувача. Відпрацьовано методику позиціонування такої електронної хмари з нанометровою точністю. У результаті взаємодії магнітних моментів електронів з магні-

² У 2007 році П.Грюнбергу та А.Феру присуджено Нобелівську премію за відкриття гігантського магніторезистивного ефекту.

тними моментами ядер останні також створюють упорядкований стан. Усі ці процеси відбуваються в дуже тонкому (~ 10 нм) шарі магнітного напівпровідника. У результаті створюється тонкий, майже двовимірний, кластер поляризованих ядер, який і може бути базовим елементом пристрою зберігання інформації. При цьому щільність запису інформації може бути більшою, ніж при використанні спінів електронів. Оскільки ядерні спіни слабо взаємодіють з оточуючим середовищем, вони добре зберігають свій стан. Електронний кластер використовується в описаній системі не для зберігання інформації, а лише як магніт, що керує магнітним станом ядер, і до того ж, не потребує додаткової енергії для підтримки створеного ним магнітного поля. А завдяки наявності в електронів електричного заряду переміщення кластеру поляризованих електронів вздовж напівпровідникової плівки не є проблемою. Проблемою є створення в напівпровіднику магнітнополяризованого електронного кластера. Для цього, зокрема, використовуються складні методи інжекції спін-поляризованих електронів у двовимірну квантову яму, що обговорювались раніше. Наприклад, можна інjektувати спін-поляризовані носії за допомогою лазерного імпульсу. Інший метод полягає в інjektуванні в напівпровідник поляризованих електронів за допомогою квантових магнітних тунельних переходів. У таких переходах із визначеним напрямком спіну електрони тунелюють між двома феромагнітними шарами.

Розглянемо детальніше фізику гігантського марніторезистивного ефекту або гігантського магнітоопору (ГМО). Нагадаємо, у структурах феромагнетик–немагнітний метал–феромагнетик (F_1 –N– F_2) провідність залежить від відносної намагніченості в феромагнітних шарах. Орієнтація намагніченості в шарах F_1 та F_2 контролюється зовнішнім магнітним полем, тобто провідність структури є функцією зовнішнього магнітного поля. А тришарова структура називається спіновим клапаном (рис. 19.14). Характеристикою ГМО є різниця опору структури в двох конфігураціях, що характеризуються паралельною та антипаралельною орієнтацією намагніченості, нормованою на опір системи з антипаралельною орієнтацією. При цьому історично так склалося, що для характеристики величини ГМО використовують дві величини:

"песимістичне" відношення

$$(\Delta R / R)_p = \frac{R_{\max} - R_{\min}}{R_{\max}} \quad (19.14)$$

та "оптимістичне" відношення

$$(\Delta R / R)_o = \frac{R_{\min} - R_{\max}}{R_{\min}}. \quad (19.15)$$

У цих виразах через R_{\max} і R_{\min} позначено максимальний і мінімальний опори через структуру, відповідно. Оскільки опори пов'язані зі

спіновим станом феромагнітних шарів, то часто (залежно від зонної структури феромагнетика) вони позначаються як $R_{\uparrow\uparrow}$ або $R_{\uparrow\downarrow}$, причому, більшому (меншому) опору може відповідати як паралельний, так й антипаралельний напрямки намагніченості в суміжних магнітних шарах. Слід підкреслити, що ефект гігантського магнітоопору є ефектом не матеріалу, а структури, що складається із кількох шарів фе-

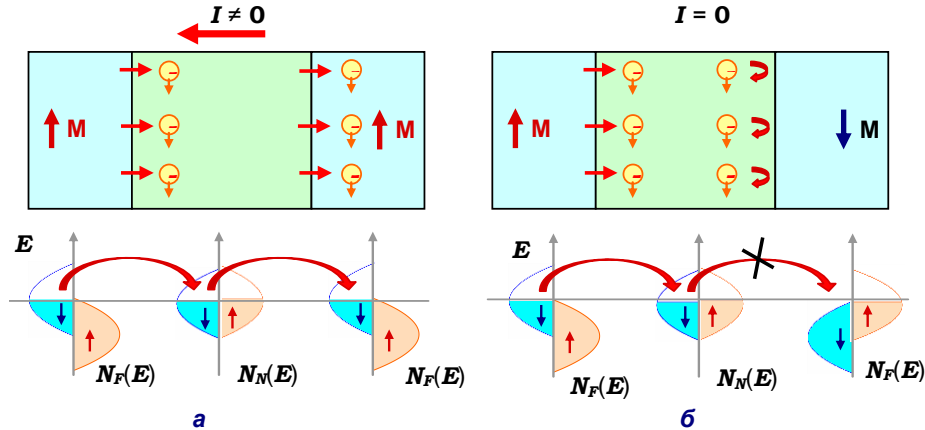


Рис. 19.16. Ненульова (а) та нульова (б) провідність спин-поляризованих електронів у структурі "феромагнетик–неферомагнітний матеріал–феромагнетик" (ідеальний спіновий клапан) залежно від стану намагніченості – випадок ненульової провідності для антипаралельних намагніченостей феромагнітних плівок.

На нижніх рисунках схематично подано хід щільності станів у різних частинах структури що забезпечує потрібні транспортні властивості

ромагнітних та неферомагнітних речовин.

Розглянемо на прикладі двох типів зонної структури феромагнітних металів (і напівпровідників), як реалізуються різні значення опору для спин-поляризованих електронів у спіновому клапані. Нехай спіновий клапан являє собою два феромагнітних шари, виготовлені з однакового матеріалу та розділені шаром немагнітного металу. Припустимо, що в структурі може реалізуватись два типи взаємної орієнтації намагніченості – феромагнітний (рис. 19.16 а) та антиферомагнітний (рис. 19.16 б). Тоді, як видно із зонної діаграми на рис. 19.16 а, у випадку паралельної орієнтації намагніченості феромагнітних шарів електрони із поляризацією спіну вниз можуть переходити з першого (ліворуч) до другого (праворуч) феромагнітного шару, оскільки існують незаповнені стани, що відповідають спіновій поляризації електронів униз як у немагнітному металі, так й у другому феромагнітному шарі. У випадку, коли намагніченість другого феромагнітного шару має протилежну до намагніченості першого шару орієнтацію, електронам з орієнтацією спінів униз немає змоги переходити до другого феромагнітного шару без перевероту спіну (на що треба витратити додаткову енергію). Таким чином електричний опір системи в антиферомагнітному стані

буде більшим, ніж у стані, коли намагніченість обох феромагнітних шарів є паралельною.

Може існувати інша ситуація, зонну діаграму для якої гіпотетично подана на ідеалізованій діаграмі рис. 19.17. У деяких магнітних матеріалах розщеплення $3d$ -зони відбувається асиметрично, так, що велика частина станів основних (за спіном) носіїв незаповнена. Тоді спіновий клапан, що складається із шарів різних феромагнітних матеріалів, розділених немагнітним металом, може мати менший опір, коли намагніченості феромагнітних шарів антипаралельні.

Магніторезистивний ефект у спіновому клапані досягає величини

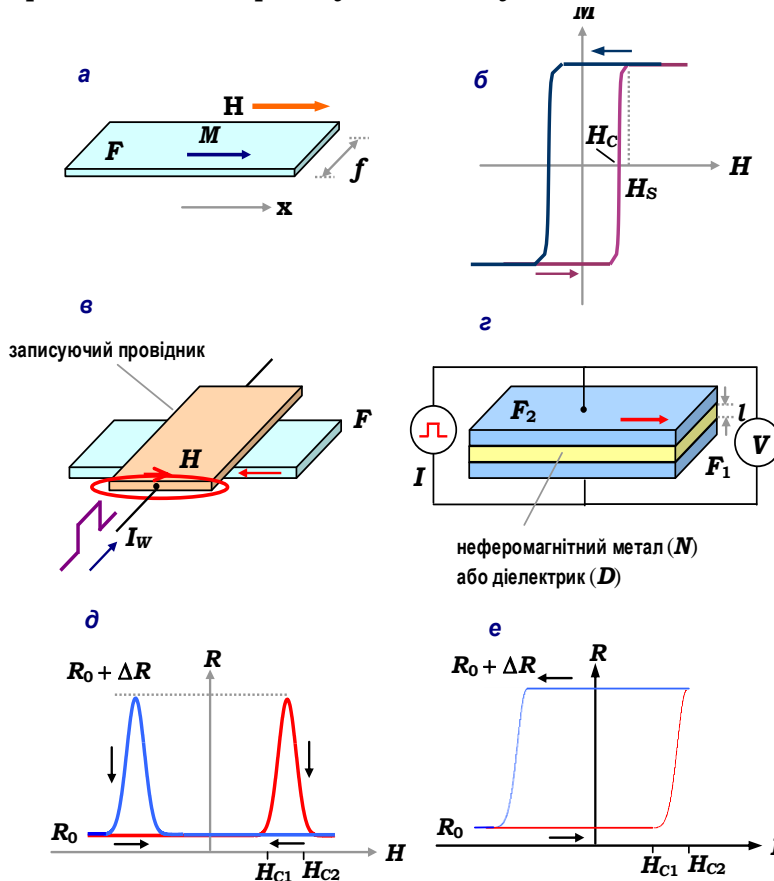


Рис. 19.18. До розуміння роботи елементарної комірки магнеторезистивного пристрою пам'яті

$\sim 10\%$, що є набагато більшим за величину анізотропного магнеторезистивного ефекту, що для феромагнітних перехідних металів становить $2\text{--}3\%$.

Покажемо, яким чином, складаючи систему спінових клапанів, можна побудувати прилад для запису інформації. Звернімося до рис.18.

Тонка феромагнітна плівка F має одновісну анізотропію намагніченості з віссю x і ширину $f(a)$. Петля гістерезиса описує намагніченість M плівки F як функцію магнітного поля H , прикладеного до зразка і направленого вздовж вісі x (б). Інтегровані магнетоелектронні прилади можуть бути виконані методом нашарування на плівку F провідників, якими протікає записуючий струм. Ці провідники індуктивно зв'язані з F і можуть перемикає стан приладу, змінюючи напрямок намагніченості (в). У скетчі (г) показано магнеторезистивний прилад, що складається із двох шарів (верхнього та нижнього) феромагнітного матеріалу, розділених тонким шаром немагнітного металу (N) у випадку спінового клапану або діелектрика (D) – у випадку магнітного тунельного переходу. Магнеторезистивний відгук симетричного тунельного переходу подано на панелі (д). Ефект пам'яті симетричного магнітного контакту (спінового клапану) подано на панелі (е). Умова $H \ll 0$ відповідає бістабільному стану із двома можливими значеннями електричного опору $R \approx R_0$ і $R \approx R_0 \pm \Delta R$. Намагніченість елемента (рис. 19.18 а) є локальною властивістю матеріалу. Для побудови приладу кожне з бінарних значень 0 або 1 необхідно пов'язувати зі стабільним станом намагніченості “ліворуч” або “праворуч” (негативний або позитивний напрямок відносно напрямку \mathbf{x}). Крім цього, узагалі кажучи, при створенні приладу пам'яті необхідна наявність систем введення–виведення. На рис. 19.18 в подано схему перспективного феромагнітного елемента, що складається із тонкої феромагнітної плівки та вбудованого записуючого дроту. У реальних пристроях комірки вбудовуються в двовимірний масив рядків і стовпчиків, що складаються із наборів тонких взаємно перехрещуваних дротів (рис. 19.19).

Подача сигналу на потрібну пару дротів із певних рядка і стовпчика задає адресу комірки, до якої записується біт інформації. При цьому записуючий дріт відокремлюється від феромагнітної плівки шаром немагнітного металу – міді або алюмінію. Коли на дріт подається записуючий сигнал (струм I_w), магнітне поле, **“навіть”** на дріт, викликає магнітний момент у феромагнітній плівці, напрямок якого визначається напрямком струму. При цьому амплітуда магнітного поля є приблизно сталою поблизу поверхні плівки, що покрита записуючим дротом. Таким чином, струм I_w забезпечує локальне магнітне поле, що можна використовувати для контролю стану намагніченості плівки F . Оскільки перемагнічування феромагнітної плівки відбувається з великою швидкістю, швидкість перемикання приладів спінтроники, що керуються короткими імпульсами струму, може бути досить великою та досягати значень ~ 1 нс. Біполярний пристрій перемикання, що використовує для керування струм, поданий на рис. 19.18 в. Імпульс

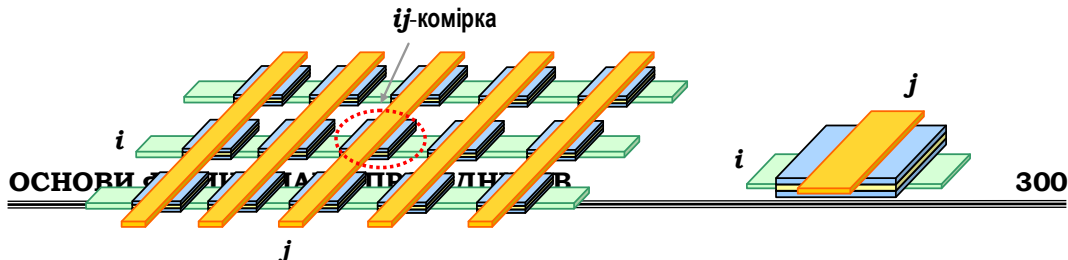


Рис. 19.19. Схема пристрою пам'яті (ліворуч) та вигляд окремої комірки (праворуч), що базується на використанні ефекту гігантського магнітоопору

струму позитивного напрямку $+2I_w$ відповідає стану 0, а імпульс струму негативного напрямку $-2I_w$ – стану 1. Невеликою, але дуже важливою рисою приладу є те, що амплітуда струму записування становить величину $2I_w$. Це дозволяє застосовувати ефективну систему адресації індивідуальної комірки в двовимірному масиві. Амплітуда записуючого імпульсу дозволяє в деякому рядку (i -му) або стовпчику (j -му) забезпечувати величину поля перемикавання $\alpha I_w = H_s/2$ (α – константа індуктивного зв'язку). Поле, що забезпечує ефект перемикавання в комірці (ij), очевидно

$$H_s = 2\alpha I_w. \quad (19.16)$$

Воно є суттєво більшим за поле, що забезпечує орієнтацію намагніченості. Між тим, поле в будь-якій іншій комірці (mn) не достатньо сильне для зміни існуючого стану намагніченості в цій комірці, тобто процес переорієнтації спінів у комірці є напіввизначеним. Для визначення процесу переорієнтації спінів в комірці необхідно, щоб стан намагніченості елементу F міг бути керованим поєднанням із зовнішньою системою керування (бажано неруйнівним чином). Біт інформації має бути „відчутий” у процесі зчитування. Для цього використовують шарувату систему, де складовою є ферромагнітна плівка F (рис. 19.18 а). Кожен із ферромагнітних шарів має товщину у кілька нанометрів. Для спінового клапану шар, що розділяє ферромагнітні плівки, є діелектричним тунельним бар'єром (D) або немагнітним металом (M). На рис. 19.18 б подано загальний магніторезистивний ефект у спіновому клапані. Якщо намагніченості плівок $F1$ і $F2$ ($M1$ і $M2$) є паралельними, ця область має відносно високу провідність, що пов'язана з транспортом основним носіїв зі спінової підзони обох ферромагнітних плівок. Якщо намагніченості відповідних областей плівок $F1$ і $F2$ ($M1$ і $M2$) є антипаралельними, провідність відповідної області приладу, що визначається транспортом неосновних носіїв однієї плівки та основної (основних) спінової підзони іншої ферромагнітної плівки є відносно невеликою.

Хід намагніченості (рис. 19.18 в) стає зрозумілий, якщо розглянути два ферромагнітні шари, що характеризуються різними значеннями коерцитивного поля H_{c1} та H_{c2} ($H_{c2} > H_{c1}$) на петлі гістерезису (рис. 19.18 б). За великих негативних значень магнітного поля \mathbf{H} намагніченості плівок \mathbf{M}_1 та \mathbf{M}_2 паралельні та направлені вздовж вісі \mathbf{x} . Ділянка приладу, що розглядається, характеризується невеликим значенням опору. Коли поле H зростає до нуля (червона крива на рис. 19.18 в), орієнтація намагніченості плівок не змінюється. За подальшого зростання поля та

досягнення ним величини H_{C1} орієнтація намагніченості в першій півці змінюється. Тепер орієнтація намагніченості півок F_1 та F_2 стає антипаралельною, і провідність дільниці приладу зменшується – бачимо значне зростання опору. За подальшого збільшення поля до величини H_{C2} орієнтація намагніченостей півок F_1 і F_2 знов стає паралельною, і провідність дільниці стає великою – опір ділянки зменшується. Аналогічний ефект спостерігається за зменшення поля до великих (за модулем) негативних значень магнітного поля – синя крива.

Розгортка поля може бути зупинена за значень поля з інтервалу $H_{C1} < H < H_{C2}$, що відповідає полю, яке встановлює антипаралельну орієнтацію намагніченостей у півках F_1 і F_2 . Далі зовнішнє поле може бути занулене. Антипаралельний стан зберігатиметься, доки \mathbf{H} не зменшиться до негативних значень $H < -|H_{C1}|$, за яких \mathbf{M}_1 змінить свій напрямок і стане паралельною до намагніченості \mathbf{M}_2 . Таке явище називають ефектом пам'яті. Цей ефект лежить в основі явища бістабільності, що використовується в магнетоелектронних пристроях.

Для керування роботою спінового клапана створюються умови, за яких намагніченість однієї із феромагнітних півок важко змінити зовнішнім полем, а другої – легко, тобто коерцитивність першого шару є високою, а інший феромагнітний шар характеризується невеликою коерцитивністю та може змінювати свою намагніченість за невеликих значень поля - він називається рухомим шаром. При цьому часто петля гістерезису шару з високою коерцитивністю зсунута відносно петлі гістерезису вільного шару несиметрично відносно $H = 0$. У такому випадку кажуть, що використовується схема з підмагнічуванням.

Спінові клапани можуть бути робочими елементами цілої низки приладів спінової електроніки. З огляду на те, що намагніченість рухомого феромагнітного шару відчутно змінюється під впливом слабких магнітних полів, їх використовують для вимірювань величини магнітного поля. При цьому зміна електроопору реальних приладів сягає 1 % на 1 Ерстед. Як реальна перспектива розглядається створення на основі спінових клапанів енергонезалежних елементів пам'яті, оскільки їх магнітний стан зберігається без додаткових витрат енергії.

Що стосується систем зі спін-поляризованим транспортом, то зрозуміло, що визначною вимогою для таких систем є достатньо великий час релаксації спінової когерентності або великі довжини, на яких утрачається спінова поляризація. Іншою вимогою ефективного спінового транспорту є здійснення інжекції поляризованих за спіном електронів. Обидві вимоги неможливо задовольнити окремо, оскільки процес інжекції має бути таким, щоб в ньому зберігався високий ступінь спінової поляризації носіїв заряду, отже треба запобігти спіновій

релаксації при проходженні носіями границі. Зазвичай, особливо для напівпровідникових структур, ці вимоги виконуються за низьких температур. Крім того, значні труднощі викликає необхідність створення досконалих меж розподілу між магнітними та немагнітними матеріалами. Одним із можливих розв'язків цієї проблеми може бути використання молекулярних систем спінової електроніки. У цьому сенсі використання молекул має дві переваги. Першою можна вважати слабку спін-орбітальну взаємодію в молекулах. Іншою перевагою є формування інтерфейсу між магнітним металом і молекулами. Дійсно, спін-орбітальна взаємодія є відповідальною за прецесію спіну, завдяки чому втрачається спінова когерентність. З іншого боку, величина спін-орбітальної взаємодії зростає зі зростанням атомного номеру Z ($\sim Z^4$). В органічних матеріалах завдяки невеликому значенню Z для вуглецю спін-орбітальна взаємодія невелика, на відміну від неорганічних матеріалів. Зокрема, величину спін-орбітальної взаємодії однозначно характеризує параметр спін-орбітального розщеплення Δ_{SO} . Параметр Δ_{SO} для GaAs становить 340 еВ, для Ge – 290 еВ. У кремнії цей параметр досить невеликий – 44 еВ, і все ж таки вуглець має менше спін-орбітальне розщеплення – всього 13 еВ. Ще одна взаємодія, що приводить до втрати спінової когерентності, – надтонка взаємодія, пропорційна скалярному добутку електронного \mathbf{s} та ядерного \mathbf{S}_N спінів $H_{HYF} \sim \mathbf{s} \cdot \mathbf{S}_N$. В органічних матеріалах ця взаємодія є слабкою, що пояснюється в основному тим, що спіновий транспорт в органічних матеріалах реалізується за рахунок π -спряжених зв'язків, які локалізовані на атомах вуглецю. Атом вуглецю ^{12}C характеризується нульовим ядерним спіном і не є активним у надтонкій взаємодії. Невелика частка ізотопу ^{13}C , що характеризується ядерним спіном $\frac{1}{2}$ в органічних речовинах, дає невеликий внесок до надтонкої взаємодії. На рис. 19.20 подано спіновий клапан, виготовлений із застосуванням органічних речовин. Як нижній феромагнітний шар FM_1 у ньому використовується сполука $\text{La}_{0.67}\text{Sr}_{0.33}\text{MnO}_3$ (у літературі такий матеріал часто позначають як LSMO). Верхній феромагнітний контакт FM_2 виготовлено з металу, що характеризується високою коерцитивністю – Co. Для поліпшення омичного контакту верхній металічний шар виконано з попереднім нанесенням алюмінію. π -спряжений органічний напівпровідник (OSE) 8-гідрокси-квінолін-алюміній (Alq_3), що широко використовується у світловипромінюючих діодах, використано як спейсер, що добре зберігає спінову поляризацію. За невеликих зміщень V дірки інжектуються з аноду на НОМО-рівень молекулярного напівпровідника Alq_3 , головним чином, шляхом тунелювання через бар'єр. Наяв-

ність майже однакових значень роботи виходу ϕ в обох ферромагнітних шарах приводить до симетричної вольт-амперної характеристики структури (рис. 19.20 а).

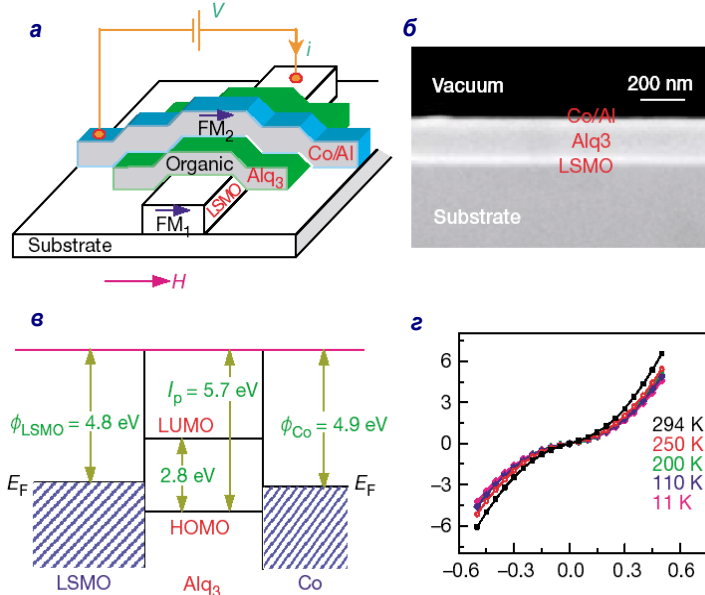


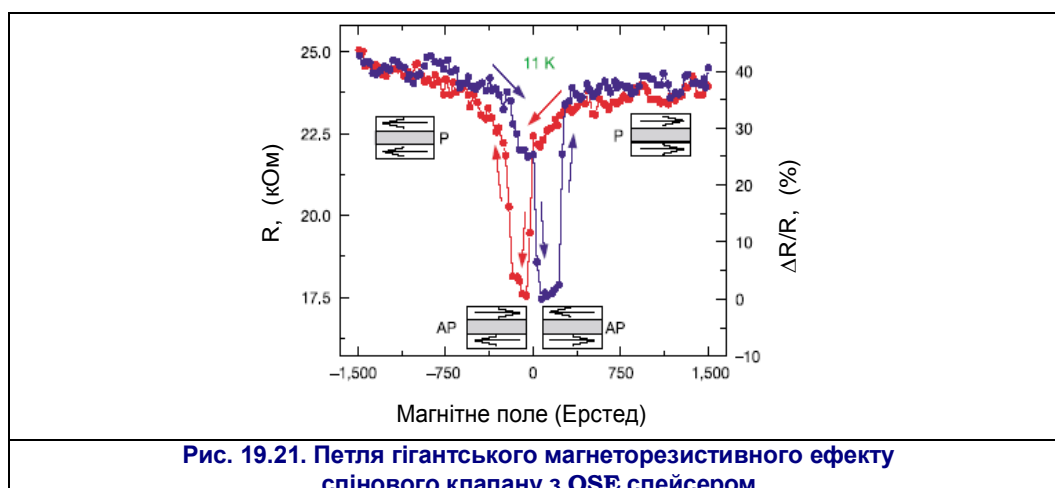
Рис. 19.20 Структура та транспортні властивості органічного спінового клапану:
а – схема приладу;

б – макрографія перерізу приладу із 60 моношарами LSMO плівки та 160 нм Alq₃ спейсера;

в – енергетична діаграма;

г – вольт амперні характеристики за різних температур

Для приладів із товщиною органічного шару, більшою за 100 нм, вольт-амперна характеристика є нелінійною, що слабо залежить від температури (рис. 19.20 г). Це свідчить про тунельний характер інжекції носіїв. На відміну від багатьох спінових клапанів, що побудовані без застосування органічних матеріалів, цей пристрій показує дуже добрі характеристики ГМР-ефекту за низьких температур. Зокрема, як видно з рис. 19.21, гігантський магніторезистивний ефект $\Delta R/R$ тут може досягати 40 %.



Надзвичайно придатними для задач спінтроники є провідні полімери. Вони цікаві, у першу чергу, через їхню високу провідність, можливість створювати парамагнітні центри в полімерах шляхом приєднання іон-радикалів, молекул у триплетному стані та комплексів перехідних металів. Крім того, елементарними збудженнями в полімерах є не зонні електрони, а електронно-коливальні збудження (солітони, полярони, біполярони), що дозволяє зберігати спінову когерентність. Наприклад, транс-ізомер поліацетилену, як очікується, характеризується солітонподібним механізмом транспорту. Електронні збудження тут мають специфічну спін-зарядову взаємодію, оскільки нейтральний солітон відповідає радикалу зі спіном $\frac{1}{2}$, у той час, коли негативно та позитивно заряджені солітони є безспіновими та діамагнітними. Вивчення спінового транспорту в системах, де використовуються контакти між перехідними металами й такими полімерами, здається надзвичайно цікавим. Іншим прикладом цікавої структури молекулярної спінтроники є молекулярний спіновий клапан.

Оскільки дотепер немає єдиного погляду на формування спінового транспорту в молекулярних структурах, роботи, де вивчаються такі ефекти, в основному є теоретичними. Надійних експериментальних даних, що могли б однозначно підтверджувати чи спростовувати наявність тих чи інших механізмів спінових явищ, немає. Відсилаємо бажаючих детальніше ознайомитись зі станом проблеми, до літератури, що наведена в кінці розділу.

19.4.4. Спіновий транзистор

Подальший розвиток спінової електроніки сприяв швидкому переходу від двоелектродних до триелектродних пристроїв спінової електроніки. Джонсон одним з перших запропонував спіновий транзистор, що складається з двох феромагнітних шарів, розділених тонким парамагнітним прошарком до яких під'єднуються контакти. По аналогії з біполярним транзистором транзистор Джонсона складається з бази, роль якої відіграє парамагнітний шар, емітера та колектора, що являють собою феромагнітні обкладинки. Якщо на колектор подається потенціал, у ланцюжку емітер-база відбувається накопичення електронів, спіни яких орієнтовані вгору (вниз). Струм колектора тепер буде залежати від того, паралельний чи антипаралельний його магнітний момент відносно намагніченості емітера. Феромагнітний емітер відіграє роль поляризатора для спінів, що накопичуються. Для зміни потенціалу в ланцюжку емітер-база необхідно прикласти магнітне поле, яке «перемкне» вектор намагніченості емітера або колектора на протилежне за напрямком. Такий транзистор є дієздатним приладом, але він має суттєвий недолік – всі напруги, що вимірюються, дуже малі і це

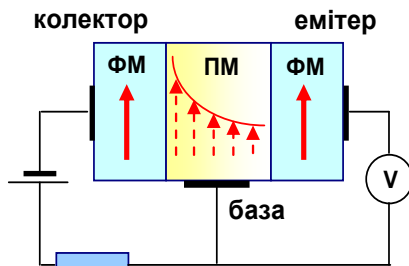


Рис.19.22. Спіновий транзистор Джонсона

збільшити їх без додаткового підсилення (тобто без застосування додаткових пристроїв) не можливо. Оскільки всі компоненти структури – металічні, то всі контакти між шарами є омичними – є суттєвий недолік, який можна подолати шляхом створення структур іншого типу – гібридних спін-електронних пристроїв. В таких пристроях реалізується інтеграція магнітних матеріалів з напівпровідниками. Феромагнетики поляризують спіни, а напівпровідники дозволяють використовувати ефект блокади напруги, струмів дифузії та тунелювання. До таких пристроїв в першу чергу можна віднести транзистор Монсма. Такий транзистор являє собою спіновий вентиль, що розташовується між двома шарами кремнію. Кремнієві шари відіграють роль емітера та колектора, а спіновий вентиль, що складається з немагнітних та магнітних шарів, відіграє роль бази (Рис.19.23). В інтерфейсах між кремнієм та металом формуються бар'єри Шотткі на яких падає напруга зміщення, що прикладаються між парами контактів. При цьому бар'єр Шотткі колектора є обернено зміщеним, а емітера – характеризується випере-

джаючим зміщенням. Такий розподіл напруг у приладі дозволяє інжектувати неполяризовані "гарячі" електрони з напівпровідникового емітера у металічну базу з енергією вищою за енергію Фермі електронів у металі. Тепер електрони можуть пройти скрізь спін-вентиль і зберегти достатньо енергії для подолання бар'єру Шотткі колектора. Якщо цього не відбудеться вони залишаються у базі і переходять у зовнішнє електричне коло. Якщо намагніченість суміжних шарів спін-вентилля направлені антипаралельно, то електрони з обидвама напрямками проекції спіну однаково розсіюються на магнітних шарах. Якщо подіяти на спін-вентиль зовнішнім магнітним полем так, що намагніченості цих шарів будуть направлені паралельно, то електрони

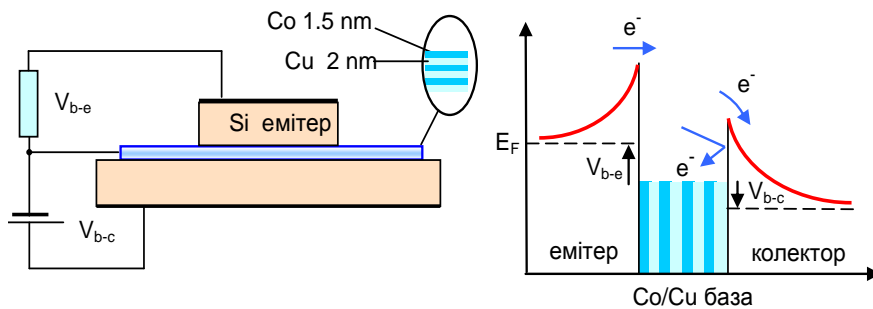


Рис. 19.23. Схема спінного транзистора Монсма (ліворуч). Енергетична схема транзистора (праворуч)

з однією проекцією спіну (наприклад, спін-вниз) будуть розсіюватись сильно. В цей же час електрони, проекція спіну яких направлена вгору будуть без розсіювання проходити всю магнітну структуру. Таким чином, за феромагнітного орієнтування магнітних моментів більша кількість електронів зі спіном орієнтованим вгору будуть з енергією, вищою за енергію бар'єра колектора і будуть проходити через базу. Тобто, як і у випадку транзистора Джонсона, ми маємо справу з транзистором, електричні властивості якого можна контролювати змінюючи магнітне поле. Але в транзисторі Монсма робоча напруга і чутливість до маг-

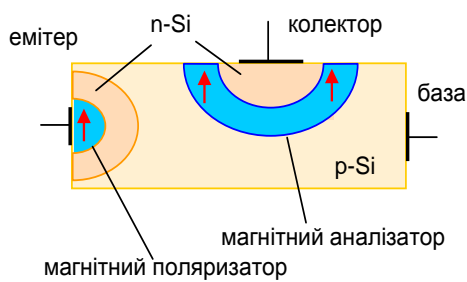


Рис. 19.24. Схема SPICE-транзистора

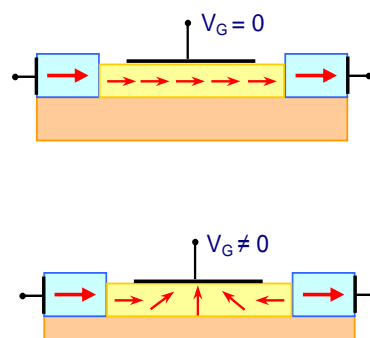


Рис.19.25. Схема роботи польового спінного транзистора

нітного поля є вищими, ніж у транзисторі Джонсона. Цей факт дозволяє сподіватись на його практичне застосування. Замітимо, що напівпровідники у розглянутому приладі використовуються лише для створення енергетичних бар'єрів на контактах метал-напівпровідник та екранування спін-залежної частини пристрою від електричних полів. Для використання всього потенціалу властивостей напівпровідників у спінових транзисторах треба задіяти і напівпровідникові шари для спін-залежного транспорту. Цю задачу вирішує створення транзистора з інжектуванням спін-поляризованого струму емітера в екрановану область пристрою, так званого SPICE-транзистора (*spin-polarized injection current emitter transistor*). Структура такого транзистора показана на рис. 19.24. Цей пристрій характеризується можливістю попереднього підсилення потужності, а його електричні характеристики можуть бути керованими зовнішнім магнітним полем. Міжфазні межі розподілу емітера та колектора в такому транзисторі можуть бути реалізованими *p-n*-переходами, бар'єрами Шотткі, або спін-тунельними переходами.

Інший тип пристрою, що використовує спін-залежні явища запропонували Датта та Дас. Цей транзистор, що базується на релятивістському ефекті спінової поляризації і є аналогом польового транзистора класичної електроніки. Такий пристрій формується з ферромагнітних витоків та стоку. Як затвор використовується напівпровідник (InAlAs плівка на InGaAs підкладинці - Рис.19.25). Коротко роботу транзистора можна описати так: спін-поляризовані електрони покидають виток маючи спіни, орієнтовані паралельно до напрямку намагніченості ферромагнетика. При цьому, завдяки ефекту Рашби під час руху електронів з високою швидкістю їх спіни прецесують. За достатньо великої напруженості магнітного поля і за високої дрейфової швидкості електронів ($\sim 0.01c$, c – швидкість світла) спіни електронів можуть змінювати орієнтацію на протилежну. Цей ефект можна просто зрозуміти, якщо перейти в рухому систему координат, що пов'язана з електроном. В такій системі координат за релятивістського руху електрона з'явиться магнітне поле

$$\mathbf{H} = (1/c) \mathbf{v} \times \mathbf{E}, \quad (19.17)$$

де \mathbf{E} – напруженість електричного поля, що створено прикладеним до затвору потенціалом. За достатньо великих значень магнітного поля спіни електронів можуть змінювати орієнтацію на протилежну. Таким чином, величина спін-поляризованого струму в приладі буде визначатись прикладеним до затвору потенціалом. Як результат електричний опір каналу зростає і струм зменшиться. Змінюючи потенціал на затворі можна змінювати провідність пристрою. Таким чином, спі-

новий польовий транзистор поводить себе як польовий транзистор з тією різницею, що диференціальна намагніченість його контактів (що означає, і його електричні характеристики) є чутливою до зовнішнього магнітного поля

Список літератури

1. Likharev K., *Single Electron Devices and Their Applications* // Proc.IEEE.-1999.- vol.87.- p.606-632
2. MacDiarmid Alan G., *Synthetic Metals: a Novel Role for Organic Polymers* // Synthetic Met.- 2002.- vol.125.-p.11-22
3. Ziese M., Thorton M.J. *Spin Electronics.*– Berlin: Springer, 2001
4. Geim A.K., Novoselov K.S., *The Rise of Graphene*// Nature Materials.- 2007.- vol.6.-p.183-191
5. Третяк О.В., Львов В.А., Барабанов О.В., *Фізичні основи спінової електроніки* // ВПЦ Київський університет.- 2002
6. S.Sanvito and A.R.Rocha, *Molecular-Spintronics: the Art of Driving Spin through Molecules*// arXiv:cond-mat/0605239 v.1, May 2006

ДОДАТКИ

■ Додаток G

Міжзонна (комбінована) щільність станів

Проаналізуємо поведінку міжзонної щільності станів (її часто, особливо в російськомовній літературі, називають комбінованою щільністю станів), що визначає кількість станів електрон-діркової системи яка взаємодіє із фотонами у частотному діапазоні $[\omega, \omega + d\omega]$

$$\rho(\omega) = \frac{V}{(2\pi)^2} \int d\mathbf{k} \delta(E_c(\mathbf{k}) - E_v(\mathbf{k}) - \hbar\omega). \quad (\text{G.1})$$

Інтеграл, що фігурує в цій формулі, зручно обчислювати, здійснивши інтегрування по поверхні сталої енергії

$$E_c(\mathbf{k}) - E_v(\mathbf{k}) = E. \quad (\text{G.2})$$

Оскільки дельта-функція задовольняє співвідношенню

$$\delta(E' - E - \hbar\omega) = \frac{\delta(\mathbf{k}' - \mathbf{k})}{|\nabla_{\mathbf{k}} (E'(\mathbf{k}) - (E'(\mathbf{k})))|_{E' - E - \hbar\omega}}, \quad (\text{G.3})$$

то для міжзонної щільності станів отримаємо

$$\rho(\omega) = \frac{V}{(2\pi)^2} \int \frac{dS}{|\nabla_{\mathbf{k}} [E_c(\mathbf{k}) - (E_v(\mathbf{k}))]|_{E_c - E_v = \hbar\omega}}, \quad (\text{G.4})$$

де інтегрування ведеться по поверхні сталої енергії. Як видно із (G.4), інтеграл являє собою гладеньку функцію частоти за всіх значень ω , крім тих, які зануляють знаменник:

$$|\nabla_{\mathbf{k}} (E_c(\mathbf{k}) - (E_v(\mathbf{k})))|_{E_c - E_v = \hbar\omega} = 0. \quad (\text{G.5})$$

Зрозуміло, що за виконання умови (G.5) інтеграл повинен мати особливість, тобто або він сам буде розбіжним, або похідні за частотою від нього. Оскільки для квадратичних законів дисперсії (що завжди є справедливим в околі стелі валентної зони та дна зони провідності) $|\nabla_{\mathbf{k}} (E_c(\mathbf{k}) - E_v(\mathbf{k}))| = v_c(\mathbf{k}) - v_v(\mathbf{k})$, то умову (G.5) можна записати у вигляді

$$|v_c(\mathbf{k}) - v_v(\mathbf{k})| = 0. \quad (\text{G.5})$$

А це означає, що особливості в міжзонній щільності станів виникають, коли міжзонний перехід відбувається між станами з однаковими швидкостями носіїв. Точки в зоні Бриллюена, для яких виконується рівність (G.6), називаються критичними. Сінгулярності в щільності станів (отже й у коефіцієнті поглинання), пов'язані з такими критичними точками, називаються сінгулярностями Ван Хова. Очевидно, що

можуть існувати два типи критичних точок. Перший тип (критичні точки першого роду) відповідає ситуації, коли обидві швидкості у зоні провідності та валентній зоні одночасно зануляються, тобто

$$\mathbf{k} \in \mathbf{k}_C^1: \quad v_c(\mathbf{k}) = 0, \quad v_v(\mathbf{k}) = 0; \quad (G.7)$$

другий тип (критичні точки другого роду) \mathbf{k}_C^2 :

$$\mathbf{k} \in \mathbf{k}_C^2: \quad v_c(\mathbf{k}) \neq 0, \quad v_v(\mathbf{k}) \neq 0, \quad v_c(\mathbf{k}) - v_v(\mathbf{k}) = 0. \quad (G.8)$$

Умови (G.7) виконуються тільки в силу тих чи інших співвідношень симетрії у зоні Брилюєна. Наприклад, оскільки енергія зонного електрона є парною функцією імпульсу, то градієнти мають бути непарними функціями імпульсу. Тоді, якщо в центрі зони Брилюєна енергетичні зони не вироджені, то там розташовується критична точка першого роду. В інших довільних точках, що не відповідають симетричним точкам зони Брилюєна, умови (G.7) можуть виникати лише випадково. Критичні точки другого роду у принципі можуть існувати в будь-яких точках зони Брилюєна, оскільки різниця $\nabla_{\mathbf{k}} E_C - \nabla_{\mathbf{k}} E_V$ може занулятися у довільних точках зони Брилюєна, наявність яких визначається конкретним виглядом закону дисперсії. При цьому, якщо в критичних точках першого роду особливості мають як міжзонна щільність станів, так і щільності станів у валентній зоні та зоні провідності, то в критичних точках другого роду особливу поведінку виявляє тільки міжзонна щільність станів.

Оскільки критичні точки першого роду визначаються нулями градієнтів енергії, то вони відповідають стелі та дну валентної зони та зони провідності, а також сідловин точкам законів дисперсії. Розглянемо поведінку міжзонної щільності станів в околі деякої критичної точки \mathbf{k}_0 . У цій точці градієнт енергії по хвильовому вектору дорівнює нулю, то, вибираючи систему координат так, щоб її осі були направлені вздовж головних осей тензора ефективної маси, поблизу цієї точки закон дисперсії можна представити у вигляді

$$E(\mathbf{k}) = E(0) + \sum_{i=1}^3 \frac{1}{2} m_i^{-1} \hbar^2 (k_i - k_i^0)^2 + \dots, \quad (G.9)$$

де через m_i^{-1} позначено компоненти тензора оберненої ефективної маси в точці $\mathbf{k} = \mathbf{k}_0$. Залежно від знаків компонентів тензора ефективної маси можуть існувати чотири типи особливих точок поверхні $E(\mathbf{k})$:

- a) особлива точка M_0 відповідає мінімуму поверхні – усі три компоненти тензора ефективної маси позитивні;
- b) особлива точка M_1 відповідає сідловій точці поверхні – одна з компонентів тензора ефективної маси від'ємна, а дві інші – позитивні;

с) особлива точка M_2 відповідає сідловій точці поверхні – одна з компонент тензора ефективної маси позитивна, а дві інші – від'ємні;

d) особлива точка M_3 відповідає максимуму поверхні – усі три компоненти тензора ефективної маси від'ємні.

Підставимо енергію (G.9) до формули для міжзонної щільності станів (G.1), обмежуючи область інтегрування невеликим околom критичної точки. Зробимо заміну змінних

$$\hbar(k_i - k_i^0) = q_i(2m_i)^{1/2}. \quad (\text{G.10})$$

Дістанемо

$$\rho(\omega) = \frac{2^{5,2} |m_1 m_2 m_3|^{3/2}}{8\pi^3 \hbar^3} \cdot I, \quad (\text{G.11})$$

де

$$I = \begin{cases} \int d\mathbf{q} \delta(q^2 + E(0) - \hbar\omega), & \text{точка } M_0, \\ \int d\mathbf{q} \delta(-q^2 + E(0) - \hbar\omega), & \text{точка } M_3, \\ \int d\mathbf{q} \delta(q_x^2 + q_y^2 - q_z^2 + E(0) - \hbar\omega), & \text{точка } M_1, \\ \int d\mathbf{q} \delta(-q_x^2 - q_y^2 + q_z^2 + E(0) - \hbar\omega), & \text{точка } M_2. \end{cases} \quad (\text{G.12})$$

У цих формулах $E(0) = E_C(\mathbf{k}_0) - E_V(\mathbf{k}_0)$.

Інтеграл для точки типу M_0 аналогічний інтегралам, які зустрічатись при обчисленні щільності станів напівпровідника із параболічним законом дисперсії. Це дає підставу написати для міжзонної щільності станів

$$\rho(\omega) \sim \begin{cases} \sqrt{\hbar\omega - E(0)}, & \hbar\omega \geq E(0), \\ 0, & \hbar\omega < E(0). \end{cases} \quad (\text{G.13})$$

Як видно з цієї формули, у разі, коли критичною є точка центру зони Бриллюена, формула (G.13) описує поглинання в прямозонному напівпровіднику біля краю смуги поглинання.

Аналогічний результат легко отримати для точки M_3 , оскільки дельта-функція є парною функцією аргументу, тобто

$$\delta(-q^2 + E(0) - \hbar\omega) = \delta(q^2 - E(0) + \hbar\omega),$$

$$\rho(\omega) \sim \begin{cases} \sqrt{\hbar\omega - E(0)}, & \hbar\omega \leq E(0), \\ 0, & \hbar\omega > E(0). \end{cases} \quad (\text{G.14})$$

Для обчислення величини I в точках M_1 і M_2 перейдемо до циліндричних координат, у результаті дістанемо $d\mathbf{q} = u du d\varphi dq_z$. Розглянемо для визначеності особливості в точці M_1 . Оскільки аргумент дельта-функції не залежить від кута, інтегрування за кутом проводиться безпосередньо. Отримаємо

$$I = 2\pi \int_{q_1}^{q_2} u du \int dq_z \delta(E(0) - \hbar\omega + u^2 - q_z^2). \quad (\text{G.15})$$

Для подальшого використаємо відоме співвідношення (див 8.3)

$$\delta(A^2 - X^2) = \frac{1}{2A} [\delta(X - A) + \delta(X + A)]$$

маємо

$$I = 2\pi \int_{q_1}^{q_2} u du \int (E(0) - \hbar\omega + u^2)^{-1/2}. \quad (\text{G.16})$$

У формулах (G.15) і (G.16) межі інтегрування q_1 та q_2 можна знайти з умови $(E(0) - \hbar\omega + u^2) > 0 - q_1$

$$q_1 = \begin{cases} \sqrt{\hbar\omega - E(0)}, & \hbar\omega \geq E(0) \\ 0, & \hbar\omega < E(0) \end{cases} \quad (\text{G.17})$$

та з розміру області навколо особливої точки, де розкладення (G.9) справедливе – q_2 . Обчислення інтеграла у (G.16) із врахуванням (G.17) дає

$$\rho(\omega) \sim \begin{cases} \sqrt{E(0) - \hbar\omega + q_2^2} - \sqrt{E(0) - \hbar\omega}, & \hbar\omega \leq E(0), \\ \sqrt{E(0) - \hbar\omega + q_2^2}, & \hbar\omega > E(0). \end{cases} \quad (\text{G.18})$$

Для точки M_2 аналогічні обчислення дають

$$\rho(\omega) \sim \begin{cases} \sqrt{\hbar\omega - E(0) + q_2^2}, & \hbar\omega < E(0), \\ \sqrt{\hbar\omega - E(0) + q_2^2} - \sqrt{\hbar\omega - E(0)}, & \hbar\omega \geq E(0). \end{cases} \quad (\text{G.19})$$

Таким чином, міжзонна щільність станів у точці $\hbar\omega = E(0)$ має залом. Схематично поведінку щільності станів поблизу точок Ван Хова подано на рис. G1.

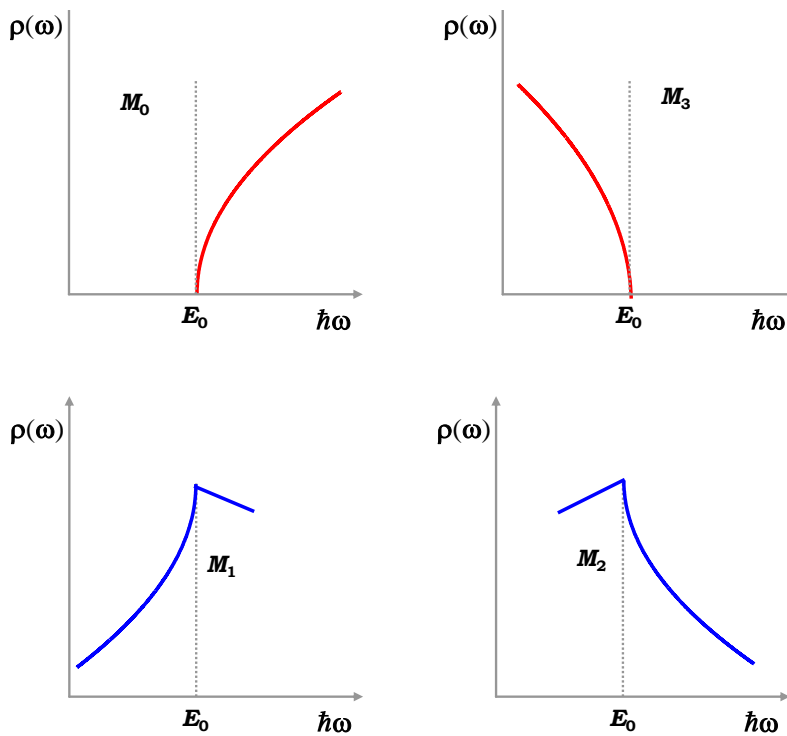


Рис. G1. Міжзонна щільність станів поблизу особливих точок. Індекс у літери **M** вказує на кількість негативних коефіцієнтів у розкладенні різниці енергій (формула (G. 9))

■ Додаток Н

Електрон в квантовій ямі

У розд. 15.3 було показано, що рух електрона в квантовій ямі може бути зведений до руху в площині ями та поперечного руху. Рух електрона в площині ями характеризується ефективною масою m^* і двовимірним хвильовим вектором $\mathbf{k}_{\parallel} = (k_x, k_y)$. У результаті визначення енергетичних станів в ямі зводиться до розв'язку одновимірного рівняння Шредингера

$$\left[-\frac{\hbar^2}{2m^*} \frac{d^2}{dz^2} + V(z) \right] \chi(z) = \epsilon \chi(z). \quad (\text{H.1})$$

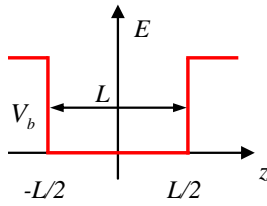


Рис. Н1. Потенціал квантової ями

Згідно із загальними властивостями рівняння Шредінгера можна очікувати наявності двох типів розв'язків цього рівняння: за $\epsilon < V_b$ електрони будуть зв'язані в ямі, а за $\epsilon > V_b$ рух електрона буде необмежений і стани – незв'язаними. Оскільки потенціал при $z = \pm L/2$ є розривним, для виконання

умов нерозривності хвильових функцій необхідно зшити розв'язок та його похідну на межах розриву потенціалу. Крім того, розв'язок має бути обмеженою функцією за умов $z \rightarrow \pm \infty$. За $\epsilon < V_b$ ззовні ями розв'язок матиме вигляд

$$\chi(z) = \begin{cases} Ae^{-k_b(z-L/2)} & \text{за } z \geq L/2, \\ Be^{k_b(z+L/2)} & \text{за } z \leq -L/2, \end{cases} \quad (\text{Н.2})$$

де $k_b = \sqrt{-2m^*(\epsilon - V_b)/\hbar^2}$. Розв'язок усередині ями буде простою комбінацією плоских хвиль

$$\chi(z) = C \cdot \cos k_w z + D \cdot \sin k_w z \quad \text{за } |z| \leq L/2, \quad (\text{Н.3})$$

де $k_w = \sqrt{-2m^* \epsilon / \hbar^2}$, A, B, C, D є константами, що визначаються з умов зшивання хвильової функції та її похідних у точках розриву потенціалу. Оскільки задача є симетричною відносно $z = 0$, можна вибрати або парну або непарну комбінації розв'язків (Н.2) і (Н.3). Виберемо парні розв'язки. Тоді з умови неперервності хвильової функції маємо: $A = B$, і розв'язок матиме вигляд

$$\chi(z) = \begin{cases} C \cos k_w z & \text{за } |z| \leq L/2, \\ Ae^{\pm k_b(z \pm L/2)} & \text{за } |z| \geq L/2. \end{cases} \quad (\text{Н.4})$$

Знаки мінус і плюс відповідають позитивним і негативним значенням координати z . Наступним кроком є зшивання функцій та їхніх похідних у точках $z = \pm L/2$. У результаті дістанемо два рівняння для визначення двох коефіцієнтів A та C

$$C \cos k_w L/2 = A, \quad (\text{Н.5})$$

$$C k_w \sin k_w L/2 = A k_b. \quad (\text{Н.6})$$

Це однорідна система двох рівнянь із двома невідомими. Для існування нетривіальних розв'язків потрібно виконання умови занулення детермінанта системи. Звідси маємо

$$\operatorname{tg}k_w L/2 = k_b/k_w. \quad (\text{H.7})$$

Аналогічно для непарного розв'язку отримуємо умову

$$\operatorname{ctg}k_w L/2 = -k_b/k_w. \quad (\text{H.8})$$

Ці рівняння можна розв'язати чисельно, але наочніше – розв'язати їх графічно. Для цього перепишемо (H.7) і (H.8) у дещо іншому вигляді. А саме

$$(\text{H.7}) \rightarrow \operatorname{tg}k_w L/2 = \sqrt{\frac{V_b}{\epsilon} - 1}; \quad \frac{\sin^2 \varphi}{\cos^2 \varphi} = \frac{V_b}{\epsilon} - 1;$$

або

$$\frac{1 - \cos^2 \varphi}{\cos^2 \varphi} = \frac{1}{\cos^2 \varphi} - 1 = \frac{V_b}{\epsilon} - 1 \rightarrow \cos k_w L/2 = \pm \sqrt{\frac{\epsilon}{V_b}}.$$

Це означає, що для $\operatorname{tg}k_w L/2 > 0$ мають виконуватись умови

$$\cos k_w L/2 = \pm k_w/k_0, \quad (\text{H.9})$$

де $k_0 = \sqrt{2m^*V_b/\hbar^2}$. У цьому рівнянні знак плюс вибирається для позитивних значень $\cos k_w L/2$, а мінус – для негативних.

Аналогічно можна записати й рівняння (H.8)

$$\sin k_w L/2 = \pm k_w/k_0 \quad \text{для} \quad \operatorname{ctg}k_w L/2 < 0. \quad (\text{H.10})$$

Тут знак плюс вибирається, якщо $\sin k_w L/2 > 0$ і мінус – якщо $\sin k_w L/2 < 0$. Ліві (L) та праві (R) частини рівнянь (H.9) і (H.10) можна представити в одному графіку як функції величини k_w . При цьому частини кривих в інтервалах

$$\frac{2\pi(l-1)}{L} < k_w < \frac{2\pi(l-1/2)}{L} \quad (\text{H.11})$$

відповідають парним розв'язкам (тобто (H.9)), а частини кривих в інтервалах

$$\frac{2\pi(l-1/2)}{L} < k_w < \frac{2\pi l}{L} \quad (\text{H.12})$$

– непарним розв'язкам (тобто (Н.10)). Тут $l = 1, 2, 3, \dots$. Праві частини обох рівнянь є прямими лініями із кутовим коефіцієнтом k_0^{-1} , що проходять через початок системи координат. Ліва частина рівнянь є синусом або косинусом. Ці лінії перетинаються в точках, що дають розв'язки $k_{w,n}$, для яких початкова задача має ненульові розв'язки. При цьому всі розв'язки знаходяться у першому квадранті, оскільки позитивні та негативні значення L і R відповідно перебувають у першому квадранті. Для аналізу цих розв'язків слід зазначити, що задача характеризується двома незалежними параметрами – глибиною V_b і шириною L ями. При цьому можна змінювати один параметр, зафіксувавши інший. Змінюватимемо параметр V_b . Це автоматично означає, що змінюватиметься величина k_0 . Таким чином ліва частина рівнянь (Н.9) і (Н.10) буде незмінною, а прями в правій частині рівнянь змінюватимуть свій нахил. Видно, що за малих значень k_0 кут нахилу прямої великий. Це означає, що може існувати лише одна точка перетину прямої і косинусоїди. Таким чином, маємо лише один розв'язок рівняння. Нехай він буде рівним $k_{w,1}$. Цьому розв'язку відповідатиме перший енергетичний рівень частинки в ямі

$$\epsilon_1 = \frac{\hbar^2 k_{w,1}^2}{2m^*}. \tag{Н.13}$$

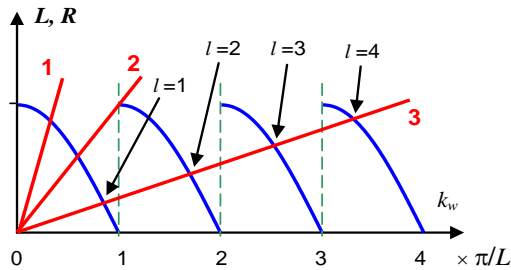


Рис.Н2. Графічний розв'язок рівнянь (Н.9) та (Н.10). Пряма 2 відповідає критичному значенню k_0 коли другий рівень з'являється в ямі як результат розв'язку рівняння (Н.10). Лінії 3 відповідають середні проміжні значення k_0 , що призводять до появи чотирьох енергетичних рівнів

При зростанні k_0 пряма все більше нахилиться до вісі OX , і при перетині її з наступним відрізком косинусоїди виникає новий енергетичний рівень (див. лінію 2 на рис. Н2). При цьому, оскільки умовою початку перетину прямої 2 і відповідного відрізка синусоїди ($l = 2$) є $k_w = 3\pi/2L$, то із (Н.10) маємо $\sin 3\pi/2 \approx -k_w/k_0 = L$, звідки

$$\epsilon_2 \approx V_b. \tag{Н.14}$$

За подальшого зростання k_0 перший і другий енергетичні рівні стають глибшими та виникає третій рівень і т. д. Дійсно, нові рівні виникають, якщо параметр

$$q = \sqrt{\frac{2m^*V_bL^2}{\pi^2\hbar^2}} \quad (\text{H.15})$$

набуває цілочисельного значення так, що число рівнів з енергією $\epsilon < V_b$

$$1 + \text{int} \left[\sqrt{\frac{2m^*V_bL^2}{\pi^2\hbar^2}} \right]. \quad (\text{H.16})$$

Додаток I

Матриця переходу (Т-матриця). Ідея та застосування

Для того, щоб чітко уявити ідею Т-матриці, розглянемо розсіяння квантової частинки на деякому потенціалі. Нехай, наприклад, це буде східцеподібний потенціальний бар'єр висотою V_b , розташований при $z = 0$ (рис. I1). Нехай частинка має енергію $E > V_b$. Введемо хвильові вектори руху частинки в передбар'єрній $k = \sqrt{2m^*E/\hbar^2}$ і підбар'єрній $k_b = \sqrt{2m^*(E - E_0)/\hbar^2}$ областях. Тоді хвильову функцію частинки можна записати у вигляді двох плоских хвиль, що рухаються назустріч одна одній:

$$\psi(z) = \begin{cases} Ae^{ikz} + Be^{-ikz}, & z \leq 0 \\ Ce^{ik_bz} + De^{-ik_bz}, & z \geq 0 \end{cases} \quad (\text{I.1})$$

Для виконання вимоги неперервності хвильової функції та її похідної прирівнюємо значення хвильових функцій та їхніх похідних на бар'єрі. Маємо

$$A + B = C + D, \quad k(A - B) = k_b(C - D) \quad (\text{I.2})$$

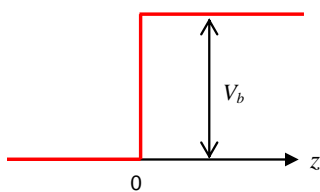


Рис. I1. Потенціальний бар'єр

Звідси можна отримати зв'язок між коефіцієнтами, що визначають амплітуду хвильових процесів у перед- та підбар'єрній областях,

$$C = \frac{1}{2} \left(1 + \frac{k}{k_b} \right) A + \frac{1}{2} \left(1 - \frac{k}{k_b} \right) B \quad (\text{I.3})$$

$$D = \frac{1}{2} \left(1 - \frac{k}{k_b} \right) A + \frac{1}{2} \left(1 + \frac{k}{k_b} \right) B \quad (1.4)$$

Припустимо, що частинка рухається з області $z < 0$ направо (див. рис.11). Вважаючи амплітуду набігаючої хвилі рівною одиниці, маємо, що амплітуда відбитої хвилі

$$r = B = \frac{k - k_b}{k + k_b}, \quad (1.5)$$

а амплітуда хвилі, що пройшла бар'єр,

$$t = C = \frac{2k}{k + k_b}. \quad (1.6)$$

Визначимо коефіцієнт проходження як відношення потоку частинок, що пройшли бар'єр, до потоку частинок, що падають на бар'єр. Тоді, аналогічно до (16.57), отримаємо

$$T = \frac{(\hbar k_b / m^*)}{(\hbar k / m^*)} |t|^2 = \frac{4kk_b}{(k + k_b)^2}. \quad (1.7)$$

Коефіцієнт відбиття є відношенням потоку частинок, відбитих від бар'єру, до потоку частинок, що падає на бар'єр,

$$R = \frac{(\hbar k / m^*)}{(\hbar k / m^*)} |r|^2 = \frac{(k - k_b)^2}{(k + k_b)^2}. \quad (1.8)$$

Легко побачити, що при цьому зберігається неперервність потоку

$$R + T = 1. \quad (1.9)$$

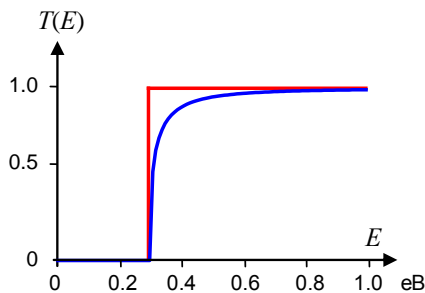


Рис. 12. Коефіцієнт проходження квантовою частинкою прямокутного бар'єру висотою 0,3 еВ (синя крива). Червона лінія відповідає коефіцієнту проходження класичної частинки

Поведінка коефіцієнта проходження бар'єру (0,3 eV) частинкою подано на рис. 12. Зрозуміло, що у випадку $E < V_b$ у кінцевих формулах треба зробити заміну

$$k_b \rightarrow ik_2 = i\sqrt{2m^*(V_b - E)/\hbar^2}.$$

Процес, що описується хвильовою функцією (1.1) з коефіцієнтами (1.3) можна подати ді-

аграмою (рис. 13).

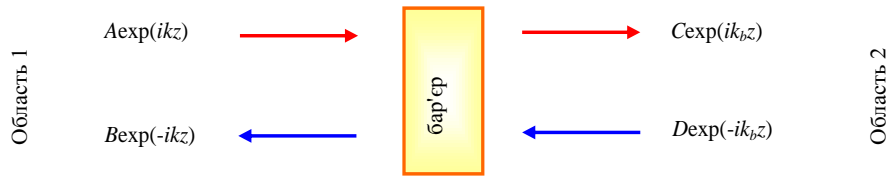


Рис.13. Діаграма процесу розсіяння квантової частинки на одnobар'єрній перешкоді

Систему рівнянь (1.3) можна переписати в матричній (або в операторній) формі

$$\begin{pmatrix} C \\ D \end{pmatrix} = \mathbf{T}^{(21)} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} T_{11}^{(21)} & T_{12}^{(21)} \\ T_{21}^{(21)} & T_{22}^{(21)} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}. \quad (1.10)$$

Для одноактного процесу розсіяння переваг введення T -матриці явно не спостерігається. Але якщо розглянемо складніший процес послідовного розсіяння на двох бар'єрах, як показано діаграмою на рис. 14, то впевнимось у перевазі методу. Амплітуди хвиль частинок, що

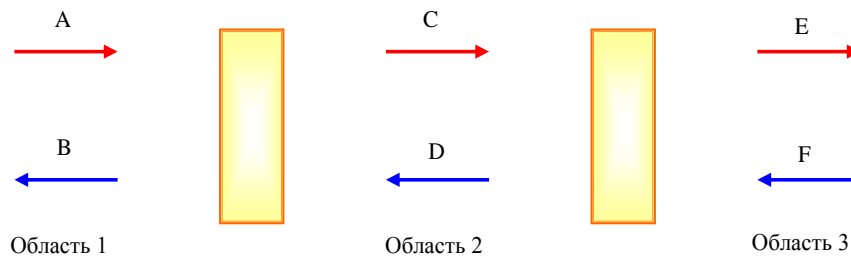


Рис.14. Діаграма розсіяння квантової частинки на двобар'єрній перешкоді

пройшли з області 1 до області 2, задаються амплітудами хвиль у першій області та T -матрицею $\mathbf{T}^{(21)}$. Амплітуди хвиль в області 3, у свою чергу, зв'язані з амплітудами хвиль в другій області матрицею $\mathbf{T}^{(32)}$. Таким чином,

$$\begin{pmatrix} C \\ D \end{pmatrix} = \mathbf{T}^{(21)} \begin{pmatrix} A \\ B \end{pmatrix}, \quad \begin{pmatrix} E \\ F \end{pmatrix} = \mathbf{T}^{(32)} \begin{pmatrix} C \\ D \end{pmatrix} \quad (1.11)$$

Тепер легко зв'язати амплітуди хвиль у третій області з амплітудами хвиль у першій області

$$\begin{pmatrix} E \\ F \end{pmatrix} = \mathbf{T}^{(32)}\mathbf{T}^{(21)} \begin{pmatrix} A \\ B \end{pmatrix} \equiv \mathbf{T}^{(31)} \begin{pmatrix} A \\ B \end{pmatrix}. \quad (1.12)$$

Тепер можна легко узагальнити метод обчислення коефіцієнту проходження квантовою частинкою багатобар'єрної структури. Нехай

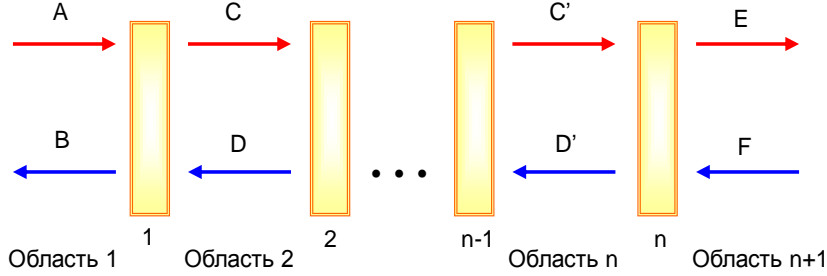


Рис.15. Діаграма розсіяння квантової частинки на n -бар'єрній перешкоді

частинка рухається у структурі, що складається з n бар'єрів, коефіцієнти проходження яких відомі (рис. 15). Розглядаючи послідовно процеси розсіяння на кожному бар'єрі, як у випадку двобар'єрної структури, можна записати

$$\begin{pmatrix} E \\ F \end{pmatrix} = \mathbf{T}^{(n,n-1)}\mathbf{T}^{(n-1,n-2)} \dots \mathbf{T}^{(32)}\mathbf{T}^{(21)} \begin{pmatrix} A \\ B \end{pmatrix} \equiv \mathbf{T}^{(n1)} \begin{pmatrix} A \\ B \end{pmatrix} \quad (1.13)$$

Отже, для того, щоб знайти амплітуду хвильового процесу за n -разового розсіянні хвиль, необхідно просто знайти відповідну T -матрицю, що є простим добутком n матриць одноразового розсіяння.

Таким чином ми отримали дуже важливий результат, що становить основу методу T -матриці.

Коефіцієнт проходження n -бар'єрної структури дорівнює добутку матриць переходу кожного із бар'єрів

Повертаючись до процесу розсіяння частинки, що, рухаючись зліва направо, розсіюється потенціальною сходинкою, можемо записати

$$\begin{pmatrix} t \\ 0 \end{pmatrix} = \mathbf{T}^{(21)} \begin{pmatrix} 1 \\ r \end{pmatrix}. \quad (1.14)$$

При цьому

$$r = -\frac{T_{21}}{T_{22}}, \quad t = -\frac{T_{21}T_{22} - T_{12}T_{21}}{T_{22}}, \quad (1.15)$$

а T -матриця визначається такими елементами:

$$\mathbf{T}^{(21)} = \begin{pmatrix} 1 \\ 2k_b \end{pmatrix} \begin{pmatrix} k_b + k & k_b - k \\ k_b - k & k_b + k \end{pmatrix}. \quad (1.16)$$

Треба зауважити, що формули для T -матриці, отримані вище, записані для випадку, коли границя бар'єру розташована на початку координат. Для подальшого нам знадобиться явний вигляд T -матриці для бар'єра, що розташований у довільній точці координатної вісі, скажімо, $z = d$. Оскільки амплітуда хвильової функції змінюється лише при переході частинкою бар'єру, то зсув бар'єру на відстань d приведе лише до зміни фази хвильової функції. Тоді можна припустити, що тільки на такі фазові множники й відрізнятимуться компоненти T -матриць $\mathbf{T}(0)$ та $\mathbf{T}(d)$. Для знаходження такого зв'язку ми спочатку транслюємо систему координат на d так, що в новій системі координат маємо $z' = z - d$. Тепер падаюча хвиля, яка мала вигляд Ae^{ikz} , перетвориться на $Ae^{ikz'}e^{ikd}$. Тобто її амплітуда домножиться на фактор e^{ikd} . Аналогічно амплітуду хвилі, що рухається в протилежному напрямку, необхідно домножити на фактор e^{-ikd} . Цю операцію можна записати як добуток діагональної матриці на вектор амплітуд:

$$\begin{pmatrix} A' \\ B' \end{pmatrix} = \begin{pmatrix} e^{ikd} & 0 \\ 0 & e^{-ikd} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}. \quad (1.17)$$

Тепер амплітуда хвиль, праворуч бар'єру буде такою

$$\begin{pmatrix} C' \\ D' \end{pmatrix} = \mathbf{T}(0) \begin{pmatrix} e^{ikd} & 0 \\ 0 & e^{-ikd} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}. \quad (1.18)$$

Для того, щоб перейти в робочу систему координат, необхідно знову помножити амплітуду хвиль на фазові множники. Але тепер варто мати на увазі, що перехід є зворотнім і відбувається в бар'єрній області. Таким чином, маємо

$$\begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} e^{-ik_b d} & 0 \\ 0 & e^{ik_b d} \end{pmatrix} \begin{pmatrix} C' \\ D' \end{pmatrix} \quad (1.17)$$

Таким чином, маємо співвідношення

$$\mathbf{T}(d) = \begin{pmatrix} e^{-ik_d d} & 0 \\ 0 & e^{ik_d d} \end{pmatrix} \cdot \mathbf{T}(0) \cdot \begin{pmatrix} e^{ikd} & 0 \\ 0 & e^{-ikd} \end{pmatrix} \quad (1.18)$$

що зв'язує T -матрицю перешкоди при $z = 0$ з T -матрицею перешкоди при $z = d$.

Надамо деякі додаткові відомості про метод T -матриці. Для вивчення аналітичних властивостей T -матриці розглянемо рівняння Шредингера, що явно залежить від часу

$$\tilde{H}\Psi(t) = i\hbar \frac{\partial}{\partial t} \Psi(t) \quad (1.19)$$

Вважаючи, що гамільтоніан системи не залежить від часу та є дійсним, використаємо інваріантність системи за часом (вимога дійсності гамільтоніана порушується, якщо розглядається дія магнітного поля на систему. У цьому випадку інваріантність системи за часом вимагає зміни знаку при магнітному полі, тобто при $t \rightarrow -t$ ми маємо записати $\mathbf{B} \rightarrow -\mathbf{B}$)

$$\tilde{H}\Psi(-t) = i\hbar \frac{\partial}{\partial t} \Psi(-t). \quad (1.20)$$

Здійснюємо комплексне спряження

$$\tilde{H}\Psi^*(-t) = i\hbar \frac{\partial}{\partial t} \Psi^*(-t). \quad (1.21)$$

Оскільки (1.19) та (1.21) формально збігаються, то рівняння Шредингера задовольняється обома функціями $\Psi(t)$ і $\Psi^*(-t)$. Припустимо для простоти, що по обидва боки від бар'єру квантові стани частинки є однаковими та описуються хвилею, що розповсюджується вздовж вісі OZ із хвилевим вектором k , тобто хвильова функція має такий вигляд

$$\Psi_k(z, t) = Ae^{i(kz - \omega t)} \quad (1.22)$$

Операція обертання часу дає

$$\Psi_k^*(z, t) = A^* e^{-i(kz - \omega t)} \quad (1.23)$$

Оскільки знак при k змінився, це означає зміну напрямку руху частинки. Менш очевидним є комплексне спряження коефіцієнта A . Але якщо взяти до уваги, що залежність хвильової функції від часу має лишатись такою самою, як до розсіяння на бар'єрі, тобто $\propto e^{-i\omega t}$ (енергія зберігається!), то зрозуміло, що для цього необхідно, щоб у функцій (1.22) та (1.23) були різні коефіцієнти. Таким чином, за часової інверсії амплітуди хвильових функцій мають бути комплексно спряженими. Таким чином в області $z < 0$ (рис. 11) хвильова функція має вигляд

$$Ae^{ikz} + Be^{-ikz}. \quad (1.24)$$

Інваріантність відносно інверсії в часі дає

$$B^* e^{ikz} + A^* e^{-ikz}. \quad (1.24)$$

Обидві функції є розв'язками рівняння Шредингера.

Визначення T -матриці дозволяє знайти коефіцієнти хвильової функції в області $z > 0$.

$$\begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} T_{11}A + T_{12}B \\ T_{21}A + T_{22}B \end{pmatrix} \quad (1.25)$$

Інваріантність відносно інверсії в часі вимагає, щоб

$$\begin{pmatrix} D^* \\ C^* \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} B^* \\ A^* \end{pmatrix} = \begin{pmatrix} T_{11}B^* + T_{12}A^* \\ T_{21}B^* + T_{22}A^* \end{pmatrix} \quad (1.26)$$

Але фізично система не змінилася, як не змінюється й T -матриця. Формально замінимо у векторах формули (1.26) елементи та візьмемо комплексне спряження. Маємо

$$\begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} T_{22}^*A + T_{21}^*B \\ T_{12}^*A + T_{11}^*B \end{pmatrix} = \begin{pmatrix} T_{22}^* & T_{21}^* \\ T_{12}^* & T_{11}^* \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}. \quad (1.27)$$

Порівнюючи (1.25) і (1.27), маємо зокрема

$$T_{22} = T_{11}^*, \quad T_{21} = T_{12}^* \quad (1.28)$$

Таким чином, тільки два із чотирьох елементів T -матриці є незалежними, тому її можна записати у формі

$$\mathbf{T} = \begin{pmatrix} T_{11} & T_{12} \\ T_{12}^* & T_{11}^* \end{pmatrix}. \quad (1.29)$$

Подальший зв'язок між двома компонентами матриці можна отримати, виходячи з умови збереження потоку частинок по обидва боки перешкоди. Можна припустити, що по обидва боки перешкоди стани частинки описуються однаковими хвильовими векторами. Тоді

$$|A|^2 - |B|^2 = |C|^2 - |D|^2. \quad (1.30)$$

Визначаючи коефіцієнти C та D через A та B і компоненти T -матриці, маємо

$$|T_{11}|^2 - |T_{12}|^2 = 1 = \det |\mathbf{T}|. \quad (1.31)$$

Ці співвідношення зводять число незалежних реальних величин, що визначають трансфер-матрицю, до трьох. Зважаючи на це, можна записати вирази до амплітудних коефіцієнтів проходження та відбиття як

$$r = -\frac{T_{12}^*}{T_{11}^*} \quad \text{і} \quad t = \frac{1}{T_{11}^*}. \quad (1.32)$$

Таким чином, можна записати T -матрицю у термінах r та t

$$\mathbf{T} = \begin{pmatrix} 1/t^* & -r^*/t^* \\ -r/t & 1/t \end{pmatrix} \quad (1.33)$$

із додатковою умовою $|r|^2 + |t|^2 = 1$, що визначає закон збереження потоку частинок. Іноді корисно використовувати T -матрицю, що зв'яже амплітуди хвиль у лівій частині від перешкоди з амплітудою в правій частині

$$\begin{pmatrix} B \\ A \end{pmatrix} = \mathbf{T}^{(12)} \begin{pmatrix} D \\ C \end{pmatrix}. \quad (1.34)$$

Для знаходження компонентів цієї матриці слід зауважити, що $\mathbf{T}^{(12)}$ не є простою інверсією матриці $\mathbf{T}^{(21)}$. Нехай компоненти матриці $\mathbf{T}^{(12)}$ позначаються як штриховані - T'_{ij} . Тоді, порівнюючи (1.26) з (1.34), можна записати

$$\begin{aligned} \mathbf{T}^{(12)} &= \begin{pmatrix} T'_{11} & T'_{12} \\ T'_{21} & T'_{22} \end{pmatrix} = \begin{pmatrix} 1/t^* & -r^*/t^* \\ -r'/t' & 1/t' \end{pmatrix} \\ &= \begin{pmatrix} T_{11} & -T_{12}^* \\ -T_{12} & T_{11}^* \end{pmatrix} = \begin{pmatrix} 1/t^* & r/t \\ r^*/t^* & 1/t \end{pmatrix} \end{aligned} \quad (1.35)$$

Отже, для зв'язку між компонентами реверсивної і прямої T -матриць маємо

$$T'_{11} = T_{11} \quad \text{і} \quad T'_{12} = T'_{12}. \quad (1.36)$$

Аналогічно зв'язок між амплітудними коефіцієнтами відбиття та проходження запишемо у вигляді

$$t' = t \quad \text{і} \quad r' = -\frac{t}{t^*} r^*. \quad (1.37)$$

■ Додаток J

Формула Кубо-Грінвуда

Вважатимемо, що вироджений електронний газ перебуває в неупорядкованій системі за нульової температури. Стани з енергіями меншими за E_F заповнені. Нехай власними хвильовими функціями електрона з енергією E в неперіодичному полі з відповідними граничними умовами є $\psi_E(\mathbf{r})$, нормовані на один електрон в об'ємі Ω . Припустимо, що на електрон діє змінне електричне поле частоти ω

$$E = F \cos \omega t \quad (J.1)$$

так, що потенціальна енергія електрона в полі $e\chi F \cos \omega t$. Тоді ймовірність того, що електрон за одиницю часу перейде зі стану E в деякий стан з енергією $E + \hbar\omega$

$$W_{E+\hbar\omega,E} = \frac{\pi}{2\hbar} e^2 F^2 \left\langle |x_{E+\hbar\omega,E}|^2 \right\rangle_{E+\hbar\omega,E} \Omega N(E + \hbar\omega), \quad (J.2)$$

де матричний елемент

$$x_{E',E} = \int \psi_{E'}^* x \psi_E d^3x. \quad (J.3)$$

Беручи до уваги, що

$$x_{E',E} = -\frac{\hbar}{m\omega} D_{E',E}, \quad (J.4)$$

де $D_{E',E} = \int \psi_{E'}^* \frac{\partial}{\partial x} \psi_E d^3x$, запишемо (J.2) у вигляді

$$W_{E+\hbar\omega,E} = \frac{\pi e^2 \hbar \Omega}{2m^2 \omega^2} F^2 \left\langle |D_{E+\hbar\omega,E}|^2 \right\rangle_{E+\hbar\omega,E} N(E + \hbar\omega). \quad (J.5)$$

Визначимо провідність так, щоб величина $(1/2)F^2 \sigma_E(\omega)$ дорівнювала середній швидкості втрат в одиниці об'єму, тобто мала сенс дисипативної функції. Домножимо (J.5) на число зайнятих станів в одиниці об'єму в інтервалі енергій $dE - N(E)f(E)dE$ та на ймовірність того, що стан з енергією $E + \hbar\omega$ не зайнятий електронами $1 - f(E + \hbar\omega)$, і на $\hbar\omega$ – енергію поля, що поглинається системою за кожен акт квантового переходу. Інтегруючи за всіма можливими значеннями енергії, отримуємо

$$\sigma_E(\omega) = \frac{2\pi e^2 \hbar^2 \Omega}{m^2 \omega} \int dE \{ f(E)[1 - f(E + \hbar\omega)] - f(E + \hbar\omega)[1 - f(E)] \} \left\langle |D_{E+\hbar\omega,E}|^2 \right\rangle_{E+\hbar\omega,E} N(E)N(E + \hbar\omega) \quad (J.6)$$

Другий доданок у фігурних дужках визначається енергією, що у вигляді фотона випромінюється за умов вимушених переходів униз. Вираз у фігурних дужках зводиться до $f(E) - f(E + \hbar\omega)$. Тоді формулу (J.6) можна переписати у вигляді

$$\sigma_E(\omega) = \frac{2\pi e^2 \hbar^3 \Omega}{m^2 \hbar \omega} \int dE \{ f(E) - f(E + \hbar\omega) \} \left\langle |D_{E+\hbar\omega,E}|^2 \right\rangle_{E+\hbar\omega,E} N(E)N(E + \hbar\omega)$$

За $T \rightarrow 0$ формула спрощується до

$$\sigma_E(\omega) = \frac{2\pi e^2 \hbar^3 \Omega}{m^2 \hbar \omega} \int_{E_F - \hbar\omega}^{E_F} dE \langle |D_{E+\hbar\omega}|^2 \rangle_{E+\hbar\omega, E} N(E)N(E+\hbar\omega). \quad (J.7)$$

Нижня межа інтегрування відповідає найнижчій енергії електрона, за якої він може поглинути квант випроміненням, а верхня межа – енергії Фермі при $T = 0$. Виконуючи граничний перехід $\omega \rightarrow 0$ за правилом

$$\begin{aligned} \lim_{\omega \rightarrow 0} \frac{1}{\hbar\omega} \int_{E_F - \hbar\omega}^{E_F} dE \langle |D_{E+\hbar\omega}|^2 \rangle_{E+\hbar\omega, E} N(E)N(E+\hbar\omega) = \\ = \frac{\hbar\omega}{\hbar\omega} \langle |D_{E+\hbar\omega}|^2 \rangle_{E+\hbar\omega, E} N(E)N(E+\hbar\omega) \Big|_{\omega \rightarrow 0} = \langle |D_E|^2 \rangle_E N^2(E), \end{aligned}$$

отримаємо із (J.7)

$$\sigma_E(0) = \frac{2\pi e^2 \hbar^3 \Omega}{m^2} \langle |D_E|^2 \rangle_E N^2(E). \quad (J.8)$$

Отриманий вираз називається формулою Кубо-Грінвуда. У цій формулі $D_E = \lim_{E' \rightarrow E} D_{E', E}$. Якщо $T = 0$, то провідність $\sigma(0)$

$$\sigma(0) = \{ \sigma_E(0) \}_{E=E_F}. \quad (J.9)$$

■ Додаток К

Плівки Ленгмюра-Блоджетт

Плівками Ленгмюра-Блоджетт називаються моно- та полі-молекулярні плівки, що отримуються за допомогою спеціальної технології, яка базується на ідеї самоорганізації плівки молекул на поверхні води. Для цього молекули повинні мати специфічну структуру – з одного кінця молекули повинні містити гідрофільну групу, а з іншого гідрофобну (рис. К1, К2) і таким чином молекула буде розташовуватись гідрофільною групою орієнтуючись до води а гідрофобний хвіст буде направлено у зовнішнє середовище (повітря). На рис. К1 показано структурну формулу стеаринової кислоти – $\text{HOOC-C}_{17}\text{H}_{35}$, що є типовим представником органічних молекул, один кінець якої містить

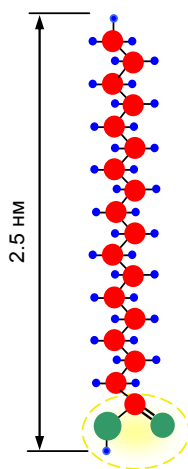


Рис. К1. Молекула стеаринової кислоти. Зеленими кульками позначено кисень, червоними – вуглець, синіми точками – водень

полярну частину COOH (виділено жовтим овалом), яка може віддавати свій протон воді, а сама заряджається від'ємно. Ця частина молекули, що називається гідрофільною частиною, таким чином буде утримуватись у воді силами кулонівського притяжіння. В той же час гідрофобний хвіст $\text{C}_{17}\text{H}_{35}$ виштовхується з води. Такі молекули, один кінець яких містить гідрофільну групу, а інший – гідрофобну, називаються амфифільними молекулами. Попадаючи на поверхню води такі речовини утворюють мономолекулярну плівку (рис. К2). Для того, щоб отримати мономолекулярну плівку амфифільних молекул, їх треба помістити на поверхню води. З цією метою амфифільну сполуку розчиняють у легко летючому розчиннику (наприклад, у бензолі). Цей розчин наносять на поверхню води – спеціальним шприцом або розпилюють. Розчинник випаровується, а амфифільна сполука утворює на поверхні води плівку. Кількість речовини підбирають так,

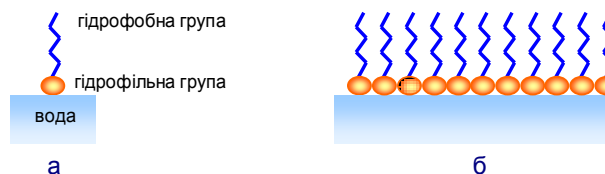


Рис. К2. Взаємодія амфифільних молекул з поверхнею води. а – орієнтація молекули гідрофільною групою до поверхні; б – утворення молекулярного моношару на поверхні води

щоб плівка була не товща за один моношар. Для того, щоб вона була однорідною (не острівцевою) спеціальною планкою у LB ванні (рис. К3) молекули притискають одну до одної. При цьому суцільність плівки контролюють характерним для даної речовини поверхневим натягом. Утворену таким чином мономолекулярну плівку занурюють пластинку, що буде відігравати роль підкладки LB-плівки. Таким чином під час занурення і витягування пластинки на обох її поверхнях утворюються однорідні в площині мономолекулярні плівки (рис. К4). Якщо

пластинку занурюють у воду, то гідрофобні хвости молекул виявляються напруженими до підкладки- отримується моношар X-типу.

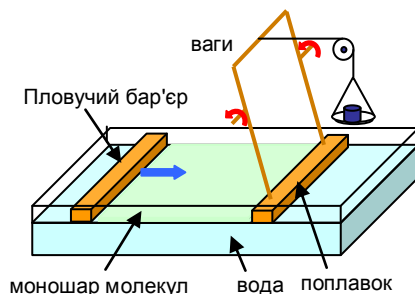


Рис. К3. Ленгмюрівська ванна для отримання ЛБ-плівок. Схема (ліворуч). Праворуч – ЛБ ванна у „чистій” кімнаті. 1 – система витягання зразка з плівкою; 2 – рухомий бар'єр, що дозволяє утримувати заданий поверхневий натяг речовини на поверхні води. Оператор шприцом наносить робочу речовину на поверхню води.

Якщо пластинку попередньо занурену у воду витягують, молекули моношару будуть орієнтованими гідрофільними групами до площини підкладки – формується моношар Z-типу. Повторюючи переніс одного моношару за іншим можна отримати багатошарові плівки Ленгмюра-Блоджетт.

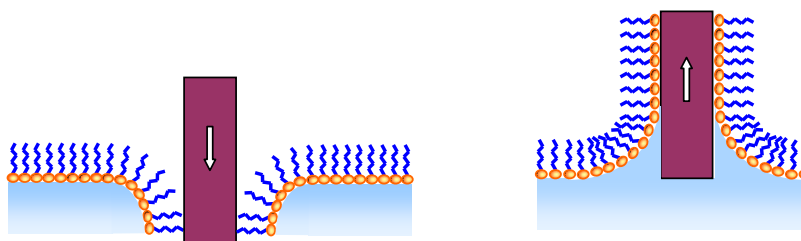


Рис. К4. Отримання мономолекулярних плівок методом занурення (ЛБ-плівка X-типу) або витягання (ЛБ-плівка Z-типу) пластинки-підкладки

Оскільки в площині плівки молекули взаємодіють між собою через так звані дисперсійні взаємодії (або Ван-дер-Ваальсові сили) то така плівка характеризується тим, що молекули майже не втрачають індивідуальні властивості, що дуже важливо для отримання ЛВ-плівок із заздалегідь заданими властивостями.

■ Додаток L

Окислювально-відтворювальні реакції

Окислювально-відтворювальні реакції – загальноприйняте позначення REDOX (*Reduction and Oxidation*) – це хімічні реакції, що протікають зі зміною ступені окислення атомів, що входять до складу реагуючих речовин. Зміна ступенів окислення відбувається шляхом перерозподілу електронів між атомом-окислювачем та атомом-відновлювачем. Нагадаємо, що в процесі окислення атом, що входить до речовини-реагенту втрачає електрон, а в процесі відновлення атом іншої речовини-реагенту набуває додатковий електрон. Таким чином, в процесі окислювально-відновлювальних реакції речовина-відновлювач віддає електрони, тобто, окислюється, а речовина-окислювач приєднує до себе електрони, тобто, відновлюється. При чому будь-яка окислювально-відновлювальна реакція являє собою єдність двох протилежних перетворювань – окислення та відновлення, що відбуваються одночасно і без відриву однієї від іншої.

Окисленням називається процес віддачі електронів атомом речовини, при цьому ступінь окислення атома підвищується. Ступінь окислення – поняття, що використовують в хімії для зручності опи-

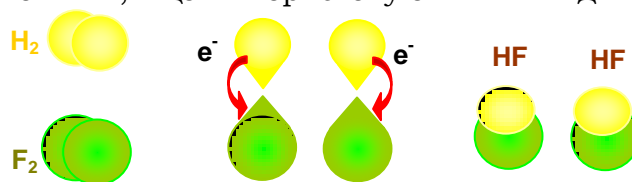
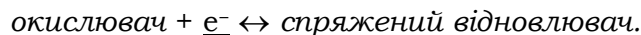


Рис. L1. Схема окислювально-відновної реакції між фтором і воднем

сання окислювально-відновлювальних реакцій і яке являє собою величину, що чисельно дорівнює формальному заряду (кількості електронів) що можна приписати елементу, виходячи з припущення, що всі електрони кожного з його зв'язків перейшли до більш електровід'ємного атома даної сполуки. При відновленні, таким, чином, ступінь окислення зменшується. Окислювач, приймаючи електрони, набуває відновлюючі властивості, перетворюючись у спряжений відновлювач, тобто

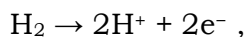


Відтворенням називається процес приєднання електронів атомом речовини, при цьому його ступінь окислення знижується. Відтворювач, віддаючи електрони, набуває окислювальних властивостей, перетворюючись у спряжений окислювач:

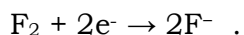
відтворювач - e^- \leftrightarrow спряжений окислювач.

Сукупність окислювача (відтворювача) з продуктом його перетворення складає спряжену окислювально-відтворювальну пару, а її взаємоперетворення є напівреакцією відтворення (окислення). В будь-якій окислювально-відтворювальній реакції беруть участь дві спряжені окислювально-відтворювальні пари, між якими має місце конкуренція за електрони, в результаті чого відбуваються дві напівреакції – одна пов'язана приєднанням електронів (відтворенням), інша – з віддачею електронів (окисленням). Слід пам'ятати, що в багатьох випадках окислювально-відтворювальних реакцій відбувається не повний переніс електронів, а тільки зміщення електронних хмар від одного атома до іншого. Як приклад окислювально-відтворювальної реакції можна навести реакцію між воднем та фтором $H_2 + F_2 \rightarrow 2HF$, що поділяється на дві напівреакції:

Окислення



та відтворення



На рис. L1 схематично показано хід цієї окислювально-відтворювальної реакції.