

В. П. Бригінець, С. О. Подласов

ФІЗИКА ДЛЯ БАКАЛАВРІВ

Міністерство освіти і науки України

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

В. П. Бригінець, С. О. Подласов

ФІЗИКА ДЛЯ БАКАЛАВРІВ

Навчальний посібник

Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для здобувачів ступеня бакалавра
за освітньою програмою «*Комп'ютерне моделювання фізичних процесів*»
спеціальності 104 Фізика та астрономія

Електронне мережеве навчальне видання

Київ
КПІ ім. Ігоря Сікорського
2026

УДК 53(075.8.034.2)
Б87

Автори: *Бригінець Валентин Петрович, канд фіз.-мат.*
Подласов Сергій Олександрович, старший викладач.

Рецензент *Кух А.М. доктор пед. наук, проф. каф. фізики Кам'янець-Подільського національного університету ім. Івана Огієнка*
Школа О. В. доктор пед. наук, завідувач кафедри фізики, математики та методики навчання Бердянського державного педагогічного університету

Відповідальний редактор *Савченко Д. В., докт. фіз.-ман. наук, доцент*

Гриф надано Методичною радою КПІ ім. Ігоря Сікорського
(протокол № 7 від 08.05.2026 р.)
за поданням вченої ради фізико-математичного факультету
(протокол № 3 від 08.04.2026 р.)

Бригінець В.П.

Б87 ФІЗИКА ДЛЯ БАКАЛАВРІВ [Електронний ресурс] : навч. посіб. для здобувачів ступеня бакалавра за освіт. програмою «G Інженерія, виробництво та будівництво» спец. G 4 Енерговиробництво / В.П. Бригінець, С. О. Подласов; КПІ ім. Ігоря Сікорського. – 2-ге вид., переробл. та доповн. – Електрон. текст. дані (1 файл). – Київ : КПІ ім. Ігоря Сікорського, 2026. – 618 с.

У посібнику представлено стислі, але вичерпні теоретичні відомості за програмою вивчення фізики в технічному університеті. В курсі представлені основні розділи загальної фізики та фундаментальні положення, але практичним застосуванням законів фізики приділено лише мінімально необхідну увагу. Навчальний матеріал є адаптованим до рівня математичної підготовки студентів, яким фізику починають викладати з першого навчального семестру. Зокрема, в багатьох випадках, довелося відмовитися від строгих доведень і обмежитися тільки якісними міркуваннями та коментуванням змісту положень, що розглядаються. З тієї ж причини математичні викладки, що наводяться, по можливості виконуються детально й супроводжуються роз'ясненнями загального змісту відповідних математичних величин та операцій.

УДК 53(075.8.034.2)

Реєстр. № НП 25/26-383. Обсяг 28 авт. арк.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
проспект Берестейський, 37, м. Київ, 03056
https://kpi.ua
Свідоцтво про внесення до Державного реєстру видавців, виготовлювачів і розповсюджувачів видавничої продукції ДК № 5354 від 25.05.2017 р.

© В. П. Бригінець, С. О. Подласов, О. В., 2026
© КПІ ім. Ігоря Сікорського, 2026

ПЕРЕДМОВА

Загальновідомо, що фізика є універсальною базою техніки — на закони фізики спираються всі прикладні й технічні науки технологічного та проектно-конструкторського спрямування. Тому вивчення фізики є конче необхідним для подальшої якісної фахової підготовки спеціалістів. Пропонований короткий курс загальної фізики є базовим і призначений для студентів усіх спеціальностей. Тому в курсі представлені основні розділи загальної фізики та фундаментальні положення, але практичним застосуванням законів фізики приділено лише мінімально необхідну увагу. Навчальний матеріал, представлений в курсі, є адаптованим до рівня математичної підготовки студентів, яким фізику починають викладати з першого навчального семестру. Зокрема, в багатьох випадках, довелося відмовитися від строгих доведень і обмежитися тільки якісними міркуваннями та коментуванням змісту положень, що розглядаються. З тієї ж причини математичні викладки, що наводяться, по можливості виконуються детально й супроводжуються роз'ясненнями загального змісту відповідних математичних величин та операцій. Така інформація, аби не перевантажувати основний текст, подається у вигляді виносок. Слід відзначити, що розуміння та засвоєння навчального матеріалу вимагає знання основних понять вищої математики – границі, похідної, диференціалу, первісної та визначеного інтегралу, уміння брати похідні та інтеграли, аналізувати функції на екстремум, розв'язувати прості диференціальні рівняння. Треба також знати основні поняття векторної алгебри, дії з векторами, вміти знаходити проекції векторів і оперувати ними.

Запропонований курс фізики включає наступні частини:

МЕХАНІКА	4
МОЛЕКУЛЯРНА ФІЗИКА І ТЕРМОДИНАМІКА	116
ЕЛЕКТРИКА І МАГНЕТИЗМ	191
КОЛИВАННЯ І ХВИЛІ	336
ЕЛЕМЕНТИ ОПТИКИ	421
ЕЛЕМЕНТИ КВАНТОВОЇ ФІЗИКИ	512

МЕХАНІКА

Зміст

МЕХАНІКА

1. КІНЕМАТИКА
1.1. Кінематика матеріальної точки
1.2. Кінематика твердого тіла
1.3. Контрольні запитання

2. ОСНОВНІ ЗАКОНИ ДИНАМІКИ
2.1. Інерціальні системи відліку. Перший закон Ньютона
2.2. Другий закон Ньютона. Сила
2..3. Третій закон Ньютона
2.4. Основне рівняння динаміки
2.5. Закони сил
2.6. Контрольні запитання

3. НЕІНЕРЦІАЛЬНІ СИСТЕМИ ВІДЛІКУ
3.1. Сили інерції
3.2. Сили інерції в обертовій системі відліку
3.3. Загальне рівняння руху точки в НСВ
3.4. Контрольні запитання

4. ДИНАМІКА СИСТЕМ
4..1. Імпульс системи
4.2. Закон збереження імпульсу
4.3. Рух тіл змінної маси (рівняння Мещерського)
4.4. Центр мас
4.5. Контрольні запитання

5. РОБОТА ТА ЕНЕРГІЯ
5.1. Робота та потужність сили
5.2. Кінетична енергія
5.3. Потенціальна енергія
5.4. Закон збереження механічної енергії
5.5. Зіткнення
5.6. Контрольні запитання

6. ЕЛЕМЕНТИ МЕХАНІКИ ТВЕРДОГО ТІЛА
6.1. Рівняння моментів
6.2. Закон збереження моменту імпульсу
6.3. Динаміка твердого тіла
6.4. Збереження моменту імпульсу відносно осі
6.5. Кінетична енергія обертального і плоского руху
6.6. Гіроскопи
6.7. Контрольні запитання

7. СПЕЦІАЛЬНАЯ ТЕОРІЯ ВІДНОСНОСТІ. КІНЕМАТИКА
7.1. Принцип відносності Галілея
7.2. Постулати СТВ, відносність простору та часу
7.3. Перетворення Лоренца
7.4. Перетворення швидкостей в СТВ, гранична швидкість
7.5. Інтервал
7.6. Контрольні запитання

8. СПЕЦІАЛЬНА ТЕОРІЯ ВІДНОСНОСТІ. ДИНАМІКА
8.1. Релятивістський імпульс і релятивістська маса
8.2. Основне рівняння релятивістської динаміки

8.3. Енергія та маса
8.4. Контрольні запитання

МЕХАНІКА

Основні закони фізики свого часу були встановлені на основі спостережень відповідних явищ та експериментів і тому, за невеликим винятком, мають обмежену сферу застосовності (мається на увазі, що фізика не торкається феномену життя та пов'язаних із ним специфічних процесів). Зокрема, закони *класичної механіки* поширюються тільки на повільні рухи частинок і тіл у макроскопічних областях простору. Швидкі рухи підпорядковані законам *спеціальної теорії відносності* (релятивістської механіки), а рухи частинок у мікроскопічних областях – законам *квантової механіки*. При цьому під повільними рухами розуміються рухи зі швидкостями, набагато меншими за швидкість світла $c = 3 \cdot 10^8$ м/с, а під швидкими — із швидкостями, що є сумірними із величиною c. Макроскопічна область простору — область, розміри якої набагато перевищують розміри атомів та молекул, а мікроскопічна — область порядку розмірів атома.

В класичній механіці ставляться і розв'язуються дві основні задачі:

1. Встановлення законів руху – співвідношень, які дозволяють визначати характер і характеристики руху будь-якого тіла чи системи тіл у залежності від його взаємодії з іншими тілами.

2. Знаходження таких загальних співвідношень між механічними характеристиками систем, які не залежать ані від складу та будови системи, ані від природи та характеру взаємодії між тілами системи.

Пошук загальних методів розв'язування першої задачі привело свого часу І. Ньютона до встановлення загальних законів руху (законів динаміки). Розв'язування ж другої задачі дозволило встановити закони збереження (імпульсу, енергії, моменту імпульсу).

Фізика є точною наукою — свої результати вона виражає не тільки словесно, а й за допомогою математичних співвідношень. Але властивості фізичних об'єктів і явищ є настільки багатогранними, що жодна теорія не може відобразити їх у всій повноті. Тому замість реальних об'єктів наука оперує фізичними моделями – ідеалізованими об'єктами, які відтворюють лише суттєві для розгляданих явищ властивості та фактори. В механіці основними моделями є *матеріальна точка* та *абсолютно тверде тіло*.

Матеріальною точкою називають тіло, розміри та форма котрого в умовах задачі, що розглядається, є не істотними. Отже, матеріальною точкою може бути не лише маленька частинка, а й велике тіло, приміром, Земля при розгляді її взаємодії з іншими небесними тілами та руху в космічному просторі. Тому в даному посібнику скрізь, за винятком механіки твердого тіла, терміни "частинка", "матеріальна точка", та "тіло" не розрізняються.

Абсолютно твердим тілом називається тіло, взаємне розташування точок якого є незмінним, тобто — тіло, що не деформується. При розгляді законів руху всі тіла вважаються абсолютно твердими.

Положення тіла у просторі та його рух можна визначити тільки відносно якогось іншого тіла, яке називають ***тілом відліку***. Задля можливості точного задання положення даного тіла відносно тіла відліку з останнім пов'язують ту чи іншу систему координат. Окрім того, для опису руху треба мати якийсь загально визнаний і узгоджений принцип фіксації моментів і вимірювання проміжків часу в усіх точках простору (принагідно відзначимо, що вимірювання часу спираються на періодичні процеси, наприклад, коливання маятника механічного годинника, або коливання напруги на конденсаторі в коливальному контурі електронного годинника, тощо.), або, як говорять, треба мати ***годинник***. При цьому

***сукупність тіла відліку, пов'язаної з ним системи координат і нерухомого відносно неї годинника називають* системою відліку.**

Отже, положення та рух тіла є визначеним тільки відносно конкретної системи відліку. В той же час вибір системи відліку є довільним (це стосується як тіла відліку, так і системи координат.) і визначається тільки зручністю опису руху в заданих конкретних умовах. Тому

положення в просторі й рух тіл є відносними за самою природою.

Далі розглянуті такі частини МЕХАНІКИ:

1. КІНЕМАТИКА
2. ОСНОВНІ ЗАКОНИ ДИНАМІКИ
3. НЕІНЕРЦІАЛЬНІ СИСТЕМИ ВІДЛІКУ
4. ДИНАМІКА СИСТЕМ
5. РОБОТА ТА ЕНЕРГІЯ
6. ЕЛЕМЕНТИ ДИНАМІКИ ТВЕРДОГО ТІЛА
7. СПЕЦІАЛЬНАЯ ТЕОРІЯ ВІДНОСНОСТІ. КІНЕМАТИКА
8. СПЕЦІАЛЬНАЯ ТЕОРІЯ ВІДНОСНОСТІ. ДИНАМІКА

1. КІНЕМАТИКА

Кінематика – це початковий розділ механіки, в якому встановлюються поняття та величини, що визначають рух, а також способи опису руху і загальні співвідношення між його характеристиками. Але в кінематиці не розглядаються умови та причини, що визначають характер руху тіл. Оскільки будь-яке тіло можна трактувати як систему матеріальних точок, спочатку, природньо, розглядається кінематика однієї матеріальної точки, а далі - кунематика твердого тіла.

1.1. Кінематика матеріальної точки

1.2. Кінематика твердого тіла

1.3. Контрольні запитання

1.1. Кінематика матеріальної точки

У цій частині ми не будемо розрізняти поняття “матеріальна точка”, “частинка” чи “тіло”. Спочатку розглянемо сособи описання руху і ті поняття, які при цьому використовують.

Існує три способи опису положення і руху точки в обраній системі відліку – векторний, координатний та природній – і відповідний набір кінематичних величин, які для цього використовуються. Розглянемо кожен із способів опису руху окремо, а також зв'язок між кінематичними характеристиками руху в різних системах відліку (перетворення Галілея).

Векторний спосіб опису руху ґрунтується на понятті радіуса-вектора і є найзручнішим для теорії, оскільки дозволяє лаконічно і повно відображати зміст кінематичних величин і зв’язок між ними.

Радіус-вектор, траєкторія, шлях, переміщення. Положення точки у просторі можна задавати її *радіусом-вектором* $\vec{r}$ **.** Радіус-вектор – це вектор, проведений із початку відліку в дану точку (рис.1.1). Рухома частинка, отже й кінець її радіуса-вектора, описує в просторі неперервну лінію, котра називається *траєкторією* руху. Можна сказати, що траєкторія є геометричним місцем точок кінця радіуса-вектора частинки, що рухається. Довжина відрізка траєкторії між двома даними точками називається ***шляхом***, пройденим частинкою за відповідний проміжок часу. Шлях визначає відстань, яку пройшла частинка вздовж траєкторії, але шлях не містить жодної інформації про її кінцеве положення. Тому для визначення зміни положення точки в просторі використовують ***переміщення*** $\Delta\vec{r}$ – вектор, який проводять із початкового в кінцеве положення точки на траєкторії (рис. 1.1). Очевидно, що

$$\Delta\vec{r} = \vec{r}_2 - \vec{r}_1. \qquad (1.1)$$

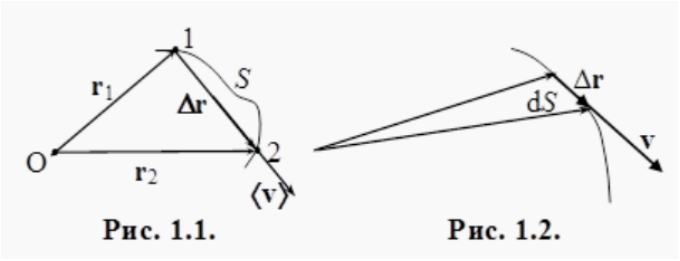


Рис. 1.1. Рис. 1.2.

Модуль вектора переміщення $|\Delta\vec{r}|$ дорівнює відстані між початковим та кінцевим положенням точки на траєкторії і в загальному випадку не дорівнює пройденому шляху S (рис. 1.1.):$|\Delta\vec{r}| \leq S$. Але переміщення $\mathrm{d}\vec{r}$ за нескінченно малий проміжок часу співпадає з відповідною нескінченно малою ділянкою траєкторії (рис. 1.2). Тому модуль вектора елементарного переміщення $|\mathrm{d}\vec{r}|$ і пройдений точкою шлях $\mathrm{d}S$ збігаються:

$$|\mathrm{d}\vec{r}| = \mathrm{d}S. \qquad (1.1а)$$

Сам вектор $\mathrm{d}\vec{r}$ напрямлений по дотичній до траєкторії, отже, вказує напрям руху в даній точці траєкторії у відповідний момент часу.

Швидкість. Стан руху точки визначається не просто зміною її положення в просторі, а тим, як воно відбувається в часі. Наближено рух точки за проміжок часу Δt характеризують відношенням здійсненого за цей час переміщення $\Delta\vec{r}$ до величини Δt, яке називають ***середнім вектором швидкості*** (або ***вектором середньої швидкості переміщення***) $\langle\vec{v}\rangle$:

$$\langle\vec{v}\rangle = \frac{\Delta\vec{r}}{\Delta t}. \qquad (1.2)$$

Напрям вектора $\langle\vec{v}\rangle$ співпадає з напрямом вектора переміщення (рис.1.1), і його модуль

$$|\langle\vec{v}\rangle| = \frac{|\Delta\vec{r}|}{\Delta t}. \qquad (1.2a)$$

Слід зазначити, що вектор $\langle\vec{v}\rangle = \Delta\vec{r}/\Delta t$ є визначеним тільки для заданого проміжку часу Δt , тож для різних проміжків може довільно змінюватись як за модулем, так і за напрямком. Але при поступовому зменшенні величини Δt відношення $\Delta\vec{r}/\Delta t$ прямує до визначеної границі $\vec{v}$, яка є точною характеристикою руху в кожну мить і називається ***миттєвою швидкістю*** (або просто ***швидкістю***):

$$\vec{v} = \lim_{\Delta t\to(0)} \frac{\Delta\vec{r}}{\Delta t} = \frac{\mathrm{d}\vec{r}}{\mathrm{d}t} \equiv \vec{r}\,'(t) \equiv \dot{\vec{r}}(t). \qquad (1.3)$$

(Похідну по часу можна записувати по-різному. У формулі (1.3) наведені всі можливі варіанти запису).

Оскільки переміщення $\Delta\vec{r}$ є приростом радіуса-вектора, то миттєва швидкість є похідною від радіуса-вектора по часу. Вектор $\vec{v}$ співнапрямлений з $\mathrm{d}\vec{r}$, тобто ***вектор миттєвої швидкості напрямлений по дотичній до траєкторії в кожній її точці***. Отже, вектор $\vec{v}$ визначає не тільки як швидко, а й у якому напрямі переміщується тіло в кожен момент часу. Тому говорять, що вектор швидкості визначає стан руху тіла. Зокрема, поведінка вектора швидкості дає загальну інформацію про характер руху. Наприклад, якщо $\vec{v} = \mathrm{const}$, тобто, ані величина, ані напрям швидкості не змінюються, то маємо рівномірний прямолінійний рух тіла. Якщо ж незмінною лишається тільки модуль вектора швидкості, то тіло здійснює рівномірний криволінійний рух, тощо.

У практичних задачах часто буває істотним не напрям руху, а лише швидкість подолання тілом шляху. Тому крім величин $\langle\vec{v}\rangle$ і $\vec{v}$ використовують середню $\langle v\rangle$ та миттєву *v скалярну*, або *шляхову* швидкості, які означають через пройдений шлях аналогічно до співвідношень (1.2) і (1.3.):

$$\langle v\rangle = \frac{S}{\Delta t}, \qquad (1.4)$$

і

$$v = \frac{dS}{dt}. \qquad (1.5)$$

Зауважимо, що із співвідношення (1.1а) випливає, що миттєва шляхова швидкість v дорівнює модулю вектора миттєвої швидкості, але для середніх швидкостей це, загалом, не так.

<u>Прискорення</u>. Ще однією характеристикою руху є прискорення. ***Вектор миттєвого прискорення*** (або просто ***прискорення***) $\vec{a}$ визначає швидкість зміни вектора швидкості у часі й уводиться аналогічно до миттєвої швидкості:

$$\vec{a} = \lim_{\Delta t\to 0} \frac{\Delta\vec{v}}{\Delta t} = \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} \equiv \vec{v}\,'(t) \equiv \dot{\vec{v}}\,(t) \equiv \frac{d^2\vec{r}}{dt^2} = \vec{r}\,''(t) = \ddot{\vec{r}}\,(t), \qquad (1.6)$$

тобто це похідна від вектора швидкості по часу, або ж друга похідна від радіуса-вектора по часу (У формулі (1.6) наведені всі варіанти запису другої похідної по часу). Вектор прискорення збігається за напрямом з вектором $\Delta\vec{v}$ і в загальному випадку складає певний кут із напрямом швидкості (рис.1.3). Якщо рух є рівнозмінним ($\vec{a} = \mathrm{const}$), то прискорення визначається виразом

$$\vec{a} = \frac{\Delta\vec{v}}{\Delta t} = \frac{\vec{v}_2 - \vec{v}_1}{\Delta t}, \qquad (1.7)$$

де $\Delta t = t_2 - t_1$ – проміжок часу, за який швидкість змінюється на величину $\Delta\vec{v}$.

При нерівномірній зміні швидкості цей вираз визначає середнє прискорення $\langle\vec{a}\rangle$ за час Δt.

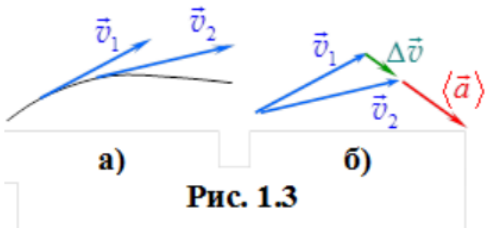


Рис. 1.3

Загальні рівняння кінематики точки. Коли задано закон руху точки, тобто – залежність радіуса-вектора від часу $\vec{r} = \vec{r}(t)$, то за допомогою співвідношень (1.2) – (1.7) легко знайти всі інші характеристики руху. Але на практиці найчастіше відоме прискорення $\vec{a} = \vec{a}(t)$, яке визначається законами динаміки, і завдання полягає у визначенні через нього решти кінематичних величин. Для цього існують загальні рівняння швидкості, переміщення (радіуса-вектора) та шляху. Вони встановлюються за допомогою методів інтегрального числення.

З формули (1.6) випливає, що за проміжок часу dt приріст вектора швидкості $\mathrm{d}\vec{v} = \vec{a}\mathrm{d}t$. Зміна швидкості за скінчений проміжок часу дорівнює сумі (точніше – інтегралу) всіх елементарних змін $\mathrm{d}\vec{v}$:

$$\Delta\vec{v} = \vec{v}_2 - \vec{v}_1 = \int_{t_1}^{t_2} \vec{a}\mathrm{d}t. \tag{1.8}$$

Якщо розглядати проміжок часу від початкового t = 0 до довільного моменту t, то з (1.8) отримаємо загальне рівняння швидкості:

$$\vec{v} = \vec{v}_0 + \int_0^t \vec{a}\mathrm{d}t, \tag{1.9}$$

де $\vec{v}_0 = \vec{v}(0)$ – початкова швидкість рухомої точки.

Аналогічними міркуваннями на основі співвідношення (1.3) встановлюються загальні рівняння для переміщення та радіуса-вектора:

$$\mathrm{d}\vec{r} = \vec{v}\mathrm{d}t,$$

$$\Delta\vec{r} = \vec{r}_2 - \vec{r}_1 = \int_{t_1}^{t_2} \vec{v}\mathrm{d}t, \tag{1.10}$$

$$\vec{r} = \vec{r}_0 + \int_0^t \vec{v}\mathrm{d}t, \tag{1.11}$$

де $\vec{r}_0 = \vec{r}(0)$ – початковий радіус-вектор, який визначає початкове положення рухомої точки.

Зауважимо, що для визначення швидкості та положення рухомої точки в будь-який момент часу замало знати лише прискорення $\vec{a} = \vec{a}(t)$. Треба також мати ***початкові умови***, тобто знати величини $\vec{r}_0$ та $\vec{v}_0$, котрі визначаються конкретною системою відліку, в якій розглядається рух.

Рівняння для пройденого точкою шляху встановлюється аналогічно через шляхову швидкість (модуль вектора швидкості) із співвідношення (1.5):

$$\mathrm{d}S = v\mathrm{d}t,$$

$$S = \int_0^t v\mathrm{d}t. \qquad (1.12)$$

Підставляючи в отримані загальні рівняння заданий закон зміни прискорення $\vec{a} = \vec{a}(t)$, можна отримати рівняння кінематики для будь-якого конкретного виду руху. Нехай, наприклад, маємо рух із відомим сталим прискоренням $\vec{a} = \text{const}$. Тоді, згідно із рівнянням (1.9),

$$\vec{v} = \vec{v}_0 + \vec{a}\int_0^t \mathrm{d}t \quad \Rightarrow \quad \vec{v} = \vec{v}_0 + \vec{a}t. \qquad (1.13)$$

Це відомий з елементарної фізики закон зміни швидкості при русі зі сталим вектором прискорення. Підставивши отриманий вираз $\vec{v}$ у рівняння (1.11), отримаємо відоме рівняння для радіуса-вектора при вказаному русі:

$$\vec{r} = \vec{r}_0 + \int_0^t (\vec{v}_0 + \vec{a}t)\,\mathrm{d}t \quad \Rightarrow \quad \Delta\vec{r} = \vec{r}_0 + \vec{v}_0 t + \frac{\vec{a}t^2}{2}. \qquad (1.14)$$

Аналогічно можна отримати рівняння кінематики й для складніших рухів, у яких прискорення змінюється з часом.

Координатний спосіб опису руху. Векторні співвідношення компактно і повно відображують фізичний зміст величин і зв'язки між ними, але в розрахунках вектори треба задавати числами. Тому, крім векторного, використовують координатний спосіб опису руху. В такому разі з тілом відліку жорстко зв'язують певну систему координат, найчастіше декартову, і положення точки в просторі визначають координатами – числами x, y, z. Ці числа дорівнюють відстаням (з тим, чи іншим знаком) від початку координат O до проекції точки на відповідну координатну вісь, рис. 1.4а. Відтак закон руху точки визначається рівняннями залежності координат від часу:

$$x = x(t), \quad y = y(t), \quad z = z(t). \qquad (1.15)$$

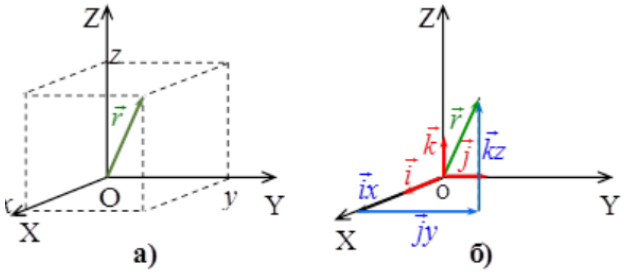


Рис. 1.4

Очевидно, що декартові координати точки є проекціями кінця радіуса-вектора на осі координат OX, OY, OZ: $x = r_x = r\cos\alpha$, $y = r_y = r\cos\beta$, $z = r_z = r\cos\gamma$, де α, β, γ – "напрямні кути", тобто кути між напрямками радіуса-вектора $\vec{r}$ та осей OX, OY, OZ (на рис. 1.4а показано тільки один з них). Тому можна записати:

$$\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}, \qquad (1.16)$$

де $\vec{i}$, $\vec{j}$, $\vec{k}$ – одиничні базисні вектори (*орти*), які визначають напрямки осей декартової системи координат (рис. 1.4б).

Отже, знаючи координати точки, можна обчислити модуль її радіуса-вектора та, через напрямні косинуси, його напрям:

$$r = \sqrt{x^2 + y^2 + z^2}; \quad \cos\alpha = x/r, \quad \cos\beta = y/r, \quad \cos\gamma = z/r, \tag{1.17}$$

Продиференціювавши вираз (1.16) по часу і врахувавши означення (1.3), отримаємо вираз швидкості точки в координатній формі:

$$\vec{v} = \vec{i}\frac{\mathrm{d}x}{\mathrm{d}t} + \vec{j}\frac{\mathrm{d}y}{\mathrm{d}t} + \vec{k}\frac{\mathrm{d}z}{\mathrm{d}t}. \tag{1.18}$$

Отже, похідні координат по часу – то є проекції вектора швидкості на відповідні осі:

$$v_x = \frac{\mathrm{d}x}{\mathrm{d}t}, v_y = \frac{\mathrm{d}y}{\mathrm{d}t}, v_z = \frac{\mathrm{d}z}{\mathrm{d}t}. \tag{1.19}$$

Тому за заданим законом руху в координатній формі (1.15) можна визначити проекції вектора швидкості (1.19), а також його модуль і напрям:

$$v = \sqrt{v_x^2 + v_y^2 + v_z^2}, \quad \cos\alpha = v_x/v, \quad \cos\beta = v_y/v, \quad \cos\gamma = v_z/v. \tag{1.20}$$

Аналогічно визначаються й параметри вектора прискорення при координатному способі опису руху:

$$\vec{a} = \vec{i}\frac{\mathrm{d}v_x}{\mathrm{d}t} + \vec{j}\frac{\mathrm{d}v_y}{\mathrm{d}t} + \vec{k}\frac{\mathrm{d}v_z}{\mathrm{d}t} = \vec{i}a_x + \vec{j}a_y + \vec{k}a_z. \tag{1.21}$$

$$a_x = \frac{\mathrm{d}v_x}{\mathrm{d}t} = \frac{\mathrm{d}^2x}{\mathrm{d}t^2}, a_y = \frac{\mathrm{d}v_x}{\mathrm{d}t} = \frac{\mathrm{d}^2y}{\mathrm{d}t^2}, a_z = \frac{\mathrm{d}v_z}{\mathrm{d}t} = \frac{\mathrm{d}^2z}{\mathrm{d}t^2}. \tag{1.22}$$

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2}, \cos\alpha = \frac{a_x}{a}, \cos\beta = \frac{a_y}{a}, \cos\gamma = \frac{a_z}{a}. \tag{1.23}$$

З усього сказаного в цьому пункті випливає наступний загальний порядок (алгоритм) розв'язування векторних рівнянь. Спочатку векторне рівняння "проектується" на координатні осі, тобто замість нього записується відповідна система алгебраїчних рівнянь для проекцій. Потім із них визначаються проекції шуканих векторів і, нарешті, якщо необхідно, через знайдені проекції визначаються модулі та напрямки шуканих векторів. У нескладних ситуаціях деякі етапи описаного алгоритму ми виконуємо автоматично, не виписуючи їх на папері. Як приклад, отримаємо з векторних рівнянь (1.9) і (1.11) рівняння для проекцій вектора швидкості та для координат точки через відомі проекції вектора прискорення:

$$v_x = v_{0x} + \int_0^t a_x \mathrm{d}t, \quad v_y = v_{oy} + \int_0^t a_y \mathrm{d}t, \quad v_z = v_{0z} + \int_0^t a_z \mathrm{d}t, \tag{1.24}$$

$$x = x_0 + \int_0^t v_x \mathrm{d}t, \quad y = y_0 + \int_0^t v_y \mathrm{d}t, \quad z = z_0 + \int_0^t v_z \mathrm{d}t. \tag{1.25}$$

Аналогічно з рівнянь (1.13) і (1.14) для *руху із сталим прискоренням* отримаємо:

$$v_x = v_{0x} + a_x t, \quad v_y = v_{0y} + a_y t, \quad v_z = v_{0z} + a_z t;$$
$$x = x_0 + v_{0x}t + \frac{a_x t^2}{2}, \quad y = y_0 + v_{0y}t + \frac{a_y t^2}{2}, \quad z = z_0 + v_{0z}t + \frac{a_z t^2}{2}. \tag{1.25a}$$

Із цих рівнянь можна також отримати корисні для розв'язування задач співвідношення:

$$v_x^2 - v_{0x}^2 = 2a_x(x - x_0), \quad v_y^2 - v_{oy}^2 = 2a_y(y - y_0), \quad v_z^2 - v_{0z}^2 = 2a_z(z - z_0). \quad (1.25б)$$

Природній спосіб описання руху. Цей спосіб є зручним при описі руху по заданій траєкторії. Положення точки на траєкторії задається криволінійною координатою, тобто відстанню l від обраного початку відліку O до цієї точки, відрахованою вздовж траєкторії. Позитивний напрям відліку криволінійної координати задається довільно, виходячи з міркувань зручності. Закон руху точки при цьому визначається залежністю криволінійної координати від часу $l = l(t)$. Зауважимо також, що модуль зміни криволінійної координати – то є пройдений точкою шлях.

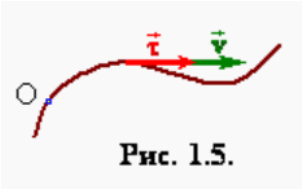


Рис. 1.5.

Для задання напряму руху і вектора швидкості точки з нею зв'язують одиничний вектор (орт) $\vec{\tau}$ дотичної до траєкторії, спрямований в бік збільшення координати рис. 1.5. В такому разі можна записати:

$$\vec{v} = v_\tau \vec{\tau}$$

де $v_\tau = \mathrm{d}l/\mathrm{d}t$ – проекція вектора швидкості на напрям $\vec{\tau}$**.** Швидкість v_τ є величиною алгебраїчною, її знак залежить від напрямку руху точки, а модуль дорівнює модулю вектора швидкості:

$$|v_\tau| = |\vec{v}| = v.$$

Вектор прискорення точки, котрий при такому способі опису називається ***повним прискоренням***, відповідно до означення (1.6), виражається як

$$\vec{a} = \frac{\mathrm{d}(v_\tau \vec{\tau})}{\mathrm{d}t} = \frac{\mathrm{d}v_\tau}{\mathrm{d}t}\vec{\tau} + v_\tau \frac{\mathrm{d}\vec{\tau}}{\mathrm{d}t}.$$

Таким чином, вектор повного прискорення має дві складові. Перша з них

$$\vec{a}_\tau = \frac{\mathrm{d}v_\tau}{\mathrm{d}t}\vec{\tau} \qquad (1.26)$$

напрямлена по дотичній до траєкторії (рис. 1.6в) і називається ***тангенціальним прискоренням***. Воно визначає ***зміну модуля вектора швидкості точки***. Тому для будь-якого рівномірного руху тангенціальне прискорення $\vec{a}_\tau = 0$.

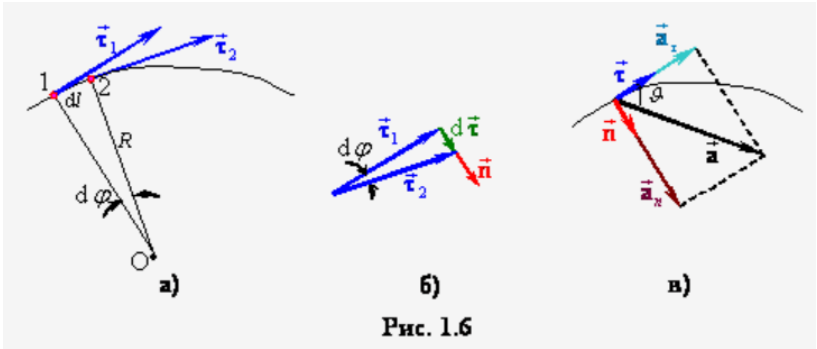


Рис. 1.6

З'ясуємо зміст другої складової вектора повного прискорення, котра називається *нормальним прискоренням*:

$$\vec{a}_n = v_\tau \frac{\mathrm{d}\vec{\tau}}{\mathrm{d}t}. \qquad (1.27)$$

Візьмемо до уваги те, що елементарна ділянка $\mathrm{d}l$ будь-якої кривої співпадає з дугою певного кола (*кола кривизни*) з відповідним центром (***центром кривизни*** O) та радіусом R - ***радіусом кривизни*** (рис. 1.6а). Урахувавши це, визначимо похідну $\mathrm{d}\vec{\tau}/\mathrm{d}t$.

Приріст орта $\vec{\tau}$ зумовлений його поворотом на нескінченно малий кут $\mathrm{d}\varphi$ при переміщенні точки по траєкторії на нескінченно малу відстань $\mathrm{d}l$ (рис. 1.6а,б) за гранично малий проміжок часу $\mathrm{d}t$. Очевидно, що при такому переміщенні кути повороту орта $\vec{\tau}$ і радіуса кривизни траєкторії R однакові, отже

$$\left|\frac{\mathrm{d}\tau}{\tau}\right| = \frac{\mathrm{d}l}{R} \quad \Rightarrow \quad |\mathrm{d}\tau| = \frac{\mathrm{d}l}{R}.$$

Оскільки кут $\mathrm{d}\varphi$ є нескінченно малим, то вектор $\mathrm{d}\vec{\tau}$ напрямлений перпендикулярно до вектора $\vec{\tau}$ (рис. 1.6б). Тому, якщо ввести орт (одиничний вектор) нормалі до траєкторії $\vec{n}$, то вектор $\mathrm{d}\vec{\tau}$ можна подати у вигляді

$$\mathrm{d}\vec{\tau} = |\mathrm{d}\vec{\tau}|\vec{n} = \frac{\mathrm{d}l}{R}\cdot\vec{n}.$$

Поділивши цей вираз на $\mathrm{d}t$, дістанемо:

$$\frac{\mathrm{d}\vec{\tau}}{\mathrm{d}t} = \frac{1}{R}\cdot\frac{\mathrm{d}l}{\mathrm{d}t}\vec{n} = \frac{v}{R}\vec{n}.$$

Отже, згідно з виразом (1.27), ***нормальне прискорення***

$$\vec{a}_n = \frac{v^2}{R}\vec{n}. \qquad (1.28)$$

Цей вектор напрямлений по нормалі до центра кривизни траєкторії у кожній її точці. **Він показує**, як швидко повертається орт $\vec{\tau}$, тобто, ***як швидко змінюється напрям руху точки***. Тому при прямолінійному русі $\vec{a}_n = 0$.

Повне прискорення $\vec{a}$ (рис.1. 6в) дорівнює сумі тангенціального та нормального прискорень:

$$\vec{a} = a_\tau\vec{\tau} + a_n\vec{n}. \qquad (1.29)$$

Його модуль

$$a = \sqrt{a_n^2 + a_\tau^2} = \sqrt{\left(\frac{v^2}{R}\right)^2 + \left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^2}. \qquad (1.29а)$$

Напрям вектора повного прискорення визначається кутом ϑ (рис. 1.6в), причому

$$\mathrm{tg}\vartheta = \frac{a_n}{a_\tau}. \qquad (1.29б)$$

Перетворення Галілея. Як уже говорилося, положення в просторі та рух точки є відносними, тобто вони є визначеними лише в обраній системі відліку. Відповідно, характеристики руху точки в двох різних системах відліку не однакові.

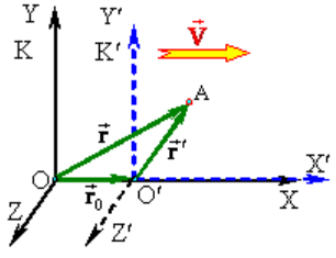


Рис. 1.7

Розглянемо дві системи відліку із збіжними осями X, X' й однаково напрямленими іншими осями координат – нерухому K і рухому K', що рухається зі швидкістю $\vec{V} = \mathrm{const}$ \ відносно K-системи у додатному напрямку осей X, X'. (рис. 1.7). Уважатимемо також, що

в початковий момент часу системи К і К’ збігалися. Тоді в довільний момент *t* положення точки О′ відносно О визначається радіусом-вектором $\vec{r}_0 = \vec{V}t$. Положення довільної точки А в К- і К′-системах визначаються радіусами-векторами $\vec{r}$ та $\vec{r}'$, які пов’язані співвідношеннями

$$\vec{r}' = \vec{r} - \vec{r}_0 = \vec{r} - \vec{V}t, \qquad (1.30)$$

або

$$\vec{r} = \vec{r}' + \vec{r}_0 = \vec{r}' + \vec{V}t. \qquad (1.30а)$$

Ці співвідношення виражають ***перетворення Галілея*** у векторній формі. Вони дозволяють визначати положення точки в одній системі відліку, якщо відоме її положення в іншій. При цьому вважається самоочевидним, що час є абсолютним, тобто тривалість будь-яких процесів, приміром руху тіл, не залежить від системи відліку:

$$t = t'. \qquad (1.31)$$

Перетворення Галілея в координатній формі , згідно з (1.16), мають вигляд:

$$x' = x - Vt,\ y' = y,\ z' = z; \qquad (1.32)$$

$$x = x' + Vt,\ y = y',\ z = z'. \qquad (1.32а)$$

Слід одразу зауважити, що ці співвідношення ґрунтуються на принципово хибних уявленнях про простір та час і є непридатними при швидкостях руху, сумірних із швидкістю поширення світла $c = 3 \cdot 10^8$ м/с (див. Лекція 1.10).

Взявши першу та другу похідну по часу від виразів (1.30) і (1.30а), знайдемо формули перетворення швидкостей і прискорень:

$$\vec{v}' = \vec{v} - \vec{V}, \qquad (1.33)$$

$$\vec{v} = \vec{v}' + \vec{V}, \qquad (1.33а)$$

$$\vec{a}' = \vec{a}, \quad \vec{a} = \vec{a}'. \qquad (1.34)$$

Зауважимо, що формули (1.33), (1.33а), які називають ще законом додавання швидкостей, стосуються будь-яких систем відліку, але співвідношення (1.34), є чинними лише для описаних систем відліку, що рухаються одна відносно одної прямолінійно, рівномірно і поступально.

1.2. Кінематика твердого тіла

Тіла, що рухаються, не завжди можна розглядати як матеріальні точки. Це стосується, наприклад, рухомих деталей та вузлів механізмів і машин. У цій лекції розглянуто наступні питання:

1. Рухи твердого тіла. Кінематика поступального руху
2. Обертальний рух твердого тіла
3. Зв'язок між лінійними та кутовими величинами
4. Загальні рівняння кінематики обертального руху
5. Плоский рух твердого тіла. Миттєва вісь

Рухи твердого тіла. Кінематика поступального руху

Різні точки твердого тіла рухаються не однаково (наприклад, розгляньте рух точок на осі та на ободі колеса, що котиться), тому механіка твердого тіла набагато складніша за механіку точки. Рух твердих тіл можна поділити на такі різновиди: 1) поступальний рух; 2) обертання навколо нерухомої осі; 3) плоский рух; 4) обертання навколо нерухомої точки; 5) вільний рух. Основними при цьому є поступальний і обертальний рухи, оскільки, як виявляється, інші різновиди руху твердого тіла можна розглядати як сукупність цих двох.

Поступальним називають рух, при якому довільна пряма, проведена між двома точками тіла, залишається паралельною до свого початкового напрямку. Прикладом може бути рух кузова автомобіля, або зісковзування бруска по похилій площині без зміни орієнтації.

Обертальним називають такий рух твердого тіла, коли в будь-яку мить усі його точки рухаються по колах із центрами на одній прямій – осі обертання, – що проходить через тіло.

При поступальному русі всі точки тіла рухаються по однакових за формою траєкторіях і в кожен момент часу мають однакову швидкість і прискорення. Отже, поступальний рух тіла визначається рухом будь-якої однієї його точки. Тому розглянута кінематика точки одночасно є й кінематикою поступального руху твердого тіла. Натомість кінематика обертального руху потребує окремого розгляду .

Обертальний рух твердого тіла

Нехай маємо якесь тіло, що обертається навколо нерухомої осі OZ (рис. 2.1). Розглянемо деяку його точку, котра рухається по коловій траєкторії з центром у точці C і радіусом *R*. Задамо положення точки радіусом-вектором $\vec{r}$ із початком в точці O на осі обертання. За час d*t* точка здійснює переміщення $\mathrm{d}\vec{r}$, яке є перпендикулярним до $\vec{r}$ і має модуль

$$|\mathrm{d}\vec{r}| = R\mathrm{d}\varphi = r\mathrm{d}\varphi \sin\vartheta. \qquad (2.1)$$

де $\mathrm{d}\varphi$ – кут повороту тіла навколо осі обертання за час d*t*.

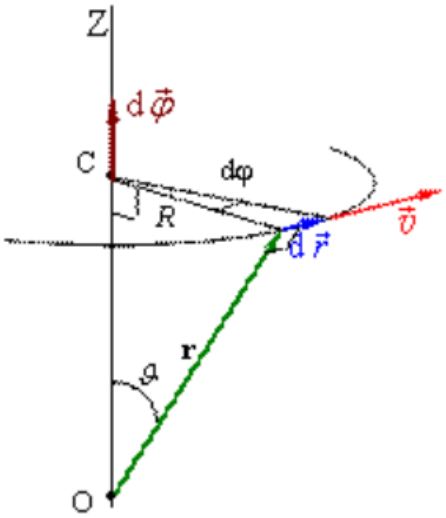


Рис. 2.1

Це співвідношення можна подати у векторній формі так, що воно буде відображати й напрям обертання тіла. Для цього величину $\mathrm{d}\varphi$ розглядають як модуль вектора елементарного кута повороту $\mathrm{d}\vec{\varphi}$, який напрямлений уздовж осі обертання згідно з правилом правого гвинта.

За цим правилом вектор $\mathrm{d}\vec{\varphi}$ *спрямований у напрямку вкручування правого гвинта при його обертанні в напрямку обертання тіла; подібні вектори називаються аксіальними. Зауважимо також, що зображувати векторами можна лише нескінченно малі повороти.*

У такому разі замість виразу (2.1) можна записати:

$$\mathrm{d}\vec{r} = [\mathrm{d}\vec{\varphi}, \vec{r}]\,. \qquad (2.1a)$$

Швидкість руху точки отримаємо, поділивши $\mathrm{d}\vec{r}$ на $\mathrm{d}t$:

$$\vec{v} = \left[\frac{\mathrm{d}\vec{\varphi}}{\mathrm{d}t}, \vec{r}\right]. \qquad (2.2)$$

З цього виразу видно, що різні точки обертового тіла рухаються з різними швидкостями, але перший множник під знаком векторного добутку, однаковий для всіх точок, і тому визначає рух не лишень окремої точки, а й усього тіла. Вектор

$$\vec{\omega} = \frac{\mathrm{d}\vec{\varphi}}{\mathrm{d}t} \qquad (2.3)$$

називається *кутовою швидкістю* тіла і є кількісною характеристикою обертального руху .

Вектор $\vec{\omega}$, як і $\mathrm{d}\vec{\varphi}$, напрямлений уздовж осі обертання згідно з правилом правого гвинта. Одиницею кутової швидкості є 1 рад/с (радіан за секунду).

Зміна вектора кутової швидкості з часом характеризується вектором кутового прискорення $\vec{\beta}$:

$$\vec{\beta} = \frac{\mathrm{d}\vec{\omega}}{\mathrm{d}t}. \qquad (2.4)$$

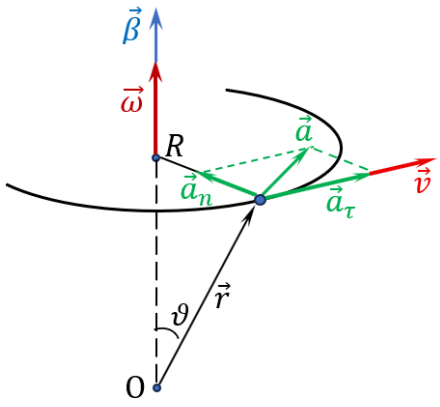


Рис. 2.2

Одиницею кутового прискорення є 1 рад/с^2.

При обертанні навколо фіксованої осі вектор $\vec{\beta}$, так само як і вектор $\vec{\omega}$, напрямлений уздовж осі обертання (рис 2.2). У такому разі зручніше використовувати не вектори, а їхні проекції на вісь обертання OZ, напрям якої пов'язаний із позитивним напрямом відліку кута повороту правилом правого гвинта:

$$\omega_z = \frac{\mathrm{d}\varphi}{\mathrm{d}t}, \quad \beta_z = \frac{\mathrm{d}\omega_z}{\mathrm{d}t}. \tag{2.5}$$

При цьому знак ω_z визначає напрям обертання, а знак β_z – характер обертання (рис. 2.2 відображає прискорене обертання в додатному напрямі осі OZ).

Якщо під час руху вісь обертання змінює напрям, то $\mathrm{d}\vec{\omega}$ і $\vec{\beta}$ напрямлені під кутом до осі).

Окремим важливим для практики випадком обертального руху є рівномірне обертання тіла навколо фіксованої осі ($\vec{\beta} = 0, \quad \vec{\omega} = \mathrm{const}$). Такий рух є періодичним, отож окрім кутової швидкості його характеризують періодом T – проміжком часу, за який здійснюється один оберт, – і частотою обертання n – кількістю обертів за одиницю часу.

Оскільки за один оберт тіло повертається на кут 2π, то

$$T = \frac{2\pi}{\omega} \text{ і } n = \frac{1}{T} = \frac{\omega}{2\pi}. \tag{2.6}$$

Зв'язок між лінійними та кутовими величинами

При розгляді обертального руху окремих точок твердого тіла величини $\mathrm{d}\vec{r}$ і $\vec{v}$ відповідно називають *лінійним* переміщенням і *лінійною* швидкістю, на відміну від кутового переміщення $\mathrm{d}\vec{\phi}$ та кутової швидкості $\vec{\omega}$. Між лінійними та кутовими величинами існують однозначні зв'язки. Зокрема, зв'язок між елементарними лінійним і кутовим переміщеннями задається виразами (2.1) і (2.1а), а зв'язок між лінійною та кутовою швидкістю – виразом (2.2) із урахуванням означення (2.3):

$$\vec{v} = [\vec{\omega}, \vec{r}]\,. \tag{2.7}$$

Для модулів маємо:

$$v = \omega r \sin\vartheta = \omega R, \tag{2.7а}$$

де $R = r\sin\vartheta$ – радіус кола, по якому рухається точка (рис. 2.2).

Вираз для повного прискорення точки через кутові величини знайдемо диференціюванням виразу (2.7):

$$\vec{a} = \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \left[\frac{\mathrm{d}\vec{\omega}}{\mathrm{d}t}, \vec{r}\right] + \left[\vec{\omega}, \frac{\mathrm{d}\vec{r}}{dt}\right] = \left[\vec{\beta}, \vec{r}\right] + [\vec{\omega}, \vec{\mathbf{v}}]\,,$$

де враховано, що $\mathrm{d}\vec{\omega}/\mathrm{d}t = \vec{\beta}$ – вектор кутового прискорення, а $\mathrm{d}\vec{r}/\mathrm{d}t = \vec{v}$ – вектор лінійної швидкості.

Оскільки при обертанні тіла навколо нерухомої осі вектори $\vec{\omega}$ і $\vec{\beta}$ лежать на осі, то вектор $\left[\vec{\beta}, \vec{r}\right]$ напрямлений по дотичній до траєкторії даної точки тіла (рис. 2.2) і є її тангенціальним прискоренням $\vec{a}_\tau$:

$$\vec{a}_\tau = \left[\vec{\beta}, \vec{r}\right]. \tag{2.8}$$

При цьому тангенціальне прискорення даної точки тіла, що обертається, називається *лінійним* прискоренням цієї точки. Його проекція на напрям дотичної до кола

$$a_\tau = \beta_z r \sin\vartheta = \beta_z R. \tag{2.8a}$$

Так само друга складова повного прискорення $[\vec{\omega}, \vec{v}]$ при нерухомій осі обертання напрямлена по нормалі до траєкторії точки (рис.2.2), то є її нормальним прискоренням:

$$\vec{a}_n = [\vec{\omega}, \vec{v}]. \tag{2.9}$$

Модуль нормального прискорення

$$a_n = \omega v = \omega^2 R. \tag{2.9a}$$

На основі співвідношень (2.8а) і (2.9а) можна визначити модуль і напрям (див. рис. 2.2) повного прискорення точок обертового тіла:

$$a = \sqrt{a_\tau^2 + a_n^2} = R\sqrt{\beta^2 + \omega^4}, \tag{2.10}$$

$$\mathrm{tg}\alpha = \frac{\omega^2}{\beta}. \tag{2.10a}$$

Загальні рівняння кінематики обертального руху

Кутове прискорення тіла, як і прискорення окремої матеріальної точки, визначається силовою дією на обертове тіло з боку інших тіл, отож його можна знайти, аналізуючи фізичні умови, в яких здійснюється обертання. Тому основне завдання кінематики обертального руху тіла полягає у визначенні інших кутових величин через відоме кутове прискорення. При обертанні навколо *нерухомої* осі це завдання розв'язується так само, як і в кінематиці точки (див. Лекцію 1). При цьому слід завважити, що за загальним змістом і формальними означеннями кутові величини – переміщення $\mathrm{d}\vec{\varphi}$, швидкість $\vec{\omega}$ і прискорення $\vec{\beta}$ – є аналогами відповідних лінійних величин, які характеризують рух матеріальної точки. Тому й зв'язки між кутовими величинами такі самі, як і між лінійними. Через це всі основні рівняння кінематики обертального руху навколо фіксованої осі аналогічні відповідним рівнянням кінематики руху точки у фіксованому напрямку (прямолінійного руху). Зокрема, проекція кутової швидкості визначається загальним рівнянням, аналогічним рівнянню v_z із (1.24):

$$\omega_z = \omega_{0z} + \int_0^t \beta_z \mathrm{d}t, \tag{2.11}$$

де ω_{0z} – проекція початкової кутової швидкості, β_z – проекція кутового прискорення на напрям осі обертання.

У простому випадку рівнозмінного обертання β_z = const, отже,

$$\omega_z = \omega_{0z} + \beta_z \int_0^t \mathrm{d}t \quad \Rightarrow \quad \omega_z = \omega_{0z} + \beta_z t, \tag{2.11а}$$

що аналогічно до (1.25а).

Кут повороту (кутове переміщення) φ, який визначає зміну положення тіла відносно осі обертання, знаходиться із загального рівняння

$$\varphi = \int_0^t \omega_z \mathrm{d}t., \tag{2.12}$$

Це рівняння теж є аналогом рівнянь (1.25) кінематики точки, що визначають зміну положення точки відносно вибраного початку відліку.

При рівнозмінному обертанні $\beta_z = \mathrm{const}$, $\omega_z = \omega_{0z} + \beta_z t$, отже,

$$\varphi = \int_0^t (\omega_{0z} + \beta_z t)\, \mathrm{d}t \quad \Rightarrow \quad \varphi = \omega_{0z} t + \frac{\beta_z t^2}{2}. \tag{2.12а}$$

Слід зауважити, що в рівняннях (2.11) – (2.12) величина φ є алгебраїчною, тож число $\varphi/2\pi$ не визначає кількості обертів N (повний "шлях"), зроблених тілом за час t. (Виняток становить тільки обертання тіла в незмінному напрямі, коли φ не змінює знаку протягом заданого часу руху). В загальному випадку кількість обертів тіла визначається через модуль кутової швидкості рівнянням

$$N = \frac{1}{2\pi} \int_0^t \omega \mathrm{d}t, \tag{2.12б}$$

яке є аналогом рівняння шляху (1.12) в кінематиці матеріальної точки.

Плоский рух твердого тіла. Миттєва вісь

Плоским рухом називається такий рух, при якому всі точки тіла переміщуються в площинах, паралельних до певної нерухомої в обраній системі відліку площини. Будемо для зручності умовно називати такі площини "площинами руху" точок тіла. Прикладом плоского руху може бути кочення циліндра: всі його точки рухаються в перпендикулярних до осі площинах.

<u>Швидкість точки тіла при плоскому русі</u>. Нехай якесь тіло здійснює плоский рух. Прослідкуємо за відрізком АВ, який з'єднує дві точки цього тіла, що знаходяться в площині руху. За деякий проміжок часу відрізок із положення A_1B_1, переміщується в положення A_2B_2 (рис. 2.3). Цю зміну положення можна розглядати як результат поступального переміщення в положення A_2B' і повороту в площині руху на деякий кут φ навколо точки А (рис. 2.3а). Але так само можна говорити про поступальне переміщення відрізка в положення $A'B_2$ та поворот навколо точки В (рис. 2.3б). При цьому переміщення точок А і В – A_1A_2 і B_1B_2 – не однакові, але кут повороту φ один і той самий. Зрозуміло, що сказане вірно й для будь якої іншої пари точок і для будь-якого проміжку часу, зокрема, й для нескінченно малого. Тому плоский рух твердого тіла можна розглядати як сукупність поступального руху та обертання навколо фіксованої осі перпендикулярної до площин руху точок тіла. При цьому

кутова швидкість обертання тіла не залежить від вибору такої осі.

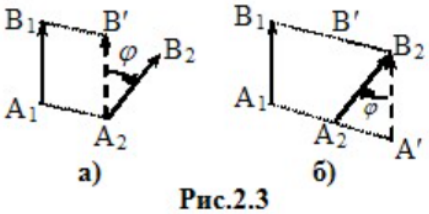


Рис.2.3

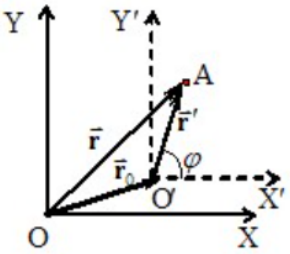


Рис. 2.4

Взявши до уваги сказане, розглянемо рух довільної точки А тіла, що здійснює плоский рух в системі відліку XOY (К-система) так, що точки тіла рухаються в площинах, паралельних XOY (рис. 2.4). Пов'яжемо з тілом рухому систему відліку X′O′Y′ (К′-система), положення початку відліку котрої O′ в К-системі визначається радіусом-вектором $\vec{r}_0$ Положення точки А відносно К-системи відліку визначається радіусом-вектором $\vec{r}$, а відносно К′-системи – радіусом-вектором $\vec{r}'$. Очевидно, що

$$\vec{r} = \vec{r}_0 + \vec{r}'.$$

Переміщення точки А за нескінченно малий проміжок часу dt

$$\mathrm{d}\vec{r} = \mathrm{d}\vec{r}_0 + \mathrm{d}\vec{r}'.$$

Переміщення $\vec{r}'$ зумовлене поворотом тіла навколо осі, що проходить через точки O′, тому $\mathrm{d}\vec{r}' = \left[\mathrm{d}\vec{\varphi}, \vec{r}'\right]$ (див. (2.1а)). Отже

$$\mathrm{d}\vec{r} = \mathrm{d}\vec{r}_0 + \left[\mathrm{d}\vec{\varphi}, \vec{r}'\right].$$

Поділивши останній вираз на проміжок часу dt, одержимо швидкість точки А в К – системі відліку:

$$\vec{v} = \vec{v}_0 + \left[\vec{\omega}, \vec{r}'\right]. \qquad (2.13)$$

Таким чином, при плоскому русі швидкість довільної точки А твердого тіла складається із швидкості $\vec{v}_0$ будь-якої іншої точки O′, що жорстко зв'язана з ним (пака точка може розміщуватись і поза тілом), і лінійної швидкості $\vec{v} = [\vec{\omega}, \vec{r}]$ обертального руху точки А навколо осі, що проходить через точку O′ перпендикулярно до площини руху.

Миттєва вісь. Оскільки вибір точки O′ є довільним, плоский рух тіла можна звести до чисто обертального. Справді, при плоскому русі вектори $\vec{v}_0$ і $\vec{v}'$ перпендикулярні до вектора кутової швидкості $\vec{\omega}$, отже обидва лежать в одній площині руху. Тому в кожну мить існує така жорстко зв'язана з тілом точка М, миттєва швидкість якої $\vec{v}'$ в К – системі рівна нулю. ЇЇ радіус-вектор $\vec{r}_m{}'$ визначається із співвідношення (2.13):

$$\vec{v}_0' + \left[\vec{\omega}, \vec{r}_m'\right] = 0 \quad \Rightarrow \quad \left[\vec{\omega}, \vec{r}_m'\right] = -\vec{v}_0' \qquad (2.14)$$

Зокрема, модуль $\vec{r}_m{}'$, тобто відстань між точками М і O′, дорівнює:

$$r_m' = \frac{v_0}{\omega} \qquad (2.14а)$$

(При потребі детальнішої інформації про положення миттєвої осі вираз (2.14) слід розписати в координатній формі за правилами розкриття векторноо добутку)

Оскільки точка М у дану мить є нерухомою, то рух тіла в цей момент можна трактувати як чисте обертання навколо осі, що проходить через цю точку перпендикулярно до площини руху. Таку вісь називають ***миттєвою віссю***. В загальному випадку положення миттєвої осі може змінюватися з часом. Наприклад, при коченні циліндра без ковзання по плоскій поверхні миттєва вісь збігається з лінією дотику циліндра до поверхні й рухається із швидкістю осі циліндра. Поняття миттєвої осі є досить продуктивним, оскільки в багатьох випадках спрощує аналіз плоского руху.

1.3. Контрольні запитання

1. Що таке система відліку?
2. Як треба розуміти відносність руху?
3. Які кінематичні величини використовують для опису руху матеріальної точки?
4. Які існують способи опису руху?
5. Який зв'язок існує між радіусом-вектором точки та її декартовими координатами? Між вектором швидкості та його проекціями на декартові осі координат?
6. Що таке нормальне, тангенціальне та повне прискорення? Який між ними існує зв'язок?
7. Що можна сказати про тангенціальне та нормальне прискорення точки, якщо вона здійснює: прямолінійний нерівномірний рух? криволінійний рівномірний рух?
8. Як змінюється кут між вектором повного прискорення та вектором швидкості у випадку прискореного руху точки по колу?
9. Як записуються формули (1.33) і (1.33*a*) в координатній формі для систем відліку, описаних в п. 1.4?
10. Які види рухів може здійснювати тверде тіло? Які з них можна вважати основними?
11. Які величини використовують для опису обертального руху?
12. Як напрямлений вектор кутової швидкості твердого тіла?
13. Як напрямлений вектор кутового прискорення? В якому випадку вектори кутової швидкості та кутового прискорення є колінеарними? Співнапрямленими?
14. Який зв'язок існує між векторами лінійної та кутової швидкості? Між їхніми модулями?
15. Як виражаються тангенціальне, нормальне та повне прискорення точки обертового тіла через кутові характеристики руху?
16. Вивести співвідношення (2.11) і (2.12).
17. Який рух твердого тіла називається плоским? Наведіть декілька прикладів.
18. Що таке миттєва вісь? Яку перевагу вона дає при розгляді плоского руху тіла?

2. ОСНОВНІ ЗАКОНИ ДИНАМІКИ



Динаміка – це розділ механіки, в якому закони руху тіл встановлюються через причини, що зумовлюють його характер. Основу динаміки складають три закони Ньютона, котрі є узагальненням результатів спостережень і спеціально поставлених експериментів, і які не можна вивести з якихось більш простих принципів. Закони Ньютона мають і велике практичне значення, на них ґрунтуються розрахунки різноманітних машин і механізмів, будівельних конструкцій, космічних апаратів, тощо. Щоправда, закони Ньютона навіть у рамках класичної механіки не є універсальними і виконуються тільки в так званих інерціальних системах відліку.

У цій лекції розглянуті наступні питання:

2.1. Інерціальні системи відліку. Перший закон Ньютона
2.2. Другий закон Ньютона. Сила
2.3. Третій закон Ньютона
2.4. Основне рівняння динаміки
2.5. Закони сил
2.6. Контрольні запитання

2.1. Інерціальні системи відліку. Перший закон Ньютона

Як відзначалося в кінематиці, для опису руху тіл можна користуватися будь-якими системами відліку. При цьому в довільній системі відліку характер руху тіла (зміна швидкості з часом, форма траєкторії, тощо) визначається не лише дією інших тіл, а й властивостями системи відліку. Тож рух даного тіла в різних системах відліку може бути складнішим чи простішим (наприклад, видимий рух планет на небосхилі, тобто - відносно Землі, відбувається по дуже складних траєкторіях, в той час як відносно Сонця вони рухаються по простих еліптичних траєкторіях). Але в деяких системах відліку характер руху тіла при заданих початкових умовах визначається тільки його взаємодією з іншими тілами. Такі системи називаються ***інерціальними системами відліку.***

Особливості цього класу систем відліку відображує *перший закон Ньютона*, за яким

в інерціальній системі відліку тіло за відсутності взаємодії з іншими тілами зберігає стан спокою або рівномірного прямолінійного руху .

Таке тіло називається вільним, і його рух теж називається вільним рухом, або рухом *за інерцією*. Тому перший закон Ньютона також називають **законом інерції**.

Як бачимо, в інерціальній системі відліку рух вільного тіла є найпростішим. При цьому виявляється, що в таких системах відліку й інші закони фізики набувають найбільш простого вигляду. Тому для фізики інерціальні системи відліку є переважними. Але тут виникає принципова складність. Інерціальність системи відліку можна встановити тільки за характером руху в ній вільного тіла, але абсолютно вільних тіл у природі, строго говорячи, не існує. Тому реально ми можемо мати справу тільки з тілами, дія на котрі з боку інших тіл є або *компенсованою*, або неістотною. Отже, теоретично довести інерціальність тієї, чи іншої системи відліку принципово неможливо. Це є питанням досліду, і саме дослід свідчить, що такі системи відліку існують (у зв'язку з цим деякі автори трактують перший закон Ньютона як твердження про існування інерціальних систем відліку). Численні спостереження показують, що з гранично великою точністю інерціальними можна вважати системи відліку, пов'язані із віддаленими "нерухомими" зірками та з Сонцем, а також такі, що рухаються відносно зірок поступально, рівномірно та прямолінійно, тобто без будь-яких прискорень. Що ж до систем відліку, пов'язаних із Землею, які найчастіше використовують на практиці, то їх можна вважати інерціальними лише наближено через наявність у них невеликих прискорень зумовлених обертанням Землі навколо власної осі та орбітальним рухом навколо Сонця.

2.2. Другий закон Ньютона. Сила

Сила. Дослідні факти вказують на те, що в інерціальній системі відліку зміна стану руху тіла зумовлюється виключно дією на нього з боку іншого тіла чи тіл. Кількісною мірою такої дії є ***сила*** – одна з основних величин динаміки. Сила має такі загальні властивості.

– Зовнішньою ознакою сили є зміна швидкості тіла (ще одною ознакою сили є деформація тіл, але при вивченні законів руху вона є несуттєвою).

– Дія одного тіла на інше завжди має певний напрям, тому сила є векторною величиною.

– Якщо матеріальна точка взаємодіє не з одним, а з декількома іншими тілами, то, як показує дослід, діюча на неї сила визначається як

$$\vec{F} = \vec{F}_1 + \vec{F}_2 + \ldots = \sum_i \vec{F}_i, \qquad (3.1)$$

де $\vec{F}_i$ – сила, яка би діяла на точку з боку i-того тіла за відсутності інших (рис. 3.1).

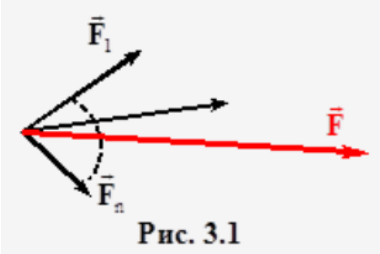


Рис. 3.1

У цьому сенсі говорять, що сили $\vec{F}_i$ підпорядковані принципу суперпозиції. Силу \(\vec{F}\) називають ***рівнодійною*** силою, а $\vec{F}_i$ – ***складовими*** силами. Поняття рівнодійної є зручним, оскільки дозволяє лаконічно записувати й формулювати відповідні закони та співвідношення механіки. З іншого боку, задану силу можна розкладати на будь-які складові так, аби виконувалося співвідношення (3.1). Але слід зауважити, що для тіл, які не можна вважати матеріальними точками, поняття рівнодійної існує тільки тоді, коли лінії дії всіх прикладених до тіла окремих сил перетинаються в одній точці.

– У механіці всі сили поділяють на два класи: сили, що виникають при безпосередньому контакті тіл (сила тертя, сила опору, сила тиску, тощо), та сили, що діють між тілами на відстані (наприклад, сила всесвітнього тяжіння, сила взаємодії між електричними зарядам та струмами). В останньому випадку кажуть, що в просторі існує силове поле (гравітаційне, електромагнітне), в якому на дане тіло діє певна сила (зазначений поділ є умовним, оскільки на молекулярному рівні поняття дотику втрачає прямий зміст, бо взаємодія між молекулами здійснюється не “безпосередньо”, а через створюване ними електромагнітне поле).

– Сили різної фізичної природи мають різні властивості, але у загальному випадку сила взаємодії між двома тілами визначається їх взаємним розташуванням і швидкістю їх руху одне відносно одного: $\vec{F} = \vec{F}(\vec{r}, \vec{v})$.

Маса та імпульс. Другий закон Ньютона

Маса та імпульс. Досліди показують, що будь-яке тіло “чинить опір” при спробах зміни його швидкості, через що швидкість тіла неможливо змінити миттєво навіть на дуже малу величину. Ця властивість називається *інертністю тіла*. Мірою інертності тіла є його ***маса***. Чим більша інертність, тим більша маса тіла, тож під дією однакової сили швидкість тіла з більшою масою змінюється повільніше, ніж у тіла з меншою масою.

Масу тіл визначають, у той чи інший спосіб порівнюючи її з масою еталона. Одиницею маси в міжнародній системі одиниць (СІ) є 1 кілограм (кг) – маса еталонного тіла, яке зберігається в міжнародному метрологічному центрі (м. Севр поблизу Парижа).

Дослід свідчить, що маса має дві важливі властивості:

1). Маса є скалярною й, у межах ньютонівської механіки, *адитивною* величиною, тобто маса тіла дорівнює сумі мас його складових частин.

2). Інертність є внутрішньою властивістю тіла, тому маса тіла не залежить від його руху, отже й від системи відліку, в якій розглядається рух.

Через наявність інертності механічний стан тіла визначається не лише швидкістю, а й масою. Наприклад, ми легко зупинимо рукою кульку для пінг-понгу, але цього не можна сказати про вагон. Тому в динаміці мірою стану руху тіла є ***імпульс*** – добуток маси тіла на вектор його швидкості:

$$\vec{p} = m\vec{v}, \tag{3.2}$$

отже ***імпульс є вектором*** напрямленим, як і швидкість, по дотичній до траєкторії руху тіла. Одиницею імпульсу є 1 кг·м/с.

Імпульс теж є адитивною величиною – імпульс системи дорівнює сумі імпульсів усіх тіл, які входять до її складу:

$$\vec{P} = \sum_i \vec{p}_i = \sum_i m_i\vec{v}_i. \tag{3.2а}$$

Другий закон Ньютона. Аналізуючи результати дослідів і спостережень за рухом тіл, Ньютон встановив, що *швидкість зміни імпульсу тіла (тут і далі мається на увазі матеріальна точка) повністю визначається силою, що діє на нього*, згідно з рівнянням:

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = \vec{F}, \tag{3.3}$$

де $\mathrm{d}\vec{p}$ - зміна імпульсу за нескінченно малий проміжок часу $\mathrm{d}t$ і $\vec{F}$ - рівнодійна сила, прикладена до тіла (зауважимо, що математично ліва частина цього виразу являє собою похідну імпульсу по часу).

Це твердження становить ***другий закон Ньютона***. Його також записують у вигляді:

$$\mathrm{d}\vec{p} = \vec{F}\mathrm{d}t. \tag{3.4}$$

Величину $\vec{F}\mathrm{d}t$ називають ***імпульсом сили*** за час $\mathrm{d}t$. Отже другий закон Ньютона можна сформулювати й так:

приріст імпульсу тіла за певний проміжок часу дорівнює імпульсу сили , що діє на тіло протягом цього часу.

При необхідності визначати зміну імпульсу за скінчений проміжок часу t_1 - t_2 вираз (3.4) треба інтегрувати на проміжку від t_1 до t_2:

$$\Delta\vec{p} = \vec{p}_2 - \vec{p}_1 = \int_{t_1}^{t_2} \vec{F}\mathrm{d}t. \tag{3.4а}$$

Якщо маса тіла не змінюється (m = const), то $\mathrm{d}\vec{p} = \mathrm{d}(m\vec{v}) = m\mathrm{d}\vec{v}$ і тоді $\mathrm{d}\vec{p}/\mathrm{d}t = m\vec{a}$, де $\vec{a} = \mathrm{d}\vec{v}/\mathrm{d}t$ – прискорення. Отже, для тіла незмінної маси другий закон Ньютона можна записати у вигляді:

$$m\vec{a} = \vec{F}, \qquad (3.5)$$

тобто,

добуток маси тіла на його прискорення дорівнює рівнодійній силі, що прикладена до тіла.

На основі цього рівняння встановлена одиниця сили - ньютон (Н): 1 Н - то є сила, котра тілу маси 1 кг надає прискорення 1 м/с^2.

2..3. Третій закон Ньютона

Важливим дослідним фактом є те, що дія одного тіла на інше завжди має взаємний характер, тобто сили є проявом ***взаємодії*** між тілами. Третій закон Ньютона виражає загальну властивість сил взаємодії між тілами:

> ***сили, з якими два тіла діють одне на одне, завжди рівні по модулю і напрямлені у протилежні боки вздовж прямої, що з'єднує тіла*** **(рис.3.2).**



Рис. 3.2

Отже

$$\vec{F}_1 = -\vec{F}_2. \qquad (3.6)$$

Щодо наведеного формулювання треба зауважити таке.

– Сили взаємодії завжди мають одну й ту саму фізичну природу та походження.

– Будь-яка сила, що діє на тіло, не залежить від інших прикладених до нього сил. Тому третій закон Ньютона виконується й у довільній системі (сукупності тіл) для кожної пари тіл.

2.4. Основне рівняння динаміки

Основне рівняння динаміки

Закони Ньютона є системою органічно взаємопов'язаних тверджень, і в будь-якій задачі динаміки "працюють" усі три закони. Але основним є другий закон Ньютона, оскільки він безпосередньо оперує характеристиками руху – імпульсом або прискоренням (рівняння (3.3), (3.5)).

Враховуючи, що прискорення $\vec{a} = \dfrac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2}$, другий закон Ньютона (3.5) можна подати у вигляді диференціального рівняння

$$\frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = \frac{\vec{F}}{m}, \qquad (3.7)$$

яке є ***основним рівнянням динаміки матеріальної точки***. Це рівняння також називають *основним рівнянням руху* матеріальної точки, оскільки воно дозволяє розв'язати основну задачу динаміки – визначити закон руху точки, тобто залежність її радіуса-вектора від часу $\vec{r} = \vec{r}(t)$. Для цього треба проінтегрувати рівняння (3.7) при відомій залежності сили від часу $\vec{F} = \vec{F}\,(t)$ і заданих початкових умовах – початкових швидкості $\vec{v}_0 = \vec{v}(0)$ і радіусі-векторі $\vec{r}_0 = \vec{r}(0)$. (Розв'язування диференціального рівняння називається його інтегруванням, оскільки зводиться до однойменної математичної процедури. На загал для цього треба знати теорію диференціальних рівнянь, але в простих задачах достатньо й початкових відомостей з інтегрального числення)

Рівняння (3.7) дозволяє розв'язувати й обернені задачі. А саме, визначати силу, що діє на точку, та її швидкість і прискорення, якщо відомий закон руху. Такі задачі є досить простими і зводяться до диференціювання по часу заданої функції $\vec{r}(t)$.

2.5. Закони сил

Для визначення закону руху точки необхідно мати повну інформацію про діючі на неї сили, а для цього треба знати властивості сил і їхні математичні вирази (закони сил).

У макроскопічному світі спостерігається багато різних сил, але всі вони є проявами лише двох фундаментальних взаємодій – гравітаційної та електромагнітної (фундаментальними називають такі взаємодії, які не зводяться ні до яких інших. Крім вказаних відомі ще два види фундаментальних взаємодій – сильна та слабка, – але в механічних явищах вони ніяк себе не виявляють). При цьому гравітація виявляє себе лишень у силі тяжіння, ***усі інші сили мають електромагнітну природу*** і зумовлені взаємодією між зарядженими частинками, з яких складаються молекули та атоми речовини. В механічних процесах найчастіше виявляють себе гравітаційна сила та сила тяжіння, пружна сила та вага тіла, сили тертя та опору.

Гравітаційна сила та сила тяжіння

Гравітаційна сила та сила тяжіння. Гравітаційна взаємодія є притаманною будь-яким матеріальним об'єктам і виявляється в тому, що всі тіла притягаються одне до одного. Згідно із законом всесвітнього тяжіння Ньютона ***гравітаційна сила*** притягання між двома **матеріальними точками** є прямо пропорційною їхнім масам m_1, m_2 і обернено пропорційною квадратові відстані r між ними:

$$F = G\frac{m_1 m_2}{r^2}, \qquad (3.8)$$

де величина G = 6,67·10^{-11} м^3/(кг·с^2), яка називається гравітаційною сталою, показує "силу" гравітаційної взаємодії. А саме, G чисельно дорівнює силі притягання двох тіл масою по 1 кг на відстані 1 м одне від одного. Отже, гравітаційні сили є гранично слабкими і відіграють істотну роль лише тоді, коли хоча б одне з тіл має астрономічну масу (планети та зірки).

Гравітаційна сила діє вздовж прямої, що з'єднує тіла — партнери взаємодії.

Варто звернути увагу на те, що величини m_1 і m_2 в законі всесвітнього тяжіння визначають не інертність тіл, а їхню здатність до взаємного притягання, тобто вони є ***гравітаційними*** масами, на відміну від ***інертних*** мас тіл, що фігурують у другому законі Ньютона. Але в дослідах встановлено, що гравітаційна та інертна маси є строго прямо пропорційні одна одній. Тому гравітаційну та інертну маси не розрізняють і говорять просто про масу тіла та вимірюють в одних і тих самих одиницях. Отже, фізична величина "маса" є і мірою інертності тіл, і мірою їхньої здатності до гравітаційної взаємодії.

Використовуючи формулу (3.8), слід пам'ятати, що вона є придатною тільки для матеріальних точок, тобто для тіл, які розташовані одне від одного на дуже великій відстані порівняно з їхніми розмірами. Виняток складає притягання двох однорідних куль (або кулі та матеріальної точки). В цьому випадку r – це відстань між центрами куль (або центром кулі та матеріальною точкою). Зокрема, це стосується ***сили тяжіння***, тобто сили гравітаційного притягання, що діє на тіла поблизу поверхні планети. Згідно з формулою (3.8) вона визначається як

$$F = G\frac{mM}{R^2}, \qquad (3.8а)$$

де m - маса тіла; M, R - маса і радіус планети.

Гравітаційна сила є прямо пропорційною масі тіла, на яке діє. Тому біля поверхні планети вона надає всім тілам ***однакового прискорення сили тяжіння*** або прискорення вільного падіння (це не зовсім точно, оскільки на прискорення вільного падіння тіла відносно Землі дещо впливає її добове обертання та орбітальний рух)

$$g = G\frac{M}{R^2}, \tag{3.9}$$

яке напрямлене вертикально вниз. Для Землі наближено $g = 9,8$ м/с^2. З урахуванням (3.9) силу тяжіння записують у вигляді

$$\vec{F} = m\vec{g}. \tag{3.10}$$

Пружна сила

Пружна сила (сила пружності) виникає при пружних деформаціях тіл, наприклад, при розтягу пружини, або еластичного шнура чи стержня (пружними називають деформації, котрі зникають після припинення дії сили, що їх спричинює). Вона зумовлена дією електромагнітних сил взаємодії між молекулами деформованого тіла. Пружна сила завжди напрямлена протилежно до напрямку деформації й за модулем є прямо пропорційною величині деформації (закон Гука). Для деформації розтягу чи стискання ***величина*** (***модуль***) пружної сили

$$F = k\Delta l, \tag{3.11}$$

де $\Delta l = |l - l_0|$ – величина деформації, l, l_0 – довжина деформованої та недеформованої пружини, відповідно, і k - коефіцієнт пропорційності, який називають жорсткістю тіла.

Для стержня (шнура) або пружини довжиною l із сталим перерізом s

$$k = E\frac{s}{l}, \tag{3.12}$$

де E – так званий модуль Юнга, що є табличною характеристикою пружних властивостей речовини тіла. У такому разі вираз закону Гука можна подати у вигляді:

$$\sigma = E\epsilon \quad \text{або} \quad \epsilon = \frac{\sigma}{E}, \tag{3.13}$$

де $\epsilon = \Delta l/l_0$ – відносна деформація і $\sigma = F/s$ – механічна напруга, що виникає в тілі внаслідок деформації.

Вага і невагомість

Якщо тіло лежить на опорі (чи підвішене на шнурі), то через його притягання до Землі виникає деформація як опори (підвісу), так і самого тіла. Як наслідок, з'являються сили взаємодії між тілом і опорою (підвісом).

> ***Сила, що діє з боку тіла на горизонтальну опору або вертикальний підвіс унаслідок притягання до Землі називається вагою тіла*** $\vec{P}$ **(рис. 3.3)**.

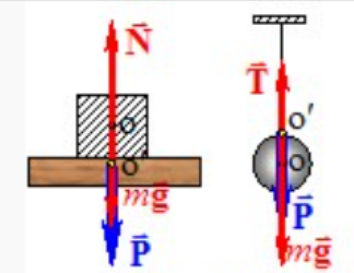


Рис.3.3

Відповідно, сила, що діє на тіло з боку опори (підвісу) називається реакцією опори $\vec{N}$ або підвісу $\vec{T}$. Отже вага тіла прикладена не до нього, а до опори чи підвісу. Якщо опора (підвіс), отож і тіло, не має прискорення відносно інерціальної системи відліку, то $m\vec{g} + \vec{N} = 0$ і, оскільки $\vec{N} = -\vec{P}$,

$$\vec{P} = m\vec{g}, \tag{3.14}$$

тобто, вага тіла дорівнює силі тяжіння, що діє на нього (з цієї причини силу тяжіння інколи теж називають вагою). Але, коли опора (підвіс) має прискорення $\vec{a}$, то згідно з другим і третім законами Ньютона $m\vec{a} = m\vec{g} + \vec{N} = m\vec{g} - \vec{P}$, і

$$\vec{P} = m(\vec{g} - \vec{a}). \tag{3.14a}$$

Отже, вага залежить від прискорення опори і може бути або більшою (при русі вгору), або меншою (при русі вниз), ніж mg. Зокрема, коли опора (підвіс) рухається вниз із прискоренням $\vec{a} = \vec{g}$, то $\vec{P} = 0$, і тіло перебуває в **невагомості**. При цьому $\vec{N} = 0$, тобто в ***стані невагомості на тіло діє тільки сила тяжіння.*** Із залежністю ваги від прискорення опори ми повсякденно стикаємося в ліфті, але особливо відчутно це для космонавтів: вони зазнають великих перевантажень при виході ракети на орбіту та перебувають у невагомості протягом усього часу орбітального польоту.

Сили тертя та опору

З дослідів відомо, що при ковзанні даного тіла по поверхні іншого, або ж при спробах такого руху виникає сила, що перешкоджає рухові. Цю силу називають ***силою сухого тертя***. Якщо при спробі примусити одне тіло ковзати по поверхні іншого закріпленого тіла перше залишається у спокої, то тертя, що виникає, називають ***тертям спокою***. Якщо ж тіла рухаються одне відносно одного, то залежно від характеру руху говорять про ***тертя ковзання***, або ***тертя кочення***. Тертя між частинами одного й того ж тіла при їх відносному русі, наприклад, тертя між сусідніми шарами рідини чи газу, що рухаються із різними швидкостями, називають ***внутрішнім тертям***.

Властивості сили тертя визначаються умовами, за яких вона виникає. Зокрема, ***сила тертя спокою*** $\vec{F}_c$ дорівнює сумі векторів решти сил (рівнодійній), що діють на дане тіло, взятій зі знаком мінус:

$$\vec{F}_c = -\sum \vec{F}_i \tag{3.15}$$

(У цьому легко впевнитися на основі другого закону Ньютона. Якщо тіло знаходиться у спокої, то його прискорення $\vec{a} = 0$,отже і сума сил, що діють на тіло $\Sigma(\vec{F}_i + \vec{F}_c) = 0$, звідси і випливає вираз (3.15)).

Окрім того, в дослідах встановлено, що величина сили тертя спокою не може перевищувати певного максимального значення

$$F_{c\ \max} = \mu N, \qquad (3.16)$$

де μ – коефіцієнт тертя, який залежить від природи і стану дотикових поверхонь (зокрема їхньої шерхлості), N – сила нормального тиску, що притискає дане тіло до поверхні іншого.

Якщо рівнодійна решти сил, прикладених до тіла, перевищує значення $F_{c\ \max}$, то це тіло починає рухатися по поверхні іншого, і тертя спокою змінюється на тертя ковзання. Величина ***сили тертя ковзання*** $\boldsymbol{F_т}$ досить складно залежить від відносної швидкості тіл (рис. 3.4а), але при достатньо малих швидкостях її можна вважати сталою (рис.3.4б) і рівною максимальній силі тертя спокою

$$F_т = \mu N, \qquad (3.16а)$$

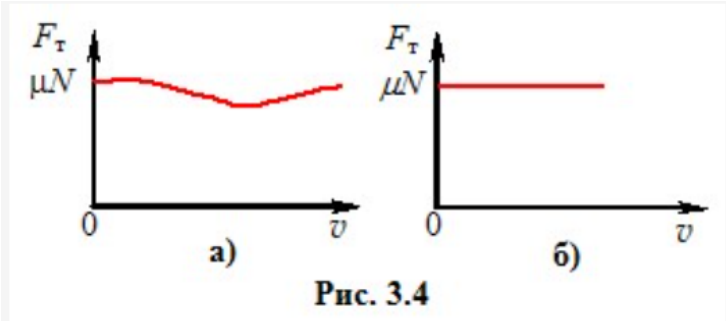


Рис. 3.4

Сила опору $\vec{F}_{оп}$ виникає при русі тіла в газоподібному чи рідкому середовищі й зумовлена взаємодією його молекул із молекулами поверхні тіла та тертям між шарами самого середовища (в'язкістю). При цьому через велику рухливість молекул газу та рідини при русі в середовищі немає тертя спокою.

Сила опору спрямована проти руху тіла й істотно залежить від його швидкості відносно середовища. При малих швидкостях можна вважати, що

$$\vec{F}_{оп} = \alpha v \vec{e}_v \qquad (3.17)$$

де α – стала величина (коефіцієнт опору), залежна від характеристик тіла і середовища, $\vec{e}_v$ – орт вектора швидкості. При великих швидкостях ця залежність є близькою до квадратичної, а при надзвукових — до кубічної

Характерною властивістю сил опору є їхня велика залежність від форми рухомого тіла. Тому рухомим апаратам (авто, літаки, ракети, тощо) надають специфічної обтічної форми.

2.6. Контрольні запитання

1. Чим вивчення механічного руху в динаміці відрізняється від такого в кінематиці?
2. Говорять, що маса є мірою інертності тіла. А що таке інертність?
3. Який фізичний зміст має поняття “сила”? Які загальні властивості вона має?
4. З рівняння (3.5) випливає, що при $\vec{F} = 0$ прискорення $\vec{a} = 0$, і $\vec{v} = \text{const}$, тобто, тіло рухається рівномірно й прямолінійно. Чи означає це, що перший закон Ньютона є наслідком другого?
5. Яка різниця існує між поняттями “прискорення сили тяжіння” й “прискорення вільного падіння”?
6. Що таке вага тіла? Від чого і як вона залежить? Чи можна сказати, що вага є властивістю тіла?
7. Чи можна сказати, що невагомість - то є стан тіла, коли на нього не діє сила тяжіння?
8. Чи діє сила тертя на тіло, що нерухомо лежить на горизонтальній площині? Похилій площині?
9. Що є спільного та відмінного між силами сухого тертя та силами опору в середовищі?

3. НЕІНЕРЦІАЛЬНІ СИСТЕМИ ВІДЛІКУ

Закони Ньютона виконуються тільки в інерціальних систем відліку. Але на практиці існує багато випадків, коли рух тіла доводиться розглядати відносно ***неінерціальної*** системи відліку, тобто системи відліку, яка ***рухається з прискоренням*** відносно інерціальної. Тому постає питання, чи не можна якось пристосувати основне рівняння динаміки й для неінерціальних систем відліку. Виявляється, що можна. Розглянемо, як саме:

3.1. Сили інерції

3.2. Сили інерції в обертовій системі відліку

3.3. Загальне рівняння руху точки в НСВ

3.4. Контрольні запитання

3.1. Сили інерції

Згадаймо добре відому кожному ситуацію. Знаходячись в автобусі, що рухається без прискорення, ми перебуваємо відносно автобуса в спокої і відчуваємо лише силу тяжіння. Але під час різкого гальмування або прискорення, чи при крутих поворотах якась інша сила відкидає нас уздовж або поперек руху. Ця сила виникає не тому, що на нас починають діяти якісь інші тіла, а тому, що автобус, який для пасажира є природньою системою відліку, набуває прискорення відносно Землі, тобто стає неінерціальною системою відліку.

Сили, що зумовлені не взаємодією між тілами, а властивостями системи відліку, називаються силами інерції.

Зауваження. У силу того, що сили інерції визначються не взаємодією між тілами, автори деяких підручників називають їх фіктивними.

Урахування сил інерції дозволяє поширити основне рівняння динаміки (3.5) або (3.7) і на неінерціальні системи відліку.

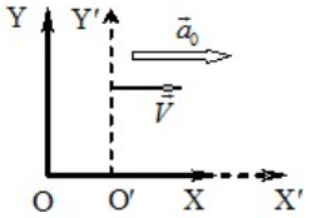


Рис. 4.1

Розглянемо сили інерції, що виникають при поступальному русі системи відліку. Нехай є дві системи відліку: нерухома інерціальна К-система й К′-система, яка рухається поступально з прискоренням $\vec{a}_0$ відносно К-системи вздовж її осі ОХ (рис.4.1). Якщо тіло рухається відносно К-системи зі швидкістю $\vec{v}$, то його швидкість відносно К′-системи, згідно з (1.33),

$$\vec{v}' = \vec{v} - \vec{V},$$

де $\vec{V}$ - швидкість К′ -системи відносно К в дану мить. Звідси, відповідно до (1.6),

$$\vec{a}' = \vec{a} - \overrightarrow{a_0}, \tag{4.1}$$

де $\vec{a}'$ – прискорення тіла відносно неінерціальної системи відліку К′, $\vec{a}$ – його прискорення відносно інерціальної системи відліку К, $\vec{a}_0$ – прискорення К′- системи відліку відносно К. Домноживши вираз (4.1) на масу тіла m, одержимо

$$m\vec{a}' = m\vec{a} - m\vec{a}_0.$$

Величина $m\vec{a} = \vec{F}$ – то є “звичайна” сила, що діє на тіло в інерціальній системі відліку з боку інших тіл. Але прискорення тіла в К′-системі відліку визначається не тільки цією силою, а ще й величиною

$$\vec{F}_i = -m\vec{a}_0, \tag{4.2}$$

яка називається (поступальною) ***силою інерції***. Ця сила зумовлена прискореним рухом К′-системи відліку, отже, вона не пов’язана із взаємодією даного тіла з іншими тілами. Тому для сили інерції третій закон Ньютона не виконується. Іншою характерною особливістю сили інерції є те, що вона визначається добутком маси тіла не на його прискорення, а на прискорення системи відліку.

Уведення сили інерції дозволяє поширити основне рівняння динаміки й на поступальні неінерціальні системи відліку:

$$m\vec{a}' = \vec{F} + \vec{F}_i, \tag{4.3}$$

де $\vec{F}$ - рівнодійна сил, що діють на тіло з боку інших тіл і полів, $\vec{F}_i$ - сила інерції.

3.2. Сили інерції в обертовій системі відліку

Розглянемо тепер неінерціальну систему відліку K′, яка обертається зі сталою кутовою швидкістю ω навколо осі нерухомої відносно інерціальної системи відліку K. Кожна точка такої обертової системи відліку має відповідне доцентрове прискорення відносно осі обертання. Тому прискорення тіла в інерціальній системі відліку K залежить не тільки від зміни швидкості, а й від його положення в системі K′.

Відцентрова сила. Спочатку розглянемо тіло нерухоме відносно обертової системи відліку.

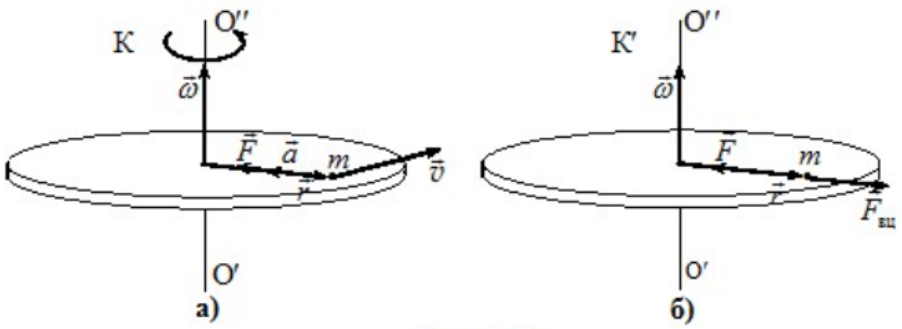


Рис. 4.2

Пов'яжемо систему відліку K′ із диском, який обертається з кутовою швидкістю ω навколо осі O′ O″, що є нерухомою в інерціальній системі відліку K (рис 4.2). Нехай на поверхні диска на відстані r від його осі обертання лежить тіло, що прикріплене до осі ниткою. Тоді відносно системи відліку K тіло рухається по колу радіуса r із нормальним прискоренням $\vec{a} = -\omega^2\vec{r}$, яке створюється силою натягу нитки $\vec{F}$ (рис. 4.2*а*). За другим законом Ньютона $\vec{F} = -m\omega^2\vec{r}$. Але відносно диска тіло є нерухомим, отже в системі відліку K′ його прискорення $\vec{a}' = 0$. Це означає, що в обертовій системі відліку на тіло, крім сили взаємодії $\vec{F}$, діє ще й рівна їй за модулем і протилежна за напрямом сила інерції

$$\vec{F}_{вц} = m\omega^2\vec{r}. \qquad (4.4)$$

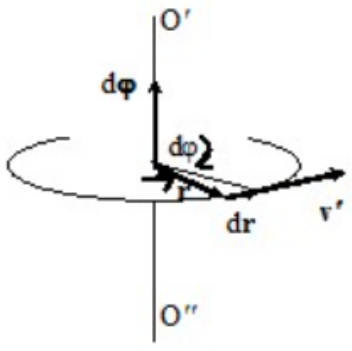


Рис. 4.3

Ця сила напрямлена від осі (центра) обертання системи відліку K′ (рис. 4.2б) і називається ***відцентровою*** силою. Саме відцентрову силу відчуває пасажир на поворотах.

Сила Коріоліса. Нехай тепер точка рухається відносно обертової системи відліку K′ із швидкістю $\vec{v}'$ (рис. 4.3). Тоді в інерціальній K-системі відліку за час d*t* радіус-вектор $\vec{r}$, який задає положення точки відносно осі обертання, набуде приросту

$$\mathrm{d}\vec{r} = \vec{v}'\,\mathrm{d}t + [\mathrm{d}\vec{\varphi}, \vec{r}]\,. \tag{4.5}$$

У цьому виразі перший доданок зумовлений рухом точки відносно К′-системи відліку, а другий – то є переміщення точки разом із К′-системою при її повороті на кут $\mathrm{d}\vec{\varphi}$ відносно К-системи відліку (див. (2.1а)). Поділивши (4.5) на dt, одержимо:

$$\vec{v} = \vec{v}' + [\vec{\omega}, \vec{r}]\,. \tag{4.5a}$$

Суттєво, що $\vec{v}$ залежить від $\vec{r}$. Тому, навіть коли точка в обертовій системі відліку К′ рухається рівномірно ($\vec{v}' = \mathrm{const}$), її швидкість відносно нерухомої К-системи відліку буде змінюватися внаслідок зміни відстані r до осі О′О″. Тому приріст вектора швидкості $\mathrm{d}\vec{v}$ у К-системі визначається як зміною вектора відносної швидкості $\mathrm{d}\vec{v}'$, так і зміною радіуса-вектора $\mathrm{d}\vec{r}$:

$$\mathrm{d}\vec{v} = \mathrm{d}\vec{v}' + [\vec{\omega}, \mathrm{d}\vec{r}]\,. \tag{4.6}$$

Підставивши в це співвідношення вираз (4.5), отримаємо:

$$\mathrm{d}\vec{v} = \mathrm{d}\vec{v}' + \left[\vec{\omega}, \vec{v}'\right]\mathrm{d}t + [\vec{\omega}\,[\mathrm{d}\vec{\varphi}, \vec{r}]] \tag{4.7}$$

Розглянемо тепер зміну вектора $\vec{v}$ відносно системи відліку К. Навіть якщо швидкість тіла $\vec{v}'$ в обертовій системі відліку К′ є сталою, внаслідок повороту системи К′ на деякий кут $\mathrm{d}\vec{\varphi}$ за час dt, вектор $\vec{v}'$ відносно К-системи відліку теж повернеться на кут $\mathrm{d}\vec{\varphi}$ і отримає приріст $\mathrm{d}\vec{v}' = \vec{a}'\mathrm{d}t + \left[\mathrm{d}\vec{\varphi}, \vec{v}'\right]$. Якщо ж тіло має відносно К′- системи відліку ще й прискорення $\vec{a}'$, то вектор $\vec{v}'$ набуває додаткового приросту $\vec{a}'\mathrm{d}t$. Тому в загальному випадку

$$\mathrm{d}\vec{v} = \vec{a}'\,\mathrm{d}t + \left[\mathrm{d}\vec{\varphi}, \vec{v}'\right]\,.$$

Підставивши цей вираз у (4.7) і поділивши на dt, отримаємо:

$$\vec{a} = \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \vec{a}' + 2\left[\vec{\omega}, \vec{v}'\right] + [\vec{\omega}\,[\vec{\omega}, \vec{r}]]\,.$$

Легко переконатися, що $[\vec{\omega}\,[\vec{\omega}, \vec{r}]] = -\omega^2\vec{r}$. Враховуючи також, що $\left[\vec{v}', \vec{\omega}\right] = -\left[\vec{\omega}, \vec{v}'\right]$, маємо:

$$\vec{a}' = \vec{a} + 2\left[\vec{v}', \vec{\omega}\right] - [\vec{\omega}\,[\vec{\omega}, \vec{r}]]\,. \tag{4.8}$$

Домноживши цей вираз на масу m точки, одержимо рівняння динаміки матеріальної точки в обертовій системі відліку:

$$m\vec{a}' = m\vec{a} + 2m\left[\vec{v}', \omega\right] + m\omega^2\vec{r}. \tag{4.9}$$

або

$$m\vec{a}' = \vec{F} + \vec{F}_{\text{к}} + \vec{F}_{\text{вц}}, \tag{4.9a}$$

де $\vec{F} = m\vec{a}$ – сила взаємодії точки з іншими тілами, $\vec{F}_{\text{вц}} = m\omega^2\vec{r}$ – вже знайома відцентрова сила інерції. Але, як бачимо, на рухому точку в обертовій системі відліку діє ще одна сила інерції

$$\vec{F}_{\text{К}} = 2m[\vec{v}', \vec{\omega}], \tag{4.10}$$

яка називається ***силою Коріоліса*** (або коріолісовою силою).

Характерною властивістю сили Коріоліса є те, що вона завжди діє перпендикулярно до напрямку руху тіла. Зокрема, при русі в меридіональному напрямку в північній півкулі вона напрямлена праворуч по ходу руху тіла, а у південній – ліворуч (рис. 4.4). З цим пов'язаний відомий ефект підмивання правих берегів річок у північній півкулі і лівих у південній. Також завдяки коріолісовій силі спостерігається обертання площини коливань математичного маятника. Саме це дозволило у свій час (1852 р.) французькому фізикові Фуко довести, що Земля обертається навколо своєї осі. Існують й інші прояви сил Коріоліса. Зокрема, при русі тіла вздовж паралелі сила Коріоліса має вертикальну складову, напрямлену вгору при русі на схід і вниз при русі на захід, що впливає на дальність польоту тіл і враховується в балістиці.

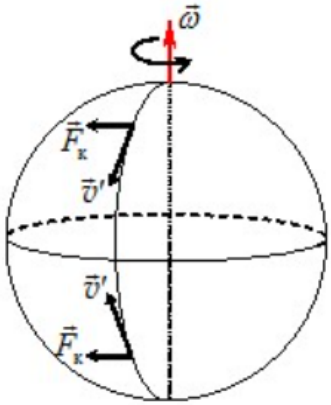


Рис. 4.4

3.3. Загальне рівняння руху точки в НСВ

У кінематиці (Лекція 2, п. 2.1) говорилося про те, що довільний рух твердого тіла можна трактувати як сукупність поступального та обертального рухів. Тому в довільній неінерціальній системі відліку на тіла, крім сил взаємодії, діють усі розглянуті види сил інерції. Відповідно, основні рівняння руху (3.5) і (3.7) у довільній НСВ мають вигляд:

$$m\vec{a}' = \vec{F} - m\vec{a}_0 + m\omega^2\vec{r} + 2m\left[\vec{v}', \vec{\omega}\right]; \tag{4.11}$$

$$\frac{\mathrm{d}\vec{r}'}{\mathrm{d}t^2} = \frac{\vec{F}}{m} - \vec{a}_0 + \omega\vec{r} + 2\left[\vec{v}', \vec{\omega}\right], \tag{4.12}$$

де $\vec{a}_0$ і $\vec{\omega}$ - поступальне прискорення та кутова швидкість системи відліку. Два інші доданки являють собою ***відцентрове прискорення***

$$\vec{a}_{\text{вц}} = \omega^2\vec{r} \tag{4.13}$$

та ***коріолісове прискорення***

$$\vec{a}_{\text{к}} = 2\left[\vec{v}', \vec{\omega}\right].$$

3.4. Контрольні запитання

1. Чи виконуються закони Ньютона в довільній системі відліку?

2. На горизонтальній платформі, що рівномірно рухається по прямій, нерухомо лежить кулька. Але коли платформа починає рухатися прискорено, кулька теж починає прискорено котитися по платформі. Як це пояснюється в системі відліку, пов'язаній із землею? із платформою?

3. Чим відрізняється основне рівняння руху тіла в неінерціальній системі відліку від рівняння руху в інерціальній системі відліку?

4. Чи скрізь напрям нитки виска збігається з точним напрямом до центра Землі?

5. Чи є різниця між поняттями "прискорення вільного падіння" та "прискорення сили тяжіння"?

6. В якому випадку коріолісова сила інерції, що діє на тіло на поверхні Землі, збігається за напрямом із відцентровою силою?

7. Чи можливо, щоби коріолісова сила, котра діє на тіло, що рухається на південь, теж була напрямлена на південь?

8. На екваторі роблять постріл із гармати один раз точно на північ, а другий - на південь. Що можна сказати про дальності польоту снарядів в обох випадках?

9. З однакової висоти роблять горизонтальний постріл із гармати точно на схід: один раз на екваторі, а інший - десь у північній півкулі. Чим будуть відрізнятися траєкторії снаряда?

4. ДИНАМІКА СИСТЕМ

На практиці найчастіше доводиться мати справу не з окремою частинкою (матеріальною точкою), а із ***системою***, тобто з виділеною сукупністю частинок або тіл, які певним чином взаємодіють між собою та з іншими тілами, що не включені до складу системи. Останні звуться *зовнішніми тілами*[1]. Підкреслимо, що вибір тіл, які включаються до системи і які розглядаються як зовнішні, не є наперед заданим, він визначається поставленими задачами. Ураховуючи які тіла включені в систему, сили, що діють на них, поділяють на ***внутрішні*** та ***зовнішні***. Якщо на тіла системи не діють зовнішні сили, то вона називається ***ізольованою***, або ***замкненою***. Отже, в ізольованих системах діють лише внутрішні сили.

Рух тіл системи формально визначається рівняннями другого закону Ньютона (3.3), в які входять усі зовнішні та внутрішні сили. При цьому внутрішні сили не є заданими, позаяк взаємне розташування та відносні швидкості тіл системи весь час змінюються. З цієї причини навіть для системи всього з трьох тіл рівняння руху (3.7) не мають точних розв'язків, не говорячи про складні системи, коли детальна інформація про сили взаємодії між тілами взагалі є невідомою. Тому при дослідженні систем особливого значення набувають деякі величини, поведінка котрих не залежить від складу системи та внутрішніх взаємодій у ній. До таких, зокрема, відноситься імпульс системи.

Далі розглянуті наступні питання:

4.1. Імпульс системи

4.2. Закон збереження імпульсу

4.3. Рух тіл змінної маси (рівняння Мещерського)

4.4. Центр мас

4.5. Контрольні запитання

4..1. Імпульс системи

Нагадаємо, що імпульсом системи називається адитивна величина, котра дорівнює сумі імпульсів усіх тіл системи:

$$\vec{P} = \sum_i \vec{p}_i = \sum_i m\vec{v}_i.$$

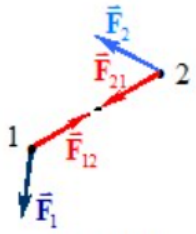


Рис. 5.1

Розглянемо поведінку імпульсу в найпростішій системі, що складається всього з двох матеріальних точок (частинок) 1 і 2, які взаємодіють із силами $\vec{F}_{12}$ та $\vec{F}_{21}$ і на які діють зовнішні сили $\vec{F}_1$ і $\vec{F}_2$ (рис. 5.1). Рух кожної частинки визначається рівнянням (3.3), отже зміна імпульсу системи виражається, як

$$\frac{\mathrm{d}\vec{P}}{\mathrm{d}t} = \sum_i \frac{\mathrm{d}\vec{p}_i}{\mathrm{d}t} = \frac{\mathrm{d}\vec{p}_1}{\mathrm{d}t} + \frac{\mathrm{d}\vec{p}_2}{\mathrm{d}t} = \vec{F}_1 + \vec{F}_{12} + \vec{F}_2 + \vec{F}_{21}.$$

Але, згідно з третім законом Ньютона (3.6), $\vec{F}_{12} + \vec{F}_{21} = 0$, тому

$$\frac{\mathrm{d}\vec{P}}{\mathrm{d}t} = \vec{F}_1 + \vec{F}_2 = \sum_i \vec{F}_i.$$

У системі, що складається з багатьох частинок, співвідношення (3.6) виконується для будь-якої пари частинок. Тому отриманий результат зберігає чинність, і можна записати:

$$\frac{\mathrm{d}\vec{P}}{\mathrm{d}t} = \vec{F}, \qquad (5.1)$$

де величину $\vec{F} = \sum_i \vec{F}_i$ назвемо ***сумарною зовнішньою силою***, що діє на тіла системи. Зауважимо, що величину $\vec{F}$ не можна розглядати як “рівнодійну” за винятком ситуації, коли лінії дії всіх зовнішніх сил перетинаються в одній точці. Таким чином

> ***швидкість зміни імпульсу довільної системи дорівнює сумарній зовнішній силі, що діє на систему.***

Це твердження й рівняння (5.1) інколи називають законом зміни імпульсу системи. Воно виражає той дуже важливий факт, що, на відміну від окремих частинок, імпульс усієї системи здатні змінювати лише зовнішні сили. Сили інерції, що діють на тіла в неінерціальних системах відліку, відносяться до зовнішніх сил.

Зміна імпульсу системи за скінчений проміжок часу $[t_1,\ t_2]$ визначається повним імпульсом зовнішніх сил за цей проміжок:

$$\Delta\vec{P} = \int_{t_1}^{t_2} \vec{F}\mathrm{d}t. \qquad (5.1a)$$

4.2. Закон збереження імпульсу

Якщо система замкнена (зовнішні сили $\vec{F}_i$ відсутні), то в рівнянні (5.1) $\vec{F} = 0$, отже

$$\frac{\mathrm{d}\vec{P}}{\mathrm{d}t} = 0 \quad \Rightarrow \quad \vec{P} = \mathrm{const}. \tag{5.2}$$

У цьому полягає **закон збереження імпульсу**:

> ***імпульс замкненої системи зберігається, тобто залишається незмінним у часі.***

При цьому імпульси окремих частинок системи можуть змінюватися внаслідок взаємодії між собою, але лише так, що імпульс системи лишається незмінним. Іншими словами, взаємодія між тілами призводить тільки до обміну імпульсами без зміни сумарного імпульсу системи. Тому рівняння (5.2) розгорнуто можна записати так:

$$m_1\vec{v}_1' + m_2\vec{v}_2' + \ldots = m_1\vec{v}_1'' + m_2\vec{v}_2'' + \ldots \tag{5.2a}$$

де ліва та права частини відносяться до двох довільних моментів часу.

З наведених міркувань і викладок можна зробити висновок, що закон збереження імпульсу є наслідком законів Ньютона, і така думка є цілком природньою. Але, як це нерідко буває в науці, наслідок може виявитися більш глибоким і загальним, ніж положення, із яких його отримано. І справді, закон збереження імпульсу є одним із небагатьох абсолютних фізичних законів. Він виконується не лише в механічних, а й у будь-яких ізольованих фізичних системах і за будь-яких умов. Наприклад, закон збереження імпульсу виконується в електромагнітних полях та у квантових системах, хоча і в першому, і в другому випадку закони Ньютона взагалі не є чинними. Це саме стосується й систем, в яких частинки рухаються із субсвітловими швидкостями. Така універсальність робить закон збереження імпульсу дуже важливим як для теорії, так і для практики. Зокрема, при розгляді маловивчених систем він дозволяє одразу відкидати всі гіпотетичні процеси, в яких не виконується умова (5.1).

Практична цінність закону збереження імпульсу визначається тим, що умова збереження імпульсу $\vec{F} = 0$, може точно чи наближено виконуватись і в незамкнених системах. Це можливо у таких випадках.

1. Зовнішні сили є компенсованими. Так буває, наприклад, при зіткненні куль, які рухаються по гладкій горизонтальній поверхні.

2. В системі імпульси окремих тіл змінюються протягом невеликого проміжку часу під дією дуже великих внутрішніх сил, так що можна нехтувати зовнішніми силами, тобто вважати, що $\vec{F} = 0$. Така ситуація, зокрема, спостерігається при розриві снарядів під дією величезних внутрішніх сил тиску порохових газів.

3. Якщо вектор сумарної зовнішньої сили $\vec{F}$ має незмінний напрям, скажімо, вертикальний OY, то її проекція на будь-який горизонтальний напрям OX $F_x = 0$. Відповідно, зберігається ***проекція*** вектора імпульсу системи на цей напрям:

$$P_x = \sum_i p_{ix} = \mathrm{const} \tag{5.3}$$

4.3. Рух тіл змінної маси (рівняння Мещерського)

Одним з цікавих і важливих наслідків збереження імпульсу є виникнення так званої реактивної сили при зміні маси тіла внаслідок витоку або притоку речовини. Це відбувається, приміром, при русі автомобіля, що поливає вулицю, навантаженні (чи розвантаженні) платформи на ходу, русі ракети з увімкненими двигунами, тощо.

Знайдемо рівняння руху такого тіла на прикладі руху космічної ракети в процесі виходу на орбіту. Позначимо як m і $\vec{v}$ масу ракети та її швидкість відносно Землі. Відповідно, імпульс ракети $\vec{p} = m\vec{v}$. Якщо за час dt двигун ракети викидає масу продуктів згоряння палива δm зі швидкістю $\vec{u}$ відносно ракети, внаслідок чого її маса змінюється на величину $-\delta m$, а швидкість – на величину d$\vec{v}$, то імпульс ракети стає рівним $(m - \delta m)\,(\vec{v} + \delta\vec{v})$. Швидкість витоку продуктів згоряння відносно Землі дорівнює $\vec{v} + \vec{u}$, а їхній імпульс складає $\delta m(\vec{v} + \vec{u})$. Отже, сумарний імпульс системи ракета-продукти згоряння стає рівним $(m - \delta m)\,(\vec{v} + \delta\vec{v}) + \delta m(\vec{v} + \vec{u})$. Зміна імпульсу системи за час dt становить

$$(m - \delta m)\,(\vec{v} + \delta\vec{v}) + \delta m(\vec{v} + \vec{u}) - m\vec{v}.$$

Розкривши дужки і, нехтуючи доданком $\delta m \cdot \mathrm{d}\vec{v}$ (мала вищого порядку відносно інших), одержимо

$$d\,(m\vec{v}) = m d\vec{v} + \vec{u} dm$$

Поділимо обидві частини цієї рівності на dt:

$$\frac{d\,(m\vec{v})}{dt} = m\frac{d\vec{v}}{dt} + \vec{u}\frac{dm}{dt}$$

Згідно з виразом (5.1), ліва частина даного рівняння дорівнює сумарній зовнішній силі $\vec{F}$, яка діє на систему (е нашому випадку – це сила тяжіння, що діє на ракету та на викинуті гази). Тому, врахувавши, що $\delta m = -\mathrm{d}m$ (спад маси ракети), отримаємо:

$$\vec{F} = m\frac{d\vec{v}}{dt} - \vec{u}\frac{dm}{dt} \;\Rightarrow\; m\frac{d\vec{v}}{dt} = \vec{F} + \vec{u}\frac{dm}{dt} \qquad (5.4)$$

Це є *основне рівняння динаміки тіла змінної маси*, яке називається рівнянням Мещерського.

У використаному прикладі зміна маси тіла (ракети) була зумовлена відокремленням речовини (газів), отже dm < 0. Але рівняння Мещерського чинне й у випадку, коли речовина приєднується, наприклад, при завантаженні рухомої платформи піском (у цьому випадку dm > 0).

Рівняння (5.4) показує, що прискорення тіла змінної маси визначається не тільки зовнішньою силою $\vec{F}$, а й величиною

$$\vec{F}_p = \vec{u}\frac{\mathrm{d}m}{\mathrm{d}t}, \qquad (5.5)$$

яку називають ***реактивною силою***. Її величина визначається величиною швидкості зміни маси тіла dm/dt, а напрям – характером зміни маси: якщо маса приєднується (dm/dt > 0), то $\vec{F}_p$ за напрямом збігається з вектором відносної швидкості $\vec{u}$, якщо ж маса відокремлюється (dm/dt < 0), то вектори $\vec{F}_p$ та $\vec{u}$ є антипаралельними. Зокрема, якщо продукти згоряння палива вилітають із сопла ракети вертикально вниз, то реактивна сила напрямлена вертикально вгору, що й забезпечує підйом ракети.

4.4. Центр мас

На загал у будь-якій системі окремі частинки (або тіла) рухаються з різними за модулем і напрямом швидкостями. Тому вираз імпульсу системи через імпульси окремих частинок $\vec{P} = \sum_i m_i \vec{v}_i$ не дає наочного уявлення про напрям та величину швидкості всієї системи. Але таку цілісну картину дає поведінка деякої зв'язаної із системою точки, що називається ***центром мас*** системи. При цьому в багатьох задачах механічну систему виявляється можливим трактувати як одну матеріальну точку.

Положення та рух центра мас. Положення центра мас можна знайти, замінивши у виразі імпульсу системи швидкості частинок $\vec{v}_i$ через їхні радіуси-вектори $\vec{r}_i$ згідно з означенням (1.3) і виконавши низку простих тотожних перетворень:

$$\vec{P} = \sum_i m_i \vec{v}_i = \sum_i m_i \vec{r}_i' = \left(\sum_i m_i \vec{r}_i' \right)' = \sum_i m_i \left(\frac{\sum_i m_i \vec{r}_i}{\sum_i m_i} \right)' = m \left(\frac{\sum_i m_i \vec{r}_i}{\sum_i m_i} \right). \quad (5.6)$$

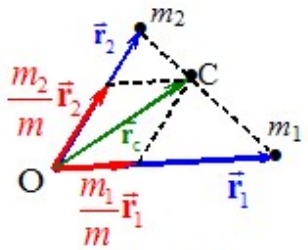


Рис. 5.2

У цих викладках штрихом позначено похідну по часу й символ $m = \sum_i m_i$ означає масу системи.

Величина під знаком похідної у двох останніх виразах являє собою вектор із початком у початку відліку O, тобто, є радіусом-вектором $\vec{r}_c$ певної точки C, яка й називається ***центром мас*** системи. Рис. 5.2 ілюструє графічне визначення радіуса-вектора $\vec{r}_c$ і розташування цієї точки для системи з двох тіл із масами m_1 та $m_2 = 2\, m_1$.

Отже, за означенням ***центр мас*** – то є точка, положення котрої визначається радіусом-вектором

$$\vec{r}_c = \frac{\sum_i m_i \vec{r}_i}{\sum_i m_i} = \frac{\sum_i m_i \vec{r}_i}{m}. \quad (5.7)$$

Відповідно, координати центра мас визначаються як

$$X_c = \frac{\sum\limits_i m_i x_i}{\sum\limits_i m_i} = \frac{\sum\limits_i m_i x_i}{m},$$

$$Y_c = \frac{\sum\limits_i m_i y_i}{\sum\limits_i m_i} = \frac{\sum\limits_i m_i y_i}{m}, \qquad (5.7a)$$

$$Z_c = \frac{\sum\limits_i m_i z_i}{\sum\limits_i m_i} = \frac{\sum\limits_i m_i z_i}{m}.$$

Відтак вираз імпльсу системи (5.6) набуває вигляду:

$$\vec{P} = m\vec{V}_c, \qquad (5.8)$$

де $\vec{V}_c = \vec{r}'_c$ – швидкість центра мас системи, яка визначається через швидкості окремих частинок системи як

$$\vec{V}_c = \frac{\sum\limits_i m_i \vec{v}_i}{m}. \qquad (5.9)$$

Підставивши вираз (5.8) у закон зміни імпульсу системи (5.1), отримаємо *рівняння руху центра мас*:

$$m\frac{d\vec{v}_c}{dt} = \vec{F},$$

або

$$ma_c = \vec{F}.$$

центр мас довільної системи рухається, як точка, в якій зосереджена вся маса системи і до якої прикладені всі зовнішні сили, що діють на тіла системи.

Це твердження називають ***теоремою про рух центра мас***.

Система відліку центра мас (Ц-система). З теореми про рух центра мас випливає, що, коли нас не цікавлять відносні рухи частинок системи, її можна розглядати як одну частинку з масою всієї системи, що розміщена в центрі мас і рухається під дією всіх зовнішніх сил, які діють на систему. Але поняття центра мас виявляється корисним і при розгляді рухів частинок усередині системи. Для цього з центром мас пов'язують систему відліку, яка називається *системою відліку центра мас,* або коротко – *Ц-системою відліку*. Очевидно, що в *Ц*-системі відліку центр мас є нерухомим, отож імпульс системи (величини, що визначені в *Ц*-системі відліку, будемо помічати позначкою ~, яка називається тильда)

$$\tilde{\vec{P}} = \sum_i \tilde{\vec{p}}_i = 0.$$

Інакше кажучи, будь-яка система тіл є нерухомою у своїй власній *Ц*-системі відліку. Ця обставина істотно полегшує розгляд багатьох процесів, зокрема, зіткнень між частинками. Наприклад, для системи з двох частинок

$$\tilde{\vec{p}}_1 + \tilde{\vec{p}}_2 = 0 \quad \Rightarrow \quad \tilde{\vec{p}}_1 = -\tilde{\vec{p}}_2, \qquad (5.11)$$

тобто, в *Ц*-системі відліку двох частинок вони завжди мають однакові за величиною й протилежні за напрямом імпульси.

Така симетричність рухів є дуже зручною, зокрема, при розгляді зіткнень між частинками. Приміром, якщо дві частинки стикаються непружно (“злипаються”), то в *Ц*-системі вони зупиняються, отже в *К*-системі їхня швидкість після зіткнення $\vec{u} = \vec{V}_c$ і визначається виразом (5.9).

Наостанку зауважимо, що *Ц*-система відліку *незамкненої* системи тіл є *неінерціальною*, оскільки $\vec{a}_c \neq 0$ (див. рівняння (5.10)). Але

Ц-система відліку замкненої системи тіл завжди є інерціальною.

4.5. Контрольні запитання

1. Який напрям має імпульс тіла? Системи тіл?

2. Чи змінюється імпульс при рівномірному русі тіла?

3. Чи однакове значення має імпульс тіла або системи тіл у різних інерціальних системах відліку? Зміна імпульсу?

4. Що таке ізольована система? Чи існують у природі строго ізольовані системи?

5. Які сили в системі частинок називаються внутрішніми та які – зовнішніми?

6. Від чого залежить зміна імпульсу системи з часом?

7. У чому полягає закон збереження імпульсу? Яке його значення?

8. Чи можна використовувати закон збереження імпульсу для неізольованих систем? Наведіть приклади.

9. Як виглядає та як записується рівняння Мещерського?

10. Що таке реактивна сила та реактивний рух? Наведіть приклади.

11. Чому при пострілі з рушниці треба щільно притискати її до плеча?

12. Від чого залежить сила тяги реактивного двигуна?

13. Якими формулами визначається положення центра мас системи?

14. Який зв'язок існує між імпульсом системи та рухом її центра мас?

15. Як записується та формулюється теорема про рух центра мас системи?

16. Як визначаються швидкості та імпульси двох частинок у їх Ц - системі відліку?

5. РОБОТА ТА ЕНЕРГІЯ

Для різних класів фізичних систем існують свої специфічні поняття та величини, що характеризують властивості об'єктів та процеси саме в таких системах. Наприклад, прискорення є суто механічною, температура – термодинамічною, а освітленість – оптичною величиною. І лише одна фізична величина однаково притаманна всім фізичним системам і є ***універсальною характеристикою їх стану та процесів*** у них. Ця величина називається терміном ***енергія***, від грецького *energia*, що означає дію, або ж діяльність. Енергія має дві визначальні властивості. По-перше, на відміну від інших величин, існують різні форми енергії: механічна, внутрішня (теплова), електромагнітна, ядерна, тощо. І, по-друге, енергія має здатність переходити з одних форм в інші, причому завжди так, що загальна кількість енергії у Природі лишається незмінною (закон збереження енергії).

Рух та взаємодія тіл характеризуються ***механічною енергією***. При цьому існує спеціальна величина, що є кількісною мірою перетворень механічної енергії - ***механічна робота***, або інакше ***робота сили***.

Далі розглянуті наступні питання:

5.1. Робота та потужність сили

5.2. Кінетична енергія

5.3. Потенціальна енергія

5.4. Закон збереження механічної енергії

5.5. Зіткнення

5.6. Контрольні запитання

5.1. Робота та потужність сили

Елементарною роботою δA сили $\vec{F}$ на нескінченно малому переміщенні $\mathrm{d}\vec{s}$ називається скалярний добуток зазначених векторів:

$$\delta A = \vec{F}\mathrm{d}\vec{s} = F\mathrm{d}s\cos\alpha = F_s\mathrm{d}s. \qquad (6.1)$$

У цій формулі $\mathrm{d}s = |\mathrm{d}\vec{s}|$ – елементарний шлях, $F_s = F\cos\alpha$ – проекція вектора сили на напрям вектора переміщення, α - кут між векторами $\vec{F}$ і $\mathrm{d}\vec{s}$ (рис.6.1). Елементарна робота позначається символом δA, а не $\mathrm{d}A$, тому що вона є малою кількістю роботи, а не її зміною (приростом); поняття "зміна (приріст) роботи" позбавлене змісту.

З означення (6.1) можна сказати, що робота є мірою дії сили на заданому переміщенні (шляху).

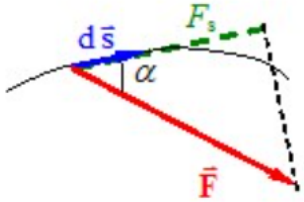


Рис. 6.1

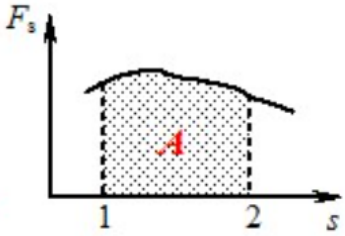


Рис. 6.2

З означення (6.1) випливає, що робота є величиною алгебраїчною, тобто може мати той, чи інший знак, залежно від знаку cosα. Зокрема, коли $\alpha > 90°$, робота сили від'ємна. Інакше говорячи, від'ємну роботу виконує сила, що напрямлена проти переміщення, отже, спричинює гальмівну дію. Зрозуміло, що в цьому випадку переміщення тіла здійснюється або за рахунок якоїсь іншої сили, або за рахунок інерції руху. Тому буває зручно замість від'ємної роботи δA розглядати ***роботу проти даної гальмівної сили***, – додатню величину $\delta A' = -\delta A$. Якщо сила перпендикулярна до переміщення, то вона роботи не виконує (cos90° = 0).

Робота є адитивною величиною. Дослід свідчить, що при переміщенні тіла по траєкторії на скінчену відстань робота сили на всій траєкторії дорівнює сумі робіт, які вона виконує на всіх ділянках траєкторії. Тому робота при переміщенні тіла по траєкторії із заданої точки 1 у задану точку 2 у загальному випадку визначається як ***криволінійний інтеграл*** від виразу (6.1):

$$A = \int_1^2 \vec{F}\mathrm{d}\vec{s} = \int_1^2 F_s\mathrm{d}s. \qquad (6.2)$$

Криволінійний інтеграл відрізняється від звичайного тим, що додаються не добутки значень функції на приріст аргументу, а добутки значень вектора сили на нескінченно малі переміщення тіла вздовж певної лінії – траєкторії руху тіла. Відповідно символи в позначенні інтеграла (границі інтегрування) – то не числа, а позначення початкової та кінцевої точок на траєкторії, вздовж якої обчислюється даний криволінійний інтеграл.

Ураховуючи геометричний зміст інтеграла, роботу можна знаходити з графіка залежності $F_s\mathrm{d}s$: площа під відповідною його ділянкою (заштрихована на рис. 6.2) чисельно дорівнює виконаній роботі.

Окрім того, робота рівнодійної декількох сил на даному переміщенні дорівнює сумі робіт кожної з них – це безпосередньо випливає з виразів (6.2) і (3.1):

$$A = \int_1^2 \vec{F}\,\mathrm{d}\vec{s} = \int_1^2 \left(\vec{F}_1 + \vec{F}_2 + \ldots\right)\mathrm{d}\vec{s} = \int_1^2 \vec{F}_1\mathrm{d}\vec{s} + \int_1^2 \vec{F}_2\mathrm{d}\vec{s} + \ldots = \\ A = A_1 + A_2 + \ldots = \sum_i A_i. \tag{6.3}$$

При русі вздовж прямої під дією сталої за величиною й напрямом сили, що діє під кутом α до напрямку руху, величина $F_s = F\cos\alpha = \text{const}$, і вираз (6.2) дає:

$$A = \int_1^2 F_s\,\mathrm{d}s = F_s\int_1^2 \mathrm{d}s = F_s\,s \quad \Rightarrow \quad A = Fs\cos\alpha. \tag{6.4}$$

Якщо сила діє в напрямку руху (α = 0), то $A = Fs$. На основі цього встановлена одиниця роботи *джоуль* (Дж): 1 Дж = 1 Н·м.

Потужність. З практики добре відомо, що ефективність механізмів і машин (наприклад, двигунів) визначається не стільки величиною виконуваної роботи, скільки тим, як швидко вони здатні її виконувати. Інтенсивність виконання роботи характеризується потужністю. Потужність визначається величиною роботи за одиницю часу й вимірюється у *ватах* (Вт): 1 Вт = 1 Дж/с. Якщо інтенсивність виконання роботи змінюється з часом, то розрізняють середню та миттєву потужності.

Середня потужність – це відношення роботи A, виконаної за проміжок часу t, до його величини:

$$\langle P\rangle = \frac{A}{t}. \tag{6.5}$$

Миттєва потужність характеризує інтенсивність виконання роботи в кожен момент часу й визначається виразом:

$$P = \frac{\delta A}{\mathrm{d}t}, \tag{6.5а}$$

тобто – це відношенням елементарної роботи δA до нескінченно малого проміжку часу $\mathrm{d}t$, за який вона виконується. Урахувавши, що $\delta A = \vec{F}\mathrm{d}\vec{s}$ і $\mathrm{d}\vec{s}/\mathrm{d}t$ – миттєва швидкість, одержимо

$$P = \vec{F}\cdot\vec{v} = Fv\cos\alpha, \tag{6.6}$$

де α - кут між векторами $\vec{F}$ та $\vec{v}$.

Очевидно, що потужність сили є алгебраїчною величиною, знак якої збігається із знаком роботи цієї сили.

5.2. Кінетична енергія

Теорема про кінетичну енергію. Позаяк при русі частинки під дією сили змінюється її швидкість, існує зв'язок між механічною роботою та станом руху частинки. Розглянемо роботу сумарної (рівнодійної) сили $\vec{F}$, яка діє на частинку маси m, на елементарному переміщенні $\mathrm{d}\vec{s}$ по довільній траєкторії (формула (6.1)):

$$\delta A = \vec{F}\mathrm{d}\vec{s}.$$

Урахувавши, що $\vec{F} = m\dfrac{\mathrm{d}\vec{v}}{\mathrm{d}t}$ і $\mathrm{d}\vec{s} = \vec{v}\mathrm{d}t$ ($\vec{v}$ – швидкість частинки), отримаємо:

$$\delta A = m\frac{\mathrm{d}\vec{v}}{\mathrm{d}t}\vec{v}\mathrm{d}\vec{v} = m\vec{v}\mathrm{d}\vec{v}.$$

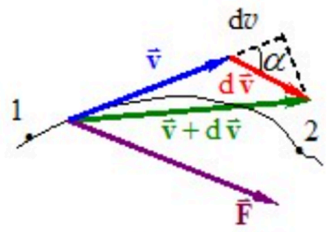


Рис. 6.3

У загальному випадку вектори $\vec{v}$ і $\mathrm{d}\vec{v}$ не збігаються за напрямом (рис. 6.3), тому

$$\vec{v}\mathrm{d}\vec{v} = v|\mathrm{d}\vec{v}|\cos\alpha = v\mathrm{d}v.$$

Цей результат є однією з важливих тотожностей векторної алгебри:

скалярний добуток вектора на його зміну дорівнює добутку модуля цього вектора на зміну його модуля).

Отже, елементарна робота

$$\delta A = mv\mathrm{d}v = \mathrm{d}\left(\frac{mv^2}{2}\right).$$

Бачимо, що елементарна робота дорівнює приросту (зміні) величини

$$K = \frac{mv^2}{2}. \qquad (6.7)$$

Через імпульс тіла $p = mv$ вона визначається, як

$$K = \frac{p^2}{2m}. \qquad (6.7а)$$

Величина K називається ***кінетичною енергією*** частинки. Очевидно, що вона вимірюється у джоулях (Дж).

Таким чином, маємо:

$$\delta A = \mathrm{d}K. \qquad (6.8)$$

Проінтегрувавши це співвідношення уздовж траєкторії від початкової точки 1 до кінцевої точки 2, дістанемо для довільного скінченого переміщення частинки:

$$A = K_2 - K_1 \text{ або } A = \Delta K. \qquad (6.8а)$$

Отримані співвідношення (6.8) і (6.8а) виражають ***теорему про кінетичну енергію***:

зміна кінетичної енергії частинки на будь-якому переміщенні дорівнює роботі всіх сил, які діють на частинку на цьому переміщенні.

Слід наголосити на тому, що співвідношення (6.8) і (6.8а), які виражають зв'язок між роботою та кінетичною енергією, є *універсальними* – вони не залежать від природи та походження сил, які діють на частинку. Зокрема, в неінерціальних системах відліку величина $\vec{F}$ включає й сили інерції. Але при цьому не слід забувати, що йдеться про роботу $рівнодійної$, тобто, про *сумарну роботу всіх сил*. Наприклад, коли ми рівномірно тягнемо санки за мотузку, то виконуємо роботу, проте кінетична енергія санок не змінюється. Але це зовсім не суперечить теоремі про кінетичну енергію, бо таку саму по модулю *від'ємну* роботу виконують сили тертя, і повна робота всіх сил дорівнює нулю.

Кінетична енергія системи. Розглянемо тепер зв'язок між роботою й станом руху для довільної системи частинок. Оскільки робота є адитивною величиною, повна робота всіх сил, що діють у системі, дорівнює сумі робіт A_i, які виконуються силами, що діють на кожну частинку системи. Тому, враховуючи співвідношення (6.8) і (6.3), маємо:

$$\delta A = \sum_i \delta A_i = \sum_i \mathrm{d}K_i = \mathrm{d}\left(\sum_i K_i\right) = \mathrm{d}K,$$

де величина

$$K = \sum_i K_i = \sum_i \frac{m_i v_i^2}{2} = \sum_i \frac{p_i^2}{2m_i}. \qquad \(6.9)$$

Кінетична енергія системи, котра, як видно, теж є адитивною величиною.

Отже, універсальний зв'язок між кінетичною енергією та роботою сил (6.8) і (6.8а) – *теорема про кінетичну енергію – є чинною для довільної системи*.

5.3. Потенціальна енергія

Потенціальні поля та консервативні сили

Якщо на рухоме тіло діє гальмівна сила, то її робота від'ємна ($\delta A < 0$), і, згідно із співвідношенням (6.8), кінетична енергія тіла у процесі руху зменшується ($\mathrm{d}K < 0$). Отже можна сказати, що тіло виконує роботу проти гальмівної сили за рахунок своєї кінетичної енергії. Тому кінетичну енергію можна трактувати як енергію руху – величину, що визначає можливість виконання тілом роботи за рахунок руху. Але через взаємодію між тілами робота може виконуватись і безвідносно до руху, просто внаслідок зміни їхнього взаємного розташування (конфігурації). Відповідно, в механіці крім кінетичної розглядають ще один різновид енергії – ***потенціальну енергію.***

Поняття про поле сил. На загал сили взаємодії залежать від взаємного розташування частинок і швидкостей їхнього відносного руху. Але в деяких випадках сила, що діє на частинку, в кожен момент часу визначається лише її положенням у просторі. В такому разі говорять, що частинка знаходиться в полі сил, або силовому полі. Інакше говорячи,

> ***область простору, в усіх точках якої на частинку діє визначена сила певної природи, називається полем сил, або силовим полем.***

У загальному випадку сили поля можуть змінюватися з часом, але ми будемо розглядати лише *стаціонарні поля*, тобто такі, сили котрих не залежать від часу. Прикладом стаціонарного поля може бути гравітаційне поле Сонця, сили якого визначаються законом всесвітнього тяжіння (3.8), або поле сил тяжіння біля поверхні планети (формула (3.9)). Можна також говорити про поле сил пружності, що виникають при розтягу закріпленого одним кінцем еластичного шнура, електричне поле, що створюється нерухомими зарядами, тощо.

Робота сил однорідного та центрального поля. Найпростішим є однорідне силове поле, тобто поле, в усіх точках якого на частинку діє однакова за величиною й напрямом сила $\vec{F} = \mathrm{const}$. Розглянемо роботу такої сили при переміщенні частинки від точки 1 до точки 2 по довільній траєкторії (рис. 6.4). Зваживши на сталість сили, із загального виразу (6.2) отримаємо:

$$A = \int_1^2 \vec{F}\mathrm{d}\vec{s} = \vec{F}\int_1^2 \mathrm{d}\vec{s}.$$

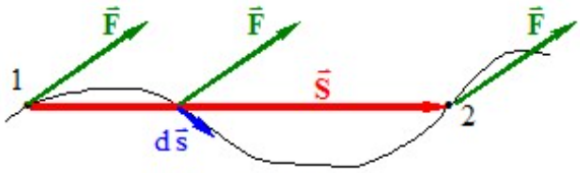


Рис. 6.4

Інтеграл в останньому виразі є сумою низки елементарних векторів $\mathrm{d}\vec{s}$ розміщених на траєкторії. За правилом додавання векторів ця сума являє собою вектор переміщення $\vec{S}$ частинки від точки 1 до точки 2: $\int_1^2 \mathrm{d}\vec{s} = \vec{S}$. Отже, робота

$$A = \vec{F}\cdot\vec{S} = FS\cos\alpha = FS_F, \qquad (6.10)$$

де S_F – проекція вектора $\vec{S}$ на напрям сили.

Очевидно, що вектор $\vec{S}$ не залежить від форми траєкторії, що з'єднує точки 1 і 2. Отже,

робота сил однорідного поля при переміщенні частинки не залежить від траєкторії і визначається тільки початковим та кінцевим положенням частинки.

Аналогічну властивість має й робота сил центрального поля. Центральним називається таке силове поле, в будь-якій точці котрого на частинку діє сила, напрям якої проходить через одну й ту саму точку О ("***силовий центр***"), а величина залежить тільки від відстані до цієї точки (рис. 6.5). Тому, для будь-якої центральної сили можна записати загальний вираз:

$$\vec{F} = F_r(r)\vec{e}_r, \tag{6.11}$$

де $\vec{e}_r$ – орт (одиничний вектор), який указує напрям радіуса-вектора $\vec{r}$ частинки відносно силового центра, $F_r(r)$ – проекція сили на напрям $\vec{r}$, $r = |\vec{r}|$ – відстань від частинки до силового центра.

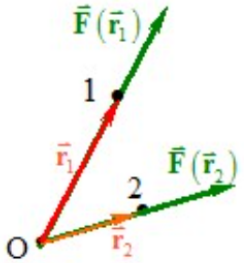


Рис. 6.5

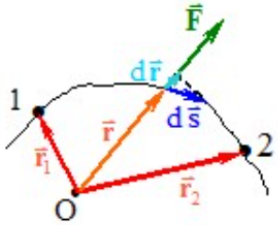


Рис. 6.6

З урахуванням (6.11) і (6.2), для роботи центральної сили при переміщенні частинки по довільній траєкторії від точки 1 до точки 2 (рис. 6.6) маємо:

$$A = \int_1^2 \vec{F}\mathrm{d}\vec{s} = \int_1^2 F_r(r)\vec{e}_r\mathrm{d}\vec{s} = \int_1^2 F_r(r)\mathrm{d}s\cdot\cos\alpha.$$

З рис.6.6 видно, що величина $\mathrm{d}s\cdot\cos\alpha$ дорівнює $\mathrm{d}r$ – зміні відстані частинки до силового центра, отже

$$A = \int_1^2 F_r(r)\mathrm{d}r. \tag{6.12}$$

Прикметно, що в цьому виразі фігурує не радіус-вектор частинки $\vec{r}$, який визначає її положення на траєкторії, а лише відстань r від частинки до силового центра. Це означає, що

робота центральної сили теж не залежить від траєкторії руху частинки.

Консервативні, неконсервативні та дисипативні сили. Розглянуті вище типи сил мають одну й ту саму особливість – незалежність роботи від траєкторії. Тому вони мають спільну загальну назву – "консервативні сили". Отже,

консервативними силами називаються такі, робота котрих не залежить від траєкторії руху частинки і визначається тільки її початковим та кінцевим положенням.

Оскільки при переміщенні частинки по будь-якій замкненій траєкторії її початкове й кінцеве положення збігаються,

робота консервативних сил на замкненій траєкторії дорівнює нулю.

Це твердження також можна розглядати як означення поняття “консервативні сили”. Конкретними прикладами консервативних сил у механіці є гравітаційна сила (3.8) і сила тяжіння (3.9), сила пружності (3.10). Поля, в яких діють консервативні сили, теж мають спеціальну назву – ***потенціальні поля*** *(тому консервативні сили іноді теж називають потенціальними силами).*

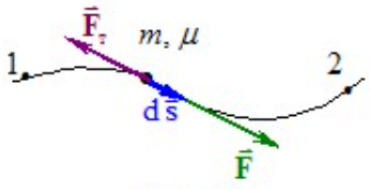


Рис. 6.7

Не всі сили є консервативними, тому всі інші називають ***неконсервативними силами***. Розглянемо, наприклад, роботу сили тертя $\vec{F}_т$ при переміщенні бруска маси m по горизонтальній поверхні з коефіцієнтом тертя μ під дією деякої сили $\vec{F}$ **(**рис. 6.7). Сила тертя в кожній точці траєкторії напрямлена протилежно до переміщення і в даному випадку має модуль $F_т = \mu mg$. Тож її робота на будь-якій ділянці траєкторії дорівнює:

$$A = -\mu mg \int_1^2 \mathrm{d}s = -\mu mgs,$$

де s – пройдений шлях. Зокрема, при русі по замкненій траєкторії s дорівнює довжині траєкторії. Отже, неконсервативність сили тертя є очевидною. Те саме стосується й сили $\vec{F}$, яка забезпечує рух бруска.

Серед усіх неконсервативних сил вирізняють так звані гіроскопічні та дисипативні сили. ***Гіроскопічними*** називають сили, котрі весь час напрямлені перпендикулярно до переміщення частинки, через що взагалі не виконують роботи. Так, гіроскопічними є: сила, що діє на рухомий заряд у магнітному полі (сила Лоренца), сила нормальної реакції опори, коріолісова сила інерції (4.10), тощо.

До ***дисипативних сил*** відносяться різні сили тертя та опору. Характерною властивістю таких сил є те, що вони завжди є *протилежними до напрямків* ***відносних*** *швидкостей* взаємодіючих тіл.

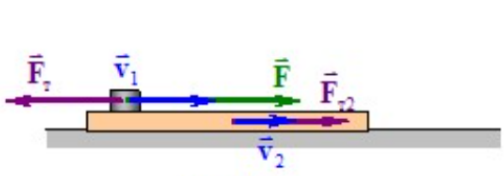


Рис. 6.8

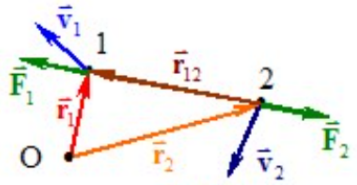


Рис. 6.9

Сказане, одначе, не стосується швидкостей руху тіл *у заданій системі відліку*, тому робота окремих дисипативних сил може бути як від’ємною, так і додатньою. Наприклад, якщо по дошці, що знаходиться на гладкій поверхні, тягти з тертям невеликий брусок (рис. 6.8), то буде рухатися й дошка. При цьому сила тертя $\vec{F}_{т2}$, що забезпечує рух дошки, виконує додатню роботу. Але сумарна робота всіх дисипативних сил взаємодії між частинками будь-якої системи завжди є від’ємною. Доведемо це. Нехай між двома частинками 1 і 2 діють дисипативні сили $\vec{F}_1$ і $\vec{F}_2$ (рис. 6.9), що задовольняють третій закон Ньютона: $\vec{F}_2 = -\vec{F}_1$. В обраній системі відліку з початком у точці О частинки в даний момент часу знаходяться в точках $\vec{r}_1$, $\vec{r}_2$ і мають швидкості $\vec{v}_1$, $\vec{v}_2$. За елементарний проміжок часу $\mathrm{d}t$ частинки здійснюють переміщення $\mathrm{d}\vec{r}_1 = \vec{v}_1 \mathrm{d}t$ і $\mathrm{d}\vec{r}_2 = \vec{v}_2 \mathrm{d}t$ (на рисунку не показані), тож дисипативні сили виконують роботу

$$\delta A_{\text{дис}} = \vec{F}_1 \mathrm{d}\vec{r}_1 + \vec{F}_2 \mathrm{d}\vec{r}_2 = \vec{F}_1 (\vec{v}_1 - \vec{v}_2)\,\mathrm{d}t = \vec{F}_1 \vec{v}_{12}\mathrm{d}t,$$

де, згідно з (1.33), $\vec{v}_{12}$ – швидкість руху першої частинки відносно другої (відносна швидкість). Оскільки сила $\vec{F}_1$ є дисипативною, вона напрямлена протилежно до вектора $\vec{v}_{12}$, отже

$$\delta A_{\text{дис}} = -F_1 v_{12}\mathrm{d}t = -F_1 |\mathrm{d}\vec{r}_{12}| < 0. \qquad (6.13)$$

Якщо сила взаємодії стала, то

$$A_{\text{дис}} = -F_1 |\vec{r}_{12}|. \qquad (6.13а)$$

Таким чином, за будь-яких умов

сумарна робота дисипативних сил взаємодії між двома частинками є від'ємною і визначається не переміщенням кожної частинки окремо, а лише їхнім відносним переміщенням.

Очевидно, що цей висновок є чинним і для довільної системи частинок.

Робота консервативних сил. Розглянемо роботу конкретних консервативних сил: тяжіння, гравітаційних і пружних.

Робота сили тяжіння. Біля поверхні планети сила тяжіння напрямлена вертикально вниз і є однорідною: $m\vec{g} = \text{const}$ (рис. 6.10). Тому її робота при переміщенні частинки з точки 1 у точку 2 по будь-якій траєкторії, згідно з (6.10), визначається як

$$A = mg \cdot s\cos\alpha = mg(z_1 - z_2) \;\Rightarrow\; A = mgz_1 - mgz_1 \qquad (6.14)$$

де z_1 і z_2 – початкова та кінцева вертикальна координата частинки.

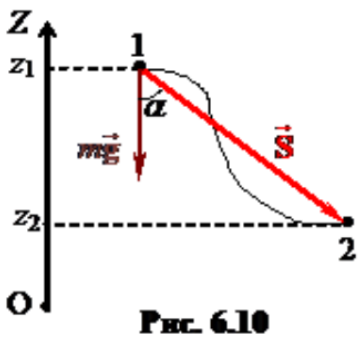


Рис. 6.10

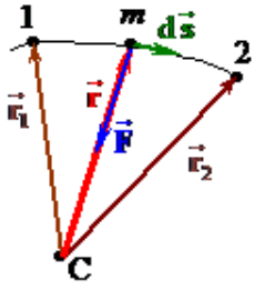


Рис. 6.11

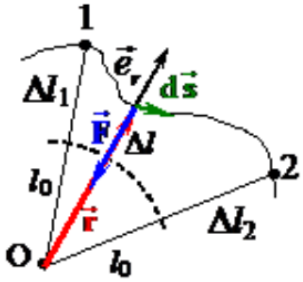


Рис. 6.12

Робота гравітаційної сили. Із закону всесвітнього тяжіння (3.8) випливає, що сила $\vec{F}$, яка діє на будь-яке тіло m (наприклад, комету) в гравітаційному полі іншого тіла M (приміром, Сонця), є центральною (рис. 6.11). Проекція сили $\vec{F}$ на напрям радіуса-вектора тіла $\vec{r}$ дорівнює $F_r = -GMm/r^2$, тому, згідно з (6.12), її робота при переміщенні тіла від точки 1 до точки 2 по довільній траєкторії дорівнює

$$A = \int_1^2 F_r(r)\mathrm{d}r = -GMm\int_1^2 \frac{\mathrm{d}r}{r_2} \quad\Rightarrow\quad A = -G\frac{Mm}{r_1} - \left(-G\frac{Mm}{r_2}\right). \qquad (6.15)$$

Робота пружної сили. Робота пружної сили $\vec{F}$ при переміщенні вільного кінця еластичного шнура (або пружини) по довільній траєкторії (рис. 6.12) визначається аналогічно. Якщо позначити через l_0 довжину недеформованого шнура і через $\Delta l = l - l_0$ величину його деформації (розтягу), то вектор пружної сили, що діє на незакріплений кінець, згідно з (3.10), можна записати як $\vec{F} = -\vec{e}_r k\Delta l$, де k – жорсткість шнура, $\vec{e}_r$ – орт (одиничний вектор), який вказує напрям радіуса-вектора вільного кінця шнура. В такому разі, згідно з (6.12), отримуємо:

$$A = \int_1^2 F_r(r)\mathrm{d}r = -k\int_1^2 \Delta l \cdot \mathrm{d}(\Delta l) \quad \Rightarrow \quad A = \frac{k\Delta l_1^2}{2} - \frac{l\Delta l_2^2}{2} \qquad (6.16)$$

Тут враховано, що проекція сили F_r = -$k\Delta l$ визначається величиною деформації Δl шнура, а зміна відстані від вільного кінця шнура до точки О — зміною цієї величини: dr =d(Δl) .

Потенціальна енергія частинки.

Усі три розглянуті вище консервативні сили утворюють відповідні силові поля. При цьому з виразів (6.14) – (6.16) випливає, що в кожному випадку існує певна фізична величина U, пов'язана з роботою сил співвідношенням:

$$A = U_1 - U_2, \qquad (6.17)$$

або, для елементарних переміщень,

$$\delta A = -\mathrm{d}U. \qquad (6.17а)$$

Така величина U називається ***потенціальною енергією*** (саме тому поля консервативних сил називають потенціальними). Тобто,

> *потенціальною енергією називається величина, спад якої при переміщенні частинки з однієї точки в іншу дорівнює роботі, котра виконується консервативними силами, що діють на частинку.*

(Спадом називають різницю між початковим і кінцевим значенням змінної величини; спад дорівнює взятому з протилежним знаком приростові (зміні) величини. Слід також зазначити, що терміни "приріст" і "спад" не слід трактувати дослівно, як "збільшення" і "зменшення", адже кожна з цих величин, залежно від умов, може виявитись як додатньою, так і від'ємною).

Порівнюючи вираз (6.17) із виразами (6.14) – (6.16), можна дійти висновку, що потенціальна енергія частинки, котра перебуває під дією сил тяжіння, гравітації, або пружності визначається, відповідно, такими формулами:

$$U = mgz, \qquad (6.18)$$

$$U = -G\frac{Mm}{r}, \qquad (6.19)$$

$$U = \frac{k\Delta l^2}{2}. \qquad (6.20)$$

Ці формули й справді є загальновживаними. Але слід зазначити, що жодне зі співвідношень (6.14) – (6.17а) не зміниться, якщо до отриманої з формул (6.18) – (6.20) величини U додати довільне число. Тож однозначно визначеною є не сама потенціальна енергія, лише її зміна ΔU при переміщенні частинки між якимось двома точками. Що ж до значення самої величини U, то його можна вказати тільки по відношенню до заздалегідь обраного ***нульового рівня***, тобто точки (або множини точок), де потенціальна енергія ***приймається*** рівною нулю. Це добре видно з формули (6.18), в якій величина z (вертикальна координата) безпосередньо залежить від вибору початку відліку. Так само формула (6.19) визначає гравітаційну потенціальну енергію не взагалі, а по відношенню до нескінченно віддалених точок (адже U = 0 при $r \to \infty$), а формула (6.20) показує, що пружна потенціальна енергія приймається рівною нулю, коли Δl = 0, тобто, за відсутності деформаціїй. Указаний вибір нульових рівнів є природнім, але не

обов'язковим. Наприклад, легко показати, що формули (6.18) і (6.19), виражають одну й ту саму енергію – потенціальну енергію тіла в гравітаційному полі, – але по відношенню до різних нульових рівнів. Справді, роботу гравітаційної сили при переміщенні тіла біля поверхні Землі можна обчислювати за загальною формулою (6.15):

$$A = -G\frac{mM}{r_1} - \left(-G\frac{mM}{r_2}\right) = GmM\left(\frac{1}{r_2} - \frac{1}{r_1}\right) = \frac{GmM(r_1 - r_2)}{r_1 r_2},$$

де m і M – відповідно, маса тіла та Землі, r_1 і r_2 – відстань від тіла до центра Землі в його початковому та кінцевому положенні. Але, якщо записати $r = R + z$ (R – радіус Землі, z – вертикальна координата тіла, відрахована від земної поверхні), і врахувати, що біля поверхні з великою точністю $r_1 \cdot r_2 = R^2$, отримаємо:

$$A = \frac{GmM}{R^2} \cdot (z_1 - z_2) = mgz_1 - mgz_2$$

де враховані формули (3.9) і (3.10).

Цей результат збігається з виразом (6.14) і випливає з (6.15) як наслідок зміни вибору нульового рівня потенціальної енергії. Так само при розгляді коливань тягарця підвішеного на пружині за нульовий рівень потенціальної енергії пружини зручніше прийняти положення статичної рівноваги тягарця, коли пружина ***має деформацію***, спричинену його вагою.

Особливість потенціальної енергії полягає також у тому, що вона, на відміну від кінетичної, є величиною алгебраїчною, тобто може бути як додатньою, так і від'ємною, залежно від характеру діючих сил і вибору нульового рівня.

Слід указати й на те, що потенціальна енергія частинки в силовому полі, створеному іншим тілом, не "належить" тільки цій частинці. За третім законом Ньютона на "інше тіло" з боку частинки діє така сама за модулем консервативна сила, і слід було би говорити також про потенціальну енергію цього тіла. Отже, насправді потенціальна енергія належить обом тілам і є ***енергією*** *їхньої* ***взаємодії.*** Але коли характеристики тіл дуже відрізняються (як, наприклад, у випадку каменя й Землі), то на стан одного з них (Землі) взаємодії практично не впливає. Тоді таке тіло розглядають не як партнера по взаємодії, а як джерело силового поля, в якому перебуває інше. Тому говоримо не про потенціальну енергію взаємодії каменя й Землі, а про потенціальну енергію каменя в полі сил тяжіння Землі.

Потенціальна енергія системи частинок. *Серед* ***консервативних*** *сил, які діють на частинки системи, можуть бути як зовнішні (сили зовнішніх потенціальних полів), так і внутрішні (сили взаємодії частинок між собою). Тому потенціальну енергію системи U поділяють на* ***зовнішню*** $U_з$ *(енергію в зовнішньому полі) і* ***внутрішню*** *або* ***власну*** $U_в$ *:*

$$U = U_з + U_в. \qquad (6.23)$$

Зовнішня потенціальна енергія визначається співвідношенням (6.17) через сумарну роботу всіх зовнішніх консервативних сил :

$$A = \sum A_i = \sum U_{1i} - \sum U_{2i} \quad \Rightarrow \quad U_3 = \sum U_i$$

У цьому виразі U_i – потенціальні енергії окремих частинок системи в зовнішньому полі. Отже, зовнішня потенціальна енергія системи є адитивною величиною. Але цього не можна сказати про власну та повну потенціальну енергію системи. Справді, нехай дві системи, що мають власні потенціальні енергії $U_{в1}$ і $U_{в2}$, об'єднуються в одну систему, як це схематично показано на рис. 6.14

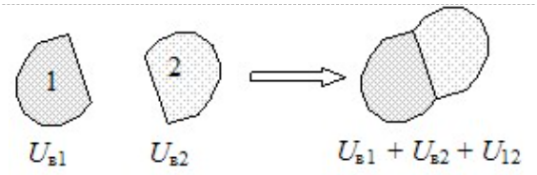


Рис. 6.14

У процесі об'єднання системи 1 і 2 будуть взаємодіяти між собою, тому власна потенціальна енергія об'єднаної системи дорівнює

$$U_{в} = U_{в1} + U_{в2} + U_{12},$$

де U_{12} – енергія взаємодії між частинами об'єднаної системи, рівна роботі, що була виконана силами взаємодії в процесі об'єднання.

Отже, власна потенціальна енергія всієї системи не дорівнює сумі енергій її частин, тобто не є адитивною величиною. Зрозуміло, що те саме може відбуватися не лише при поділі системи чи її складання з окремих "блоків", а й при зміни конфігурації, тобто взаємного розташування частинок системи та відстаней між ними.

Зв'язок між потенціальною енергією та силою.

Оскільки роботу консервативної сили можна знайти як безпосередньо через силу, так і через потенціальну енергію, між цими величинами існує прямий зв'язок. Розглянемо роботу ΔA довільної консервативної сили $\vec{F}$ на невеликому переміщенні $\Delta\vec{S}$ (рис. 6.13). Наближено вона визначається як $A \approx \vec{F} \cdot \Delta\vec{S} = F \cdot \Delta S \cos\alpha = F_S \Delta S$ (неточність зумовлена тим, що в різних точках навіть невеликого переміщення сила може мати відмінні значення), звідки $F_s \approx \dfrac{\Delta A}{\Delta S}$. Але, відповідно до співвідношення (6.17), $A = -\Delta U$, тож $F_s \approx -\dfrac{\Delta U}{\Delta S}$. Точний вираз отримаємо у границі $\Delta S \to 0$:

$$FS = -\lim_{\Delta S \to 0} \frac{\Delta U}{\Delta S}.$$

Така границя в математиці позначається як $\dfrac{\partial U}{\partial S}$ і називається ***похідною функції U по напрямку S***. Вона показує швидкість зміни величини U при переміщенні в заданому напрямку S. Отже, враховуючи цю символіку, можемо записати:

$$F_s = -\frac{\partial U}{\partial S} \qquad (6.21)$$

тобто,

проекція консервативної сили на будь-який напрям дорівнює взятій з протилежним знаком похідній потенціальної енергії в цьому напрямі.

W6-13

Це означає, що сила спрямована в бік зменшення потенціальної енергії і є тим більша, чим швидше змінюється потенціальна енергія в цьому напрямі.

Вираз (6.21) дозволяє аналогічно визначити проекції сили на координатні осі:

$$F_x = -\frac{\partial U}{\partial x}, \quad F_y = -\frac{\partial U}{\partial y}, \quad F_z = -\frac{\partial U}{\partial z}.$$

При цьому елементарні переміщення вздовж координатних осей є нескінченно малими змінами (диференціалами) координат, тобто аргументів функції потенціальної енергії $U = U(x, y, z)$. Тому похідна по кожному з напрямків x, y, z визначається за звичайними правилами диференціювання, але дві інші координати при цьому розглядаються як константи.

За знайденими проекціями вектор сили виражається як $\vec{F} = -\left(\vec{i}\frac{\partial U}{\partial x} + \vec{j}\frac{\partial U}{\partial y} + \vec{k}\frac{\partial U}{\partial z}\right)$. Потреба в подібних операціях і в математиці, і у фізиці виникає доволі часто, тому вираз у дужках має свою назву й позначення. Він називається ***градієнтом*** потенціальної енергії[7] й позначається як **grad*U***:

$$\text{grad}U = \left(\vec{i}\frac{\partial U}{\partial x} + \vec{j}\frac{\partial U}{\partial y} + \vec{k}\frac{\partial U}{\partial z}\right).$$

Зрозуміло, що поняття градієнта відноситься й до інших величин, які є функціями координат.

Таким чином, зв'язок між силою, та потенціальною енергією записується у вигляді:

$$\vec{F} = -\text{grad}U, \qquad (6.22)$$

тобто,

консервативна сила, що діє на частинку, дорівнює взятому з протилежним знаком градієнту потенціальної енергії цієї частинки.

У математиці доводиться, що вектор градієнта даної функції напрямлений у бік її найшвидшого зростання й по модулю дорівнює цій швидкості. Тому

будь-яка консервативна сила спрямована в бік найшвидшого зменшення потенціальної енергії.

Зокрема, із цієї причини вода під дією сили тяжіння стікає по лінії найбільшої крутизни схилу.

Робота сил і механічна енергія частинки.
Сума кінетичної та потенціальної енергії називається ***повною механічною енергією*** (слово "повною" часто опускають):

$$E = K + U. \qquad (6.25)$$

Це означення стосується як окремої частинки, так і довільної системи. При цьому для системи інколи окремо вирізняють суму кінетичної та власної потенціальної енергії, називаючи її *власною механічною енергією* системи.

Механічна енергія частинки. У загальному випадку на окрему частинку діють консервативні сили з боку потенціального поля, в якому вона перебуває, та інші (неконсервативні) сили, які будемо називати ***сторонніми силами.*** Згідно з теоремою про кінетичну енергію (6.8а), сумарна робота усіх сил дорівнює зміні кінетичної енергії частинки:

$$A_к + A_{ст} = \Delta K$$

де $A_к$ і $A_{ст}$ – робота рівнодійної консервативних та рівнодійної сторонніх сил, відповідно. Але робота консервативних сил дорівнює спадові потенціальної енергії частинки (6.17), тому

$$-\Delta U + A_{ст} = \Delta К \quad \rightarrow \quad \Delta K + \Delta U = A_{ст} \quad \rightarrow \quad \Delta(K + U) = A_{ст}$$

$K + U = E$ – то є повна механічна енергія частинки E, отже

$$\Delta E = E_2 - E_1 = A_{ст}, \qquad (6.26)$$

Таким чином,

зміна повної механічної енергії частинки дорівнює роботі всіх сторонніх сил, які діють на неї.

Для елементарного переміщення

$$\mathrm{d}E = \delta A_{\text{ст}} . \qquad (6.26а)$$

Коли сторонні сили відсутні, $\delta A_{\text{ст}} = 0$, $\mathrm{d}E = 0$ і $E = \text{const}$. Отже,

якщо на частинку діють тільки консервативні сили, то її повна механічна енергія зберігається [9]***, тобто, не змінюється під час руху.***

Інакше говорячи, енергія є законсервованою. Звідси походить термін "консервативні сили"

Робота сил та механічна енергія довільної системи.

Розглянемо тепер роботу всіх сил у довільній системі. В загальному випадку як серед внутрішніх, так і серед зовнішніх сил можуть бути і консервативні, і неконсервативні сили. Будь-які зовнішні неконсервативні сили, як і раніше, будемо називати сторонніми силами. Що ж до внутрішніх неконсервативних сил, то в реальних системах – це завжди тільки дисипативні сили та гіроскопічні сили, що роботи не виконують (див. п. 6.3). Отже, повна робота всіх сил у довільній системі складається з роботи консервативних сил (як внутрішніх, так і зовнішніх), внутрішніх дисипативних сил і сторонніх (включно із зовнішніми дисипативними) сил:

$$A = A_{\text{кон}} + A_{\text{дис}} + A_{\text{ст}},$$

де $A_{\text{к}}$ – робота всіх консервативних сил (як внутрішніх, так і зовнішніх), $A_{\text{дис}}$ – робота внутрішніх дисипативних сил, і $A_{\text{ст}}$ – робота сторонніх сил.

Як уже згадувалося, робота всіх сил дорівнює приросту кінетичної енергії, а робота консервативних сил – спадові потенціальної енергії системи, тому

$$\Delta K = -\Delta U + A_{\text{дис}} + A_{\text{ст}} \rightarrow \Delta K + \Delta U = A_{\text{дис}} + A_{\text{ст}} \rightarrow$$
$$\Delta(K + U) = A_{\text{дис}} + A_{\text{ст}}.$$

Урахувавши, що $K + U = E$ – повна механічна енергія системи, отримуємо:

$$\Delta E = E_2 - E_1 = A_{\text{дис}} + A_{\text{ст}}. \qquad (6.27)$$

При елементарній зміні стану системи

$$\mathrm{d}E = \delta A_{\text{дис}} + \delta A_{\text{ст}}. \qquad (6.27а)$$

Таким чином,

зміна повної механічної енергії довільної системи дорівнює сумарній роботі всіх внутрішніх дисипативних сил і всіх сторонніх сил.

Це найбільш загальне (і важливе) співвідношення між механічною енергією та роботою сил.

5.4. Закон збереження механічної енергії

В замкненій (ізольованій) системі зовнішні сили відсутні, тож $\delta A_{ст} = 0$. Тому

$$dE = \delta A_{дис}, \qquad (6.28)$$

$$\Delta E = E_2 - E_1 = A_{дис}. \qquad (6.28а)$$

Отже, механічна енергія замкненої системи ***не зберігається***, її зміна дорівнює роботі дисипативних сил взаємодії між частинками системи. Оскільки завжди $\delta A_{дис} < 0$ (див. п. 6.3), під час руху тіл механічна енергія замкненої системи невпинно зменшується. Але, якщо дисипативні сили відсутні ($\delta A_{дис} = 0$), то $dE = 0$ і $E = const$. У цьому полягає закон збереження механічної енергії :

За відсутності дисипативних сил повна механічна енергія будь-якої замкненої системи зберігається, тобто, не змінюється під час руху частинок системи.

З іншого боку, за відповідних умов механічна енергія може зберігатися і в *незамкненій* системі. Згідно з (6.27) і (6.27а), це буде тоді, коли сумарна робота внутрішніх дисипативних сил та сторонніх сил дорівнює нулю. Тому треба чітко усвідомити, що, на відміну від імпульсу, умовою збереження механічної енергії є не замкненість (ізольованість) системи, а характер діючих в системі сил і величина виконуваної ними роботи.

Як відмічалося, за наявності внутрішніх дисипативних сил механічна енергія замкненої системи поступово зменшується. Але дослід свідчить, що вона не зникає, а перетворюється на інші види, найчастіше в тепло – кожному добре відоме нагрівання рухомих тіл унаслідок тертя (отримана теплова енергія далі передається довкіллю, тобто, відбувається розпорошування, інакше – "**дисипація**" енергії; цим пояснюється назва "**дисипативні сили**"). Це є одним із проявів універсального закону збереження енергії:

енергія не виникає з нічого й не зникає безслідно, вона тільки переходить з одних форм в інші та від одних тіл до інших у рівних кількостях.

Інакше говорячи,

***сума всіх видів енергії у будь-якій замкненій фізичній системі зберігається*, тобто не змінюється з часом.**

Розглянутий аналіз висвітлює роль роботи дисипативних сил у процесах перетворення механічної енергії:

робота дисипативних сил є мірою перетворення механічної енергії в інші, не механічні види.

Якщо в замкненій системі діють тільки консервативні сили, то механічна енергія не змінюється, тобто не переходить у інші види. Але за рахунок роботи консервативних сил відбувається перетворення кінетичної енергії в потенціальну і навпаки. Отже,

робота консервативних сил є мірою перетворення одного виду механічної енергії в інший без зміни її загальної величини

При цьому, коли консервативні сили виконують додатню роботу, потенціальна енергія, переходить у кінетичну, а коли – від'ємну, то кінетична енергія переходить у потенціальну. Це наочно видно у випадку руху тіл під дією сили тяжіння.

5.5. Зіткнення

Закон збереження механічної енергії має не тільки теоретичне, а й велике практичне значення. Разом із законом збереження імпульсу він дозволяє досліджувати складні процеси, в яких неможливо застосувати закони динаміки. Прикладом таких процесів є зіткнення. Процеси зіткнення частинок доводиться розглядати в різних розділах фізики – механіці, молекулярній фізиці, теорії електричного струму, фізиці атомного ядра та елементарних частинок, тощо.

Імпульс і кінетична енергія в системі відліку центра мас. Раніше відмічалося, що процеси зіткнень зручно розглядати в системі відліку центра мас (Ц-системі). Тому з'ясуємо, як визначаються імпульс та кінетична енергія частинок у Ц-системі відліку.

Нехай частинки масами m_1 і m_2, рухаються в певній К-системі відліку зі швидкостями $\vec{v}_1$ і $\vec{v}_2$. Згідно з перетвореннями (1.33), їх швидкості в Ц-системі дорівнюють [11]:

$$\tilde{\vec{v}}_1 = \vec{v}_1 - \vec{V}, \quad \tilde{\vec{v}}_2 = \vec{v}_1 - \vec{V}, \tag{6.29}$$

де $\vec{V}$ – швидкість Ц-системи відліку відносно К-системи. Відповідно, імпульси частинок у Ц-системі

$$\tilde{\vec{p}}_1 = m_1(\vec{v}_1 - \vec{V}) \text{ і } \tilde{\vec{p}}_2 = m_1(\vec{v}_2 - \vec{V}).$$

У записаних виразах і надалі величини, що визначені в Ц-системі відліку, будемо помічати позначкою ~ (тильда) над символом величини.

Підставивши у попередню формулу вираз $\vec{V}$ (5.9), після нескладних перетворень одержимо

$$\tilde{\vec{p}}_1 = \frac{m_1 m_2}{m_1 + m_2}(\vec{v}_1 - \vec{v}_2), \quad \tilde{\vec{p}}_2 = \frac{m_1 m_2}{m_1 + m_2}(\vec{v}_2 - \vec{v}_1). \tag{6.30}$$

Величина

$$\mu = \left(\frac{1}{m_1} + \frac{1}{m_2}\right)^{-1} = \frac{m_1 m_2}{m_1 + m_2} \tag{6.31}$$

називається ***зведеною масою*** двох частинок, тому

$$\tilde{\vec{p}}_1 = \mu(\vec{v}_1 - \vec{v}_2), \quad \tilde{\vec{p}}_2 = \mu(\vec{v}_2 - \vec{v}_1) \tag{6.32}$$

Величина

$$v_r = |\vec{v}_1 - \vec{v}_2| \tag{6.33}$$

- то є відносна швидкість частинок. Отже, імпульси частинок у Ц-системі протилежні за напрямом і мають однаковий модуль

$$p = \mu v_r. \tag{6.34}$$

Відносною швидкістю та зведеною масою визначається й кінетична енергія двох частинок у Ц-системі відліку:

$$\tilde{K} = \tilde{K}_1 + \tilde{K}_2 = \frac{\tilde{p}_1^2}{2m_1} + \frac{\tilde{p}_2^2}{2m_2} = \frac{\tilde{p}^2}{2\mu} = \frac{\mu v_r^2}{2}. \tag{6.35}$$

Абсолютно непружне зіткнення

Тут і далі будемо вважати, що, по-перше, на частинки не діють зовнішні сили (або вони є компенсованими) і, по-друге, частинки взаємодіють тільки в процесі зіткнення, тобто під час контакту. Тоді, хоча імпульс кожної частинки при зіткненні змінюється, їхній сумарний імпульс зберігається (див. п. 5.2).

Абсолютно непружним називають зіткнення, при якому частинки далі рухаються разом, як одне складене тіло (“злипаються”). Розглянемо таке зіткнення двох частинок масами m_1 і m_2, що рухаються в інерціальній К-системі відліку зі швидкостями $\vec{v}_1$ і $\vec{v}_2$. Оскільки імпульс будь-якої системи тіл у Ц-системі відліку дорівнює нулю, то в ній після абсолютно непружного зіткнення швидкість частинок $\tilde{\vec{v}}' = 0$. Це, згідно з (1.33а), означає, що в К-системі відліку їхня швидкість після зіткнення $\vec{v}$ дорівнює швидкості центра мас системи $\vec{V}$. Отже, згідно з (3.9),

$$\vec{v} = \frac{m_1\vec{v}_1 + m_2\vec{v}_2}{m_1 + m_2}. \tag{6.36}$$

Абсолютно непружне зіткнення супроводжується втратами кінетичної енергії, яка частково переходить у внутрішню (теплову) енергію тіл, і вони в тій, чи іншій мірі нагріваються. В Ц-системі відліку внаслідок зіткнення тіла зупиняються, тобто втрачають усю кінетичну енергію. Отже, кількість теплоти, що виділяється при абсолютно непружному зіткненні, дорівнює початковій кінетичній енергії частинок у Ц-системі відліку:

$$Q = \tilde{K} = \frac{\mu v_r^2}{2}. \tag{6.37}$$

У К-системі відліку ця величина складає лише частину кінетичної енергії тіл і, відповідно до (6.31) і (6.33), визначається виразом

$$Q = K_1 - K_2 = \frac{m_1 m_2}{2(m_1 + m_2)}(v_1 - v_2)^2. \tag{6.38}$$

При цьому не втрачена частина кінетичної енергії - то є енергія руху системи частинок як цілого.

Абсолютно пружне центральне зіткнення

Абсолютно пружним називають зіткнення, при якому між тілами діють тільки пружні сили. Як наслідок, після пружного зіткнення тіла розлітаються, тобто, рухаються окремо одне від одного. Під час самого зіткнення в тілах виникають чисто пружні деформації, котрі повністю зникають після розльоту, і внутрішній стан тіл повністю відновлюється. Тому не змінюється й механічна енергія системи. Крім того, через швидкоплинність процесу, під час зіткнення положення тіл у просторі практично не змінюється. Тому не змінюється й потенціальна енергія системи, тож її можна взагалі не враховувати. Таким чином ***при абсолютно пружному зіткненні одночасно зберігаються сумарний імпульс і кінетична енергія тіл, що стикаються***. Це можна розглядати як критерій абсолютно пружного зіткнення.

Існують два різновиди абсолютно пружного зіткнення: центральне (лобове, пряме) та нецентральне (нелобове, косе). При лобовому зіткненні швидкості частинок як до, так і після зіткнення, напрямлені вздовж прямої, що з’єднує частинки. При нелобовому зіткненні напрями руху частинок до та після зіткнення не лежать на одній прямій.

Розглянемо лобове зіткнення двох частинок. В Ц-системі відліку дві частинки завжди мають однакові за величиною і протилежні за напрямом імпульси (див. (6.30)), тому при центральному пружному зіткненні напрям імпульсу кожної частинки

змінюється на протилежний. А оскільки при цьому кінетична енергія системи зберігається, то модулі імпульсу та швидкості частинок лишаються незмінними. Отже,

$$\overrightarrow{\tilde{v}'}_i = -\vec{\tilde{v}}_i, \text{ де } i = 1, 2.$$

Урахувавши це й, скориставшись формулами (6.29), знайдемо швидкості частинок після зіткнення в К-системі відліку:

$$\overrightarrow{v'}_i = \overrightarrow{\tilde{v}'}_i + \vec{V} = \vec{V} - \vec{\tilde{v}}_i = \vec{V} - \left(\vec{v}_i - \vec{V}\right) = 2\vec{V} - \vec{v}_i,$$

де $\vec{V}$ – швидкість центра мас частинок, яка визначається формулою (3.9). Отже,

$$.\vec{v}'_i = 2 \cdot \frac{\sum\limits_i m_i \vec{v}_i}{\sum\limits_i m_i} - \vec{v}_i = \frac{2\sum\limits_i m_i \vec{v}_i - \vec{v}_i \sum\limits_i m_i}{\sum\limits_i m_i}.$$

Підставивши значення індексів, одержимо вираз швидкості для кожної частинки:

$$\vec{v}'_1 = \frac{(m_1 - m_2)\vec{v}_2 + 2m_1\vec{v}_1}{m_1 + m_2}, \quad \overrightarrow{v'}_2 = \frac{(m_2 - m_1)\vec{v}_2 + 2m_1\vec{v}_1}{m_1 + m_2}. \qquad (6.39)$$

Для зіткнення частинок однакової маси ці формули дають $\vec{v}'_1 = \vec{v}_2$ і $\vec{v}'_2 = \vec{v}_1$, тобто, частинки обмінюються швидкостями. Зокрема, якщо друга частинка перед зіткненням перебувала у спокої, то після зіткнення вона почне рухатись із швидкістю першої, а перша зупиниться. У загальному ж випадку для обчислень формули (6.39) треба записати в проекціях на осі вибраної системи координат.

5.6. Контрольні запитання

1. Дайте означення понять “робота” та “потужність” сили й укажіть загальні властивості цих величин. Який фізичний зміст має від’ємна робота сили?

2. Як обчислюється робота сили на заданій траєкторії довільної форми? В яких випадках це можна зробити без застосування інтегрального числення?

3. Який геометричний зміст має робота сили?

4. Укажіть зв’язок між роботою сил і кінетичною енергією? Чи залежить він від фізичної природи сил?

5. Які сили називаються консервативними? Наведіть приклади.

6. Який загальний зміст має поняття “потенціальна енергія” та як вона пов’язана з роботою сил.

7. Які сили називаються дисипативними? Яку особливість вони мають?

8. Що таке “гіроскопічні сили”? Наведіть приклали.

9. Охарактеризуйте зв’язок між повною механічною енергією та роботою сил у незамкненій та в замкненій механічній системі.

10. Сформулюйте загальний закон збереження енергії. Який загальний фізичний зміст має механічна робота?

11. Яка відміна існує між непружним та пружним зіткненням тіл?

12. Виведіть формули для швидкості після непружного зіткнення двох тіл (6.36) і для втрати їхньої кінетичної енергії (6.38).

13. Виведіть формули (6.39) для швидкостей двох тіл після їхнього пружного центрального зіткнення.

6. ЕЛЕМЕНТИ МЕХАНІКИ ТВЕРДОГО ТІЛА

При розгляді рухів твердого тіла поряд із основними лінійними кінематичними величинами (переміщення, швидкість, прискорення) використовують кутові характеристики руху – кут повороту, кутову швидкість і кутове прискорення. В динаміці системи частинок і твердого тіла також використовують відповідні кутові величини, що називаються "моментами" і є аналогами розглянутих раніше динамічних величин. А саме, стан руху визначають моментом імпульсу, силову дію – моментом сили, а інертність – моментом інерції.

У цій лекції розглянуті наступні питання:

6.1. Рівняння моментів

6.2. Закон збереження моменту імпульсу

6.3. Динаміка твердого тіла

6.4. Збереження моменту імпульсу відносно осі

6.5. Кінетична енергія обертального і плоского руху

6.6. Гіроскопи

6.7. Контрольні запитання

6.1. Рівняння моментів

Розглянемо зміст динамічних кутових величин і зв'язок між ними.

Момент імпульсу. Нехай рухома частинка маси m має швидкість $\vec{v}$ та імпульс $\vec{p} = m\vec{v}$, а її положення відносно обраного початку відліку О визначається радіусом-вектором $\vec{r}$ (рис. 7.1). Тоді за означенням

моментом імпульсу частинки відносно даної точки О називається векторний добуток векторів $\vec{r}$ і $\vec{p}$:

$$\vec{L} = [\vec{r}\vec{p}]. \tag{7.1}$$

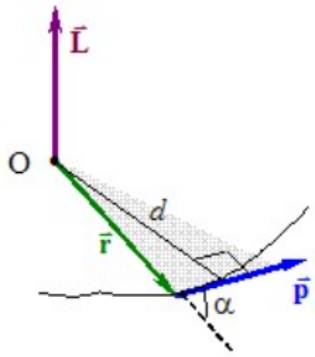


Рис. 7.1

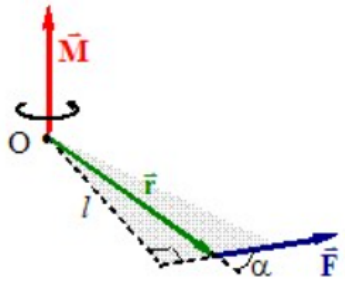


Рис. 7.2

З цього означення випливає, що момент імпульсу є аксіальним вектором, тобто, його напрям визначається правилом правого гвинта, а модуль дорівнює:

$$L = rp\sin\alpha = pd, \tag{7.1a}$$

де α – кут між напрямами векторів $\vec{r}$ і $\vec{p}$, а $d = r\sin\alpha$ – відстань від точки О до лінії вектора $\vec{p}$, яка називається ***плечем вектора*** $\vec{p}$.

Правило правого гвинта: якщо обертати правий гвинт від напрямку вектора $\vec{r}$ до напрямку вектора $\vec{p}$, то він буде вгвинчуватись у напрямку вектора $\vec{L}$.

Вектор $\vec{L}$, початок якого прийнято розміщувати в точці О, є перпендикулярним до площини векторів $\vec{r}$ і $\vec{p}$, тобто, до площини, в якій відбувається рух частинки. Тому при русі частинки по плоскій замкненій траєкторії (орбіті) навколо заданої точки О напрям вектора моменту імпульсу відносно цієї точки визначає розташування (орієнтацію) орбіти частинки у просторі.

Момент імпульсу, як і імпульс, є адитивною величиною: момент імпульсу системи частинок відносно даної точки дорівнює векторній сумі моментів імпульсу всіх частинок системи відносно тієї ж точки:

$$\vec{L} = \sum_i \vec{L}_i. \tag{7.2}$$

Момент сили. Аналогічно до означення (7.1), ***моментом сили відносно точки О*** називають вектор

$$\vec{M} = \left[\vec{r}\vec{F}\right]. \tag{7.3}$$

Вектор $\vec{M}$ також є аксіальним вектором: його напрям визначається правилом правого гвинта (рис. 7.2), а модуль – формулою

$$M = rF\sin\alpha = lF, \qquad (7.4)$$

де l – *плече сили* $\vec{F}$ відносно точки О, тобто, відстань від точки О до лінії дії цієї сили.

Із означення (7.3) випливають наступні загальні властивості моменту сили.

– ***момент центральної сили*** (див. лекцію 6, п. 6.3) ***відносно силового центра дорівнює нулю;***

– ***момент сили не змінюється при перенесенні вектора сили уздовж його лінії*** (рис. 7.3а)**.** Це дозволяє при розгляді задач для зручності розміщувати точку прикладання сили не там, де вона виникає, наприклад у точці дотику нитки та прив'язаного до неї тіла, а в будь-якому місці на лінії дії сили.

Окрім того, ***сумарний момент сил взаємодії між двома частинками*** $\vec{F}_2 = -\vec{F}_1$ (рис. 7.3б) ***відносно будь-якої точки дорівнює нулю***:

$$\vec{M} = \vec{M}_1 + \vec{M}_2 = \left[\vec{r}_1, \vec{F}_1\right] + \left[\vec{r}_2, \vec{F}_2\right] = \left[\left(\vec{r}_1 - \vec{r}_2\right), \vec{F}_1\right] = \left[\vec{r}_{12}, \vec{F}_1\right] = 0. \quad (7.5)$$

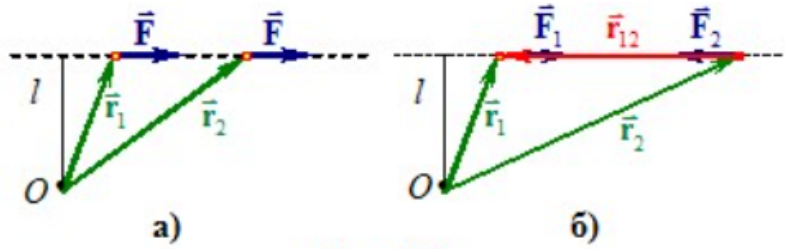


Рис. 7.3

Рівняння моментів для частинки. За винятком рівномірного прямолінійного руху, момент імпульсу частинки змінюється з часом. Аби з'ясувати, чим визначається ця зміна, знайдемо похідну по часу від вектора моменту імпульсу $\vec{L}$:

$$\frac{\mathrm{d}\vec{L}}{\mathrm{d}t} = \left[\frac{\mathrm{d}\vec{r}}{\mathrm{d}t}, \vec{p}\right] + \left[\vec{r}, \frac{\mathrm{d}\vec{p}}{\mathrm{d}t}\right].$$

Знаходження похідної від векторного добутку виконується так само, як для скалярного добутку двох функцій. Єдина особливість полягає в тому, що порядок множників не можна змінювати.

Похідна $\mathrm{d}\vec{r}/\mathrm{d}t = \vec{v}$ – це швидкість частинки. Тому, позаяк вектори $\vec{v}$ і $\vec{p} = m\vec{v}$ є збіжними за напрямом, перший доданок дорівнює нулю. У другому доданку $\dfrac{\mathrm{d}\vec{p}}{\mathrm{d}t} = \vec{F}$ – сила, що діє на частинку. Отже,

$$\frac{\mathrm{d}\vec{L}}{\mathrm{d}t} = \left[\vec{r}, \vec{F}\right],$$

або

$$\frac{\mathrm{d}\vec{L}}{\mathrm{d}t} = \vec{M}. \qquad (7.6)$$

Таким чином,

> ***швидкість зміни вектора моменту імпульсу частинки відносно даної точки дорівнює моменту прикладеної до частинки рівнодійної сили відносно цієї точки.***

З рівняння (7.6) випливає також, що

$$\mathrm{d}\vec{L} = \vec{M}\mathrm{d}t, \qquad (7.7)$$

де величина в правій частині називається ***імпульсом моменту сили*** за проміжок часу dt.

Отже,

> ***зміна моменту імпульсу частинки дорівнює імпульсу моменту сил, що діють на неї***.

Зміна моменту імпульсу частинки за скінченний проміжок часу $[t_1,\ t_2]$ визначається інтегруванням виразу (7.7):

$$\Delta\vec{L} = \vec{L}_2 - \vec{L}_1 = \int_{t_1}^{t_2} \vec{M}(t)\mathrm{d}t. \qquad (7.7а)$$

Рівняння (7.6), (7.7) і (7.7а) називаються ***рівняннями моментів***.

Як за змістом, так і за формою вказані рівняння є аналогами відповідних рівнянь другого закону Ньютона (3.3), (3.4) і (3.4а), причому роль імпульсу відіграє момент імпульсу, а роль сили – момент сили. Зауважимо також, що рівняння моментів можна використовувати і в неінерціальних системах відліку, якщо крім моментів сил взаємодії враховувати й моменти сил інерції.

Рівняння моментів для довільної системи частинок легко одержати, узагальнюючи рівняння (7.6), (7.7) і (7.7а) для системи. Згідно з (7.2) і (7.6), швидкість зміни моменту імпульсу системи

$$\frac{\mathrm{d}\vec{L}}{\mathrm{d}t} = \sum_i \frac{\mathrm{d}\vec{L}_i}{\mathrm{d}t} = \sum_i \vec{M}_{i_{\text{зовн}}},$$

де $\vec{L}_i$ – моменти імпульсів окремих частинок системи відносно заданої точки, $\vec{M}_i$ – моменти всіх сил, які діють на кожну частинку, відносно тієї ж точки. Очевидно, що повний момент сил можна подати, як суму моментів внутрішніх та зовнішніх сил:

$$\sum_i \vec{M}_i = \sum_i \vec{M}_{i_{\text{вн}}} + \sum_i \vec{M}_{i_{\text{зовн}}}.$$

Оскільки в будь-якій системі повний момент внутрішніх сил є сумою моментів сил взаємодії між кожною парою тіл, то, згідно з (7.5), перший доданок у цьому виразі дорівнює нулю, і для $\mathrm{d}\vec{L}/\mathrm{d}t$ маємо:

$$\frac{\mathrm{d}\vec{L}}{\mathrm{d}t} = \vec{M}_{\text{зовн}} \qquad (7.8)$$

де

$$\vec{M}_{\text{зовн}} = \sum_i \vec{M}_{i_{\text{зовн}}}.$$

Таким чином,

> ***швидкість зміни моменту імпульсу довільної системи відносно даної точки дорівнює сумарному (повному) моменту зовнішніх сил, що діють на тіла системи, відносно тієї самої точки.***

Відповідно, ***зміна моменту імпульсу*** системи за нескінченно малий, або скінченний проміжок часу визначається ***імпульсом моменту зовнішніх сил*** за цей проміжок:

$$d\vec{L} = \vec{M}_{\text{зовн}} dt \tag{7.9}$$

$$\Delta\vec{L} = \vec{L}_2 - \vec{L}_1 = \int_{t_1}^{t_2} \vec{M}(t) dt \tag{7.9a}$$

Рівняння (7.8), (7.9) і (7.9а) **є *рівняннями моментів для довільної системи*.**

6.2. Закон збереження моменту імпульсу

Якщо система замкнена, то $\vec{F}_{i\,\text{зовн}} = 0$ і $\vec{M}_{i\,\text{зовн}} = 0$. Тоді з рівняння моментів (7.8) маємо:

$$\frac{d\vec{L}}{dt} = 0 \quad \Rightarrow \quad \vec{L} = \sum L_i = const \tag{7.10}$$

У цьому полягає *закон* ***збереження моменту імпульсу***:

момент імпульсу замкненої системи зберігається,

тобто, не змінюється з часом. При цьому моменти імпульсу окремих частинок, звичайно, можуть змінюватися внаслідок взаємодії з іншими частинками системи. Але ці зміни мають характер обміну, так що повний момент імпульсу системи лишається незмінним.

Слід зазначити, що момент імпульсу зберігається і в незамкнених системах ($\vec{F}_{i_{\text{повн}}} \neq 0$) за умови, що зовнішні сили або не створюють моментів ($\vec{M}_{i_{\text{зовн}}} = 0$), або сума цих моментів дорівнює нулю (але ця умова, на відміну від моментів внутрішніх сил, може виконуватися лише для певної точки (або точок) відліку). Прикладом може слугувати рух планет. На кожну планету з боку Сонця діє зовнішня сила тяжіння. Напрям цієї сили проходить через центри як Сонця, так і планети, тож вона на створює моментів відносно цих точок. Тому вектори моментів імпульсу орбітального руху планет та їхнього добового обертання навколо власної осі є незмінними. Одним із наслідків цього є незмінність просторової орієнтації площини орбіти планети та орієнтації осі добового обертання планети відносно площини орбіти. Зокрема, земна вісь зберігає свій нахил до площини орбіти, що спричинює зміну пір року.

За логікою наведених вище міркувань, закон збереження моменту імпульсу випливає з рівняння моментів (7.8), тобто постає як наслідок основних законів динаміки. Але насправді закон збереження моменту імпульсу, як і закон збереження імпульсу (див. лекцію 5, п. 5.2), є *фундаментальним* законом природи. Зокрема, закон збереження моменту імпульсу виконується і в немеханічних системах, які не підпорядковані законам Ньютона, таких як електромагнітне випромінювання, атоми, ядра, тощо.

6.3. Динаміка твердого тіла

Твердим тілом називається тіло, відстань між будь-якими двома точками котрого не змінюється в процесі руху. Отже, тверде тіло можна розглядати як систему жорстко зв'язаних матеріальних точок, яка задовольняє розглянуті вище положення механіки системи частинок.

У кінематиці відмічалося, що довільний рух твердого тіла можна трактувати як сукупність поступального та обертального рухів. Тому для визначення руху твердого тіла треба задати рівняння руху будь-якої жорстко зв'язаної з ним точки та рівняння руху тіла відносно цієї точки. Найзручнішою такою точкою є центр мас тіла (зручність такого вибору зумовлена тим, що, як виявляється, сумарний момент сил інерції відносно центра мас завжди дорівнює нулю). Відповідно, загальними рівняннями динаміки твердого тіла є рівняння руху центра мас (3.10) і рівняння моментів (7.8) відносно центра мас:

$$m\vec{a}_c = \vec{F}, \tag{7.11}$$

$$\frac{\mathrm{d}\vec{L}}{\mathrm{d}t} = \vec{M}_c, \tag{7.12}$$

де $\vec{F}$ – сума всіх сил, які діють на тіло (включно із силами інерції), а $\vec{M}_c$ – **сумарний момент *лише сил взаємодії***. При цьому немає потреби говорити, що йдеться про зовнішні сили, оскільки внутрішні сили взаємодії між частинками твердого тіла ніяк не впливають на його механічну енергію та рух, і їх взагалі не розглядають.

Ці рівняння виглядають досить простими, але насправді механіка довільного руху твердого тіла є доволі складною. Тому обмежимося розглядом лише простих рухів твердого тіла – ***обертання навколо нерухомої осі та плоского руху***. Ще простіший поступальний рух повністю визначається рухом однієї точки тіла (рівняння (7.11)) і не потребує спеціального розгляду.

При обертанні навколо нерухомої осі та плоскому рухах напрям осі обертання тіла не змінюється. Тому замість векторного рівняння (7.12) зручно використовувати рівняння моментів у проекціях на вісь обертання Z:

$$\frac{\mathrm{d}L_z}{\mathrm{d}t} = M_z. \tag{7.13}$$

Проекції векторів моменту імпульсу L_z та моменту сили M_z на задану вісь Z називають ***моментом імпульсу відносно осі*** та ***моментом сили відносно осі***. Отже, моменти відносно осі є не векторними, а алгебраїчними величинами. Розглянемо, чим вони визначаються.

Момент імпульсу відносно осі та момент інерції. Обертове тверде тіло можна розглядати як сукупність маленьких часточок масами m_i, що рухаються навколо спільної осі Z по колах радіусів R_i із швидкостями $v_i = \omega R_i$, де ω – модуль кутової швидкості тіла (рис. 7.4). Відносно довільної точки відліку O на осі обертання окрема частинка у певний момент часу має вектор імпульсу $\vec{p}_i = \Delta m_i \vec{v}_i$ та вектор моменту імпульсу $\vec{L} = [\vec{r}_i,\ \vec{p}_i]$. Модуль моменту імпульсу цієї частинки відносно осі Z

$$|L_{zi}| = L_i \cos\alpha_i = r_i \Delta m_i v_i \sin\vartheta_i = \Delta m_i v_i R_i = \Delta m_i R_i^2 \omega,$$

де враховано, що вектори $\vec{r}_i$ та $\vec{p}_i$ взаємно перпендикулярні, $\alpha = 90^\circ - \vartheta$ та $r\sin\vartheta = R$.

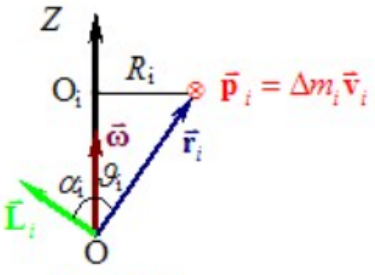


Рис. 7.4

Модуль моменту імпульсу всього тіла відносно осі Z визначається як

$$|L_z| = \sum_i |L_{zi}| = \sum_i \Delta m_i v_i R_i = \left(\sum_i \Delta m_i R_i^2 \right) \omega.$$

Урахувавши, що знак L_z збігається зі знаком ω_z, одержимо вираз моменту імпульсу твердого тіла відносно осі обертання:

$$L_z = \left(\sum \Delta m_i R_i^2 \right) \omega_z \tag{7.14}$$

Величина

$$I = \sum_i \Delta m_i R_i^2 \tag{7.15}$$

називається ***моментом інерції*** тіла відносно заданої осі. Отже,

момент імпульсу твердого тіла відносно осі обертання дорівнює добутку його моменту інерції відносно цієї осі на проекцію кутової швидкості:

$$L_z = I\omega_z. \tag{7.16}$$

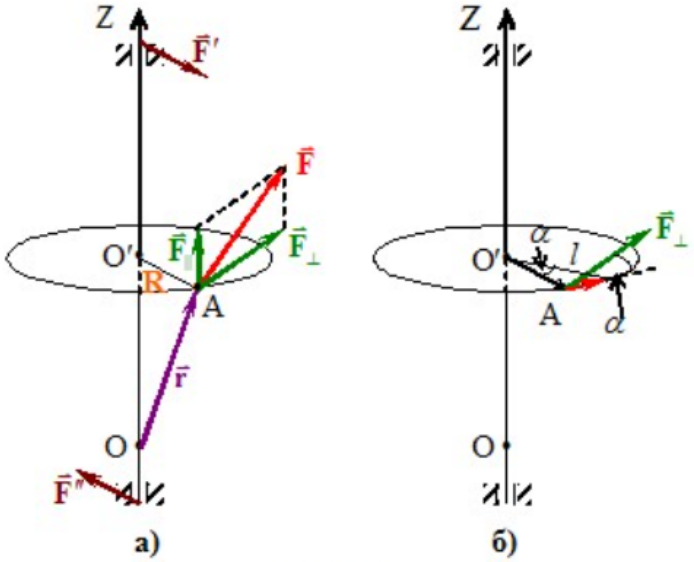


Рис. 7.5

Момент сили відносно осі

Розглянемо тепер, як обчислюється момент сили M_z відносно закріпленої (фіксованої) осі обертання тіла. Нехай до деякої точки А тіла, що обертається навколо ***закріпленої*** осі Z, прикладена довільна сила $\vec{F}$ (рис. 7.5а). Силу $\vec{F}$ можна розглядати як суму двох складових – перпендикулярної $\vec{F}_\perp$ та паралельної $\vec{F}_\parallel$ до осі обертання, так що $\vec{F} = \vec{F}_\perp + \vec{F}_\parallel$. Неважко збагнути, що складова $\vec{F}_\perp$ обертає тіло навколо осі Z, а $\vec{F}_\parallel$ намагається повернути саму вісь. Але, оскільки вісь обертання тіла є закріпленою, то

момент сили $\vec{F}_{\parallel}$ компенсується моментом сил реакції $\vec{F}'$ і $\vec{F}''$ у точках закріплення осі (підшипниках), тож силу $\vec{F}_{\parallel}$ можна не враховувати. Крім того можна показати, що момент сили відносно осі не залежить від положення початку відліку О на цій осі. Тому величину M_z можна розглядати як проекцію вектора моменту сили $\vec{F}_{\perp}$ відносно точки О′ – центра кола, по якому рухається точка А (рис. 7.5б). Отже, згідно з (7.4),

$$M_z = \pm F_{\perp} \cdot l, \qquad (7.17)$$

де $F_{\perp}$ – модуль сили $\vec{F}_{\perp}$ і l – плече цієї сили відносно осі обертання. Знак у виразі (7.17) співпадає із знаком проекції вектора $\vec{M}'_{\perp}$ на вісь Z. Цю формулу можна подати й інакше, якщо врахувати, що $l = R\cos\alpha$, де α – кут між вектором $\vec{F}_{\perp}$ й ортом τ напрямку обертання точка А по колу, і що $F\cos\alpha = F_{\tau}$ – то є проекція сили $\vec{F}_{\perp}$ (і $\vec{F}$) на орт $\vec{\tau}$:

$$M_z = F_{\tau} R \qquad (7.17а)$$

Рівняння динаміки обертального руху твердого тіла.

Підставивши вираз (7.16) у рівняння моментів (7.13), отримаємо:

$$\frac{\mathrm{d}(I\omega)}{\mathrm{d}t} = M_z \quad \Rightarrow \quad I\frac{\mathrm{d}\omega_z}{\mathrm{d}t} = M_z,$$

або

$$I\beta_z = M_z, \qquad (7.18)$$

де I – момент інерції тіла відносно осі обертання, β_z – проекція кутового прискорення на цю вісь.

Рівняння (7.18) як за формою, так і за змістом, є аналогом основного рівняння динаміки матеріальної точки (3.5) і називається ***основним рівнянням динаміки обертального руху твердого тіла***. Воно дозволяє визначити кутове прискорення, а потім, за допомогою рівнянь кінематики (2.11) і (2.12), й інші характеристики обертального руху твердого тіла.

З рівняння (7.18) випливає ще й те, що момент інерції тіла при обертальному русі відіграє таку саму роль, як маса при поступальному русі, тобто, ***момент інерції є мірою інертності обертового тіла***. При цьому така “обертальна” інертність залежить не тільки від маси тіла, а й від її розподілу відносно осі обертання.

Укажемо й на те, що рівняння (7.18), разом із (7.11), використовується також при розгляді плоского руху тіла, замість загального рівняння моментів (7.12).

Обчислення моменту інерції. Теорема Штайнера.

Згідно з формулою (7.15) момент інерції визначається не тільки величиною маси тіла, але й розташуванням окремих його частин відносно осі обертання, внаслідок чого момент інерції одного й того ж тіла має різні значення відносно різних осей. Тому для застосування рівняння (7.18) спочатку треба визначити момент інерції тіла відносно заданої осі.

Для довільного тіла ця задача аналітично не розв’язується, і момент інерції доводиться визначати дослідним шляхом. Але для тіл симетричної форми та із симетричним розподілом маси момент інерції можна обчислити. Для цього, найперше, треба взяти до уваги, що маса суцільного тіла неперервно розподілена по об’єму, і тому вираз (7.15) має тільки символічний зміст. Для реальних обчислень замість малих частинок тіла Δm треба розглядати елементарні маси $\mathrm{d}m$, і дискретне додавання замінити інтегруванням:

$$I = \int r^2 \mathrm{d}m = \int_V r^2 \rho(r) \mathrm{d}V, \tag{7.19}$$

де r – відстань від даної точки тіла до осі, відносно якої обчислюється момент інерції, $\rho(r)$ – густина речовини тіла в даній точці, яка в загальному випадку може залежати від r, і $\mathrm{d}V$ – об'єм нескінченно малої ділянки тіла в околі даної точки.

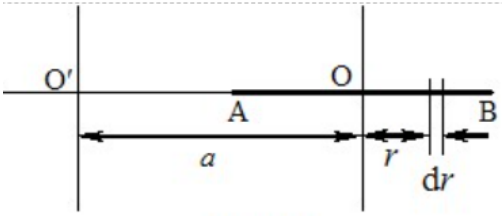


Рис. 7.6

Як приклад обчислимо момент інерції однорідного тонкого стержня маси m і довжини l відносно перпендикулярної до нього осі, що проходить через центр інерції O (центр мас), та відносно паралельної осі O′, що розташована на довільній відстані a від точки O (рис 7.6). Напрямимо координатну вісь r уздовж стержня й розмістимо початок відліку в його центрі мас O, тобто посередині. На довільній відстані r від початку відліку виділимо нескінченно малу ділянку стержня довжини $\mathrm{d}r$, маса якої дорівнює $\mathrm{d}m = (m/l) \cdot \mathrm{d}r$. Тоді для осі O, згідно з (7.19), і, враховуючи координати кінців стрижня A та B, отримаємо:

$$I_0 = \int r^2 \mathrm{d}m = \frac{m}{l} \int_{-\frac{l}{2}}^{\frac{l}{2}} r^2 \mathrm{d}r = \frac{ml^2}{12}. \tag{7.19a}$$

При обчисленні моменту інерції I стержня відносно осі O′ координата r відраховується від точки O′, отже $r_A = a - l/2$ і $r_B = a + l/2$, отже

$$I = \int r^2 \mathrm{d}m = \frac{m}{l} \int_{a-\frac{l}{2}}^{a+\frac{l}{2}} r^2 \mathrm{d}r = \frac{ml^2}{12} + ma^2.$$

У цьому виразі перший доданок – то є момент інерції стержня I_0 відносно осі, що проходить через центр мас, тобто:

$$I = I_0 + ma^2 \tag{7.20}$$

Можна довести, що цей результат є чинним і для будь-якого іншого тіла, незалежно від його форми. Він складає важливу ***теорему Штайнера***:

> ***момент інерції довільного тіла відносно заданої осі дорівнює сумі його моменту інерції відносно паралельної осі, що проходить через його центр інерції, та добутку маси тіла на квадрат відстані між цими осями.***

Ця теорема полегшує знаходження моментів інерції симетричних тіл відносно осей, паралельних до осей симетрії тіла.

На завершення наведемо вирази моментів інерції Iо деяких однорідних симетричних тіл відносно осей, що проходять через центр інерції і є осями симетрії тіла:

Табл. 1.

Тіло	**Вісь**	**Момент інерції I_0**
Тонки стержень довжини l	Перпендикулярна до стержня	$\frac{1}{12}ml^2$
Суцільний циліндр радіуса R	Співпадає з віссю циліндра	$\frac{1}{2}mR^2$
Суцільна куля радіуса R	Проходить через центр кулі	$\frac{2}{5}mR^2$
Суцільний однорідний конус із радіусом основи R	Співпадає з віссю конуса	$\frac{3}{10}mR^2$

6.4. Збереження моменту імпульсу відносно осі

З рівняння моментів (7.13), за умови, що сумарний момент зовнішніх сил $M_z = 0$, випливає закон збереження моменту імпульсу відносно осі для окремого тіла та системи тіл:

$$\frac{\mathrm{d}L_z}{\mathrm{d}t} = 0 \quad \Rightarrow \quad L_z = \mathrm{const},$$

$$\sum L_z = \mathrm{const}.$$

При цьому повний момент імпульсу може змінюватися з часом (вектор $\vec{L} \neq \mathrm{const}$). У розгорнутому вигляді і з урахуванням виразу (7.16) закон збереження моменту імпульсу відносно осі можна записати так:

$$I_1\omega_{z1} + I_2\omega_{z2} + \ldots = I_1\omega'_{1z} + I_2\omega'_{z2} + \ldots \qquad (7.21)$$

Величини в лівій та правій частині цього рівняння визначені для спільної осі Z і відносяться до двох довільних моментів часу.

Цікавий прояв закону збереження моменту імпульсу відносно осі обертання тіла спостерігається при зміні його моменту інерції під дією внутрішніх сил. У цьому випадку завдяки збереженню моменту імпульсу змінюється кутова швидкість обертання тіла:

$$I_1\omega_1 = I_2\omega_2 \quad \Rightarrow \quad \frac{\omega_2}{\omega_1} = \frac{I_1}{I_2}. \qquad (7.22)$$

Цей ефект, зокрема, широко використовують спортсмени в різних видах спорту. Наприклад, для пришвидшення свого обертання фігурист здіймає руки догори, зводячи їх над головою, або ж притискає їх до тіла, а для уповільнення – розводить їх у сторони. Подивитися відео Цікава демонстрація американських астронавтів

6.5. Кінетична енергія обертального і плоского руху

Кінетична енергія обертального руху тіла

Кінетична енергія твердого тіла, як адитивна величина, дорівнює сумі кінетичних енергій окремих його частинок:

$$K = \sum_i \frac{\Delta m_i v_i^2}{2} \qquad (7.23)$$

де Δm_i та v_i – маси та швидкості окремих частинок тіла, відповідно.

Якщо тіло обертається навколо нерухомої осі, то всі його точки рухаються по колах радіусів R_i з однаковою кутовою швидкістю ω, отже $v_i = \omega R_i$, і кінетична енергія обертального руху

$$K = \frac{1}{2}\left(\sum_i \Delta m_i R_i^2\right)\omega^2 \qquad (7.23а)$$

Вираз у дужках – то є момент інерції тіла I відносно осі обертання (див. (7.15)), отже, кінетична енергія обертального руху твердого тіла визначається формулою

$$K = \frac{I\omega^2}{2} \qquad (7.24)$$

Урахувавши вираз (7.16), можна також записати:

$$K = \frac{L_z^2}{2I} \qquad (7.24а)$$

Кінетична енергія плоского руху

При плоскому русі кожна частинка Δm_i твердого тіла рухається разом із центром мас із швидкістю $\vec{v}_c$ і обертається відносно центра мас (у Ц-системі відліку) із швидкістю $\tilde{\vec{v}}_i$ по колу навколо осі перпендикулярної до площини руху центра мас. Отже, в обраній К-системі відліку кожна частинка твердого тіла, згідно з (1.33а), має швидкість $\vec{v}_i = \tilde{\vec{v}}_i + \vec{v}_c$. Підставивши цей вираз у (7.21), отримаємо:

$$K = \sum_i \frac{\Delta m_i\left(\vec{v}_c + \tilde{\vec{v}}_i\right)^2}{2} = \left(\sum_i \Delta m_i\right)\frac{v_c^2}{2} + \sum_i \Delta m_i \tilde{\vec{v}}_i + \sum_i \frac{\Delta m_i \tilde{v}_i^2}{2}$$

Оскільки $\sum_i \Delta m_i$ – то є маса всього тіла m, перший доданок у цьому виразі визначає кінетичну енергію поступального руху $K_{\text{п}} = mv_c^2/2$. Другий доданок дорівнює нулю, бо в ньому другий множник є виразом імпульсу тіла в Ц-системі відліку, що завжди дорівнює нулю (див. (5.11)). Швидкості $\tilde{v}_i$ – це лінійні швидкості обертального руху точок тіла навколо фіксованої осі, що проходить через центр мас. Тому останній доданок, згідно з (7.23) і (7.23а), визначає кінетичну енергію обертального руху тіла, і його можна замінити виразом (7.24). Таким чином для кінетичної енергії плоского руху тіла маємо:

$$K_{\text{пл}} = \frac{I\omega^2}{2} + \frac{mv_c^2}{2}, \qquad (7.25)$$

де I – момент інерції відносно осі обертання, що проходить через центр мас, v_c – швидкість руху центра мас.

Корисно звернути увагу на те, що формули кінетичної енергії обертального руху твердого тіла (7.24) і (7.24а) є аналогами формул (6.7) і (6.7а) кінетичної енергії руху матеріальної точки чи поступального руху твердого тіла. Ця аналогія поширюється й на зв'язок між кінетичною енергією та роботою (теорема про кінетичну енергію, (лекція 6, п. 6.2), у чому легко переконатись. Розглянемо роботу, що виконується при обертанні тіла навколо нерухомої осі. В цьому випадку прикладена до тіла сила $\vec{F}$ є перпендикулярною до осі обертання. При елементарному повороті тіла на кут dφ точка прикладання сили А здійснює переміщення d$\vec{s}$ по колу радіуса R (рис. 7.7), і сила виконує роботу $\delta A = \vec{F} \cdot \mathrm{d}\vec{s} = F\mathrm{d}s\cos\alpha$. У цьому виразі d$s$ = Rdφ і F cosα = F_τ – проекція сили $\vec{F}$ на дотичну до кола, по якому рухається точка А. Отже, врахувавши вираз (7.17а), маємо:

$$\delta A = F_\tau R\mathrm{d}\varphi \quad \Rightarrow \quad \delta A = M_z\mathrm{d}\varphi. \tag{7.26}$$

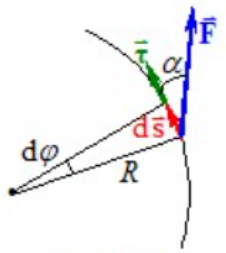


Рис. 7.7

При повороті на скінченний кут

$$A = \int_0^{\varphi} M_z\mathrm{d}\varphi. \tag{7.26а}$$

Ці вирази, що визначають *роботу моменту сил відносно осі*, теж є аналогічними до відповідних виразів (6.1), (6.2) із механіки точки.

Зв'язок між роботою моменту сил і кінетичною енергією обертального руху отримаємо після нескладних викладок, виразивши в (7.26), величину M_z із рівняння (7.18) і врахувавши співвідношення кінематики твердого тіла (2.5):

$$\delta A = M_z\mathrm{d}\varphi = I\beta_z\mathrm{d}\varphi = I\beta_z\omega_z\mathrm{d}t = I\omega_z\mathrm{d}\omega_z = \mathrm{d}\left(\frac{I\omega^2}{2}\right) \quad \Rightarrow \quad \delta A = \mathrm{d}K_{об}. \tag{7.27}$$

При повороті на скінченний кут

$$A = \Delta K_{об} = K_{об2} - K_{об1}. \tag{7.27а}$$

Отже, теорема про кінетичну енергію зберігає чинність і для обертального руху твердого тіла.

6.6. Гіроскопи

У механізмах і машинах окрім тіл, які обертаються навколо нерухомої осі чи здійснюють плоскі рухи (колеса, маховики, шатуни), використовують і так звані **гіроскопи** – масивні симетричні обертові тіла, осі обертання яких не є жорстко закріпленими і можуть змінювати напрям під час руху. Строга теорія гіроскопа є досить складною, тому обмежимося лише спрощеним розглядом окремих властивостей гіроскопів.

Зв'язок між векторами $\vec{L}$ і $\vec{\omega}$**.** Визначимо вектор моменту імпульсу $\vec{L}$ симетричного однорідного тіла, яке обертається зі сталою кутовою швидкістю $\vec{\omega}$ навколо закріпленої осі, що збігається з його віссю симетрії. Для цього розглянемо сумарний момент імпульсу $\Delta\vec{L}_i$ будь-якої пари однакових симетричних частинок тіла масами Δm_i, які мають швидкості $\vec{v}_i$ та $\vec{v}_i' = -\vec{v}_i$ (рис 7.8):

$$\Delta\overrightarrow{L_i} = [\vec{r}_i, \Delta m_i\vec{v}_i] + \left[\vec{r}_i{}', \Delta m_i\vec{v}'_i\right] = [\vec{r}_i, \Delta m_i\vec{v}_i] - \left[\vec{r}_i{}', \Delta m_i\vec{v}_i\right] =$$
$$\left[\left(\vec{r}_i - \vec{r}_i{}'\right), \Delta m\vec{v}\right] = [\vec{r}_{ii'}, \Delta m\vec{v}].$$

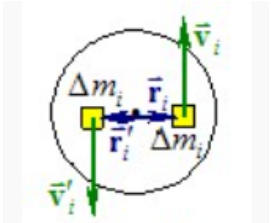


Рис. 7.8

Оскільки вектор $\vec{r}_i{}' = (\vec{r}_i - \vec{r}_i{}')$ лежить на діаметрі кола, по якому рухаються частинки (див. рис. 7.9а), вектор їхнього сумарного моменту імпульсу $\Delta\vec{L}_i$ за напрямом збігається з вектором кутової швидкості тіла $\vec{\omega}$. Із подібних парних елементів складається все симетричне тіло, тому його повний момент імпульсу а $\vec{L}$ напрямлений так само. В такому разі, враховуючи (7.16), маємо:

$$\vec{L} = I\vec{\omega}. \qquad (7.28)$$

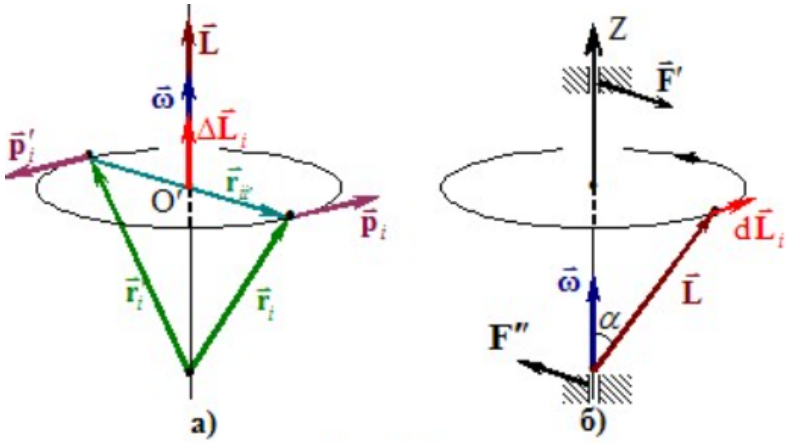


Рис. 7.9

Отже, вектор моменту імпульсу симетричного тіла, що обертається навколо закріпленої осі симетрії, дорівнює добутку його моменту інерції відносно цієї осі на вектор кутової швидкості і не залежить від положення точки відліку на осі обертання. Якщо ж вісь обертання ***не збігається*** з віссю симетрії тіла, то вектор $\vec{L}$ складає якийсь кут α з вектором $\vec{\omega}$ (рис. 7.9б). Тому, обертаючись разом із тілом, вектор $\vec{L}$ описує конус, а його кінець рухається по колу в перпендикулярній до осі обертання площині. Отже,

вектор моменту імпульсу тіла за будь-який елементарний проміжок часу dt набуває приросту, вектор якого $\mathrm{d}\vec{L}$ є дотичним до цього кола і перпендикулярним до осі обертання. У відповідності до рівнянням (7.6), це означає, що на тіло діє певний момент сил однакового з $\mathrm{d}\vec{L}$ напрямку. Він створюється поперечними силами реакції опор (підшипників) $\vec{F}'$ і $\vec{F}''$, які виникають через асиметричність розподілу маси тіла відносно осі обертання. Відповідно, з боку осі на підшипники діють такі самі сили протилежного напрямку, котрі можуть призвести до швидкого зносу чи навіть руйнування підшипників і спричинюють шкідливі вібрації пристрою, в якому закріплені підшипники. Тому при виготовленні обертових деталей та вузлів механізмів і машин забезпечують їхнє балансування, тобто максимально симетричний розподіл маси відносно осі обертання.

Вільний зрівноважений гіроскоп.

Як говорилося, гіроскопом називається масивне симетричне тіло (маховик), що може обертатися з великою швидкістю навколо осі симетрії. Зазвичай гіроскоп уміщують у так званий кардановий підвіс, який являє собою систему з трьох концентричних рухомих кілець[9] (рис. 7.10а). Маховик і кільця можуть вільно обертатися навколо взаємно перпендикулярних осей AA′, BB′ і DD′ (рис.7.10б), що забезпечує можливість орієнтації осі гіроскопа AA′ у будь-якому напрямі. Тому гіроскоп у кардановому підвісі називається вільним гіроскопом. Зазвичай одна з точок осі гіроскопа є закріпленою, і її називають *точкою опори* гіроскопа. У показаному на рис. 7.10 гіроскопі точкою опори є спільний центр кілець. Нарешті, якщо точка опори збігається з центром мас маховика, то гіроскоп називають *зрівноваженим*.

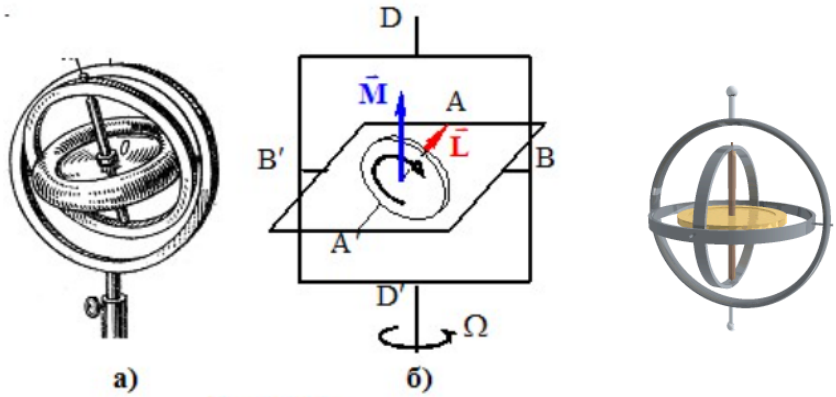


Рис. 7.10

Очевидно, що сума моментів усіх зовнішніх сил, які діють на вільний зрівноважений гіроскоп, дорівнює нулю. Тому, згідно із законом збереження моменту імпульсу, напрям осі гіроскопа буде лишатися незмінним навіть, якщо його каркас буде рухатися довільним чином. На цьому базується робота гірокомпаса – приладу, який дозволяє визначати напрям у просторі, не використовуючи магнітне поле Землі.

З властивостями гіроскопа можна познайомитися, переглянувши фільми за наведеними нижче посиланнями

Гіроскопічний ефект.

Якщо на гіроскоп почне діяти зовнішній момент сил, то його момент імпульсу буде змінюватись у відповідності до рівняння моментів (7.7):

$$\mathrm{d}\vec{L} = \vec{M}\mathrm{d}t. \qquad (7.29)$$

При цьому вісь гіроскопа поводиться досить несподівано. Якщо у зображеному на рис. 7.11а вільному зрівноваженому гіроскопі (кардановий підвіс не показано) спробувати повернути вісь гіроскопа AA′ за годинниковою стрілкою навколо перпендикулярної до неї горизонтальної осі BB′, приклавши дві антипаралельні сили (пару сил) $\vec{F}_1$ – $\vec{F}_2$, то

всупереч очікуванням вісь гіроскопа повернеться не у вертикальній площині навколо горизонтальної осі BB′, а в горизонтальній площині навколо вертикальної осі DD′. Така поведінка осі гіроскопа під дією зовнішніх сил називається ***гіроскопічним ефектом***. Хоча цей ефект важко сприймається, він прямо випливає з рівняння моментів. Справді, прикладена пара сил створює відносно точки опори гіроскопа O момент $\vec{M}$, який за правилом правого гвинта напрямлений уздовж горизонтальної осі BB′ (рис. 7.11а). Тож за з рівнянням (7.29) такий самий напрям має і вектор приросту моменту імпульсу гіроскопа $d\vec{L}$. А це означає, що вектор $\vec{M}$, а з ним і вісь гіроскопа AA′ , повертається в горизонтальній площині навколо вертикальної осі $_{DD}$,.

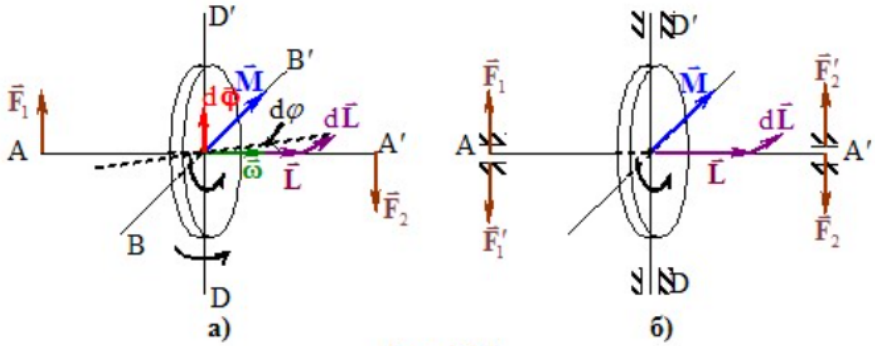


Рис. 7.11

Уявімо тепер тепер, що кільце, в діаметральних точках якого A, A′ зафіксована в підшипниках вісь гіроскопа, розташоване горизонтально і рівномірно обертається в підшипниках навколо вертикальної осі DD ′ (рис 7.11б). Тоді разом із кільцем у горизонтальній площині обертається і вісь гіроскопа AA′ та вектор його моменту імпульсу $d\vec{L}$. Позаяк він залишається в горизонтальній площині, вектор його приросту $d\vec{L}$ теж обертається в горизонтальній площині і є перпендикулярним до $\vec{L}$, тобто до осі гіроскопа. Згідно з (7.29), це означає, що при вимушеному обертанні осі гіроскопа на нього діє якийсь напрямлений ***горизонтально*** момент сил . Цей момент створюється ***вертикальними*** силами реакції , в підшипниках осі гіроскопа (рис. 7.11б). У свою чергу такі самі протилежно напрямлені сили і діють з боку осі на підшипники. Ці сили називають ***гіроскопічними силами***, а створюваний ними момент – гіроскопічним моментом. Значні гіроскопічні сили виникають, наприклад, у підшипниках масивних роторів турбін кораблів при різких змінах курсу та у штормову погоду, або в підшипниках карданних валів автомобілів на крутих віражах, що може призвести до руйнування підшипників.

Але гіроскопічний ефект знаходить і корисні застосування зумовлені тим, що гіроскоп чинить спротив зміні напрямку осі обертання. Приміром, масивні гіроскопи використовують для стабілізації положення корабля, або вагонів монорейкової залізниці. Прояви гіроскопічного ефекту спостерігаються й при їзді на велосипеді чи мотоциклі, позаяк колеса швидко обертаються і “працюють” як гіроскопи. Це не дає байкеру падати при русі по прямій. При поворотах він також використовує гіроскопічний ефект. Відхиляючись при повороті від вертикалі, велосипедист повертає вісь переднього (кермового) колеса у вертикальній площині навколо напрямку руху, внаслідок чого виникає гіроскопічний момент, який повертає вісь колеса в горизонтальній площині, спонукаючи велосипед повертати в бік нахилу велосипедиста.

Прецесія гироскопа

Ще одним проявом гіроскопічного ефекту є так званий ***прецесійний рух*** осі гіроскопа, або просто ***прецесія***. Розглянемо це явище на прикладі відомої дитячої іграшки дзиґи – симетричного тіла, яке може швидко обертатись навколо власної осі на горизонтальній опорі як незрівноважений гіроскоп. Поставлена вертикально нерухома дзиґа неодмінно завалюється під дією сили тяжіння. Але розкручена дзиґа не падає навіть якщо відхилена – її вісь описує конус навколо вертикалі (рис. 7.12). Такий рух

називається ***прецесією*** *осі.*

Поведінка осі дзиґи принципово не відрізняється від поведінки осі зрівноваженого гіроскопа і повністю пояснюється рівнянням моментів . Справді, сила тяжіння $m\vec{g}$, яка прикладена в центрі мас дзиґи, створює зовнішній момент $\vec{M}$ відносно точки опори О. Вектор $\vec{M}$, тож і вектор приросту моменту імпульсу $d\vec{L}$, весь час є перпендикулярними до вертикальної площини, в якій лежить вісь дзиґи та збіжний із нею за напрямом вектор $d\vec{L}$. Тому вектор моменту імпульсу та вісь дзиґи під дією сили тяжіння не наближаються до площини точки закріплення, а ***прецесують*** – описують конуси навколо вертикалі OZ із незмінним кутом розкриття. (Насправді через неодмінну наявність тертя в точці О та опору повітря обертання дзиґи буде поступово уповільнюватись, а кут α – зростати, і згодом дзиґа таки впаде).

Знайдемо кутову швидкість (частоту) прецесії Ω, скориставшись рис. 7.12 (початок вектора $\vec{L}$ знаходиться в т. О). За час dt кінець вектора $\vec{L}$ рухається навколо осі OZ по колу радіуса $L\sin\alpha$ і за час dt повертається на кут $\mathrm{d}\varphi = \mathrm{d}L/L\sin\alpha$. А за рівнянням моментів $\mathrm{d}L = M\mathrm{d}t$. Звідси для кутової швидкості прецесії

$\Omega = \mathrm{d}\varphi/dt$ виходить:

$$\Omega = \frac{M}{L\sin\alpha}. \qquad (7.30)$$

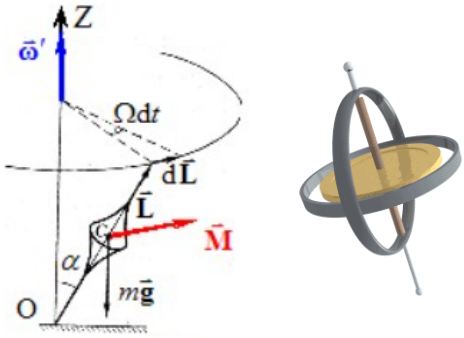


Рис. 7.12

Модуль моменту сили тяжіння відносно точки О дорівнює добутку її величини на відстань від центра мас дзиґи до осі OZ (плече), отже, у виразі (7.30) $M = mgb\sin\alpha$, де b – відстань від точки опори О до центра мас дзиґи С. Урахувавши також, що модуль моменту імпульсу $L = I\omega$, де ω – кутова швидкість обертання дзиґи навколо власної осі, отримаємо вираз кутової швидкості прецесії дзиґи

$$\Omega = \frac{mgb}{I\omega}. \qquad (7.31)$$

Варто зауважити, що завдяки прецесії кожна точка гіроскопа, крім обертання навколо власної осі із кутовою швидкістю ω, рухається ще і з кутовою швидкістю прецесії Ω навколо осі, що проходить через точку опори. Тому використаний при виведенні формули (7.31) вираз є наближеним. Але при великій частоті обертання $\Omega << \omega$, і похибка є неістотною.

Зрозуміло, що вираз (7.30) чинний не лише для дзиґи, на яку діє момент сил тяжіння, а й для будь-якого незрівноваженого гіроскопа, що перебуває під дією зовнішнього моменту сил $\vec{M}$. Тому доцільно співвідношення (7.30) записати векторно в загальному вигляді. Урахувавши напрямки векторів $\vec{L}$, $\vec{M}$ і $\vec{\Omega}$, отримаємо:

$$M = \Omega L\sin\alpha \quad \Rightarrow \quad \vec{M} = \left[\vec{\Omega}, \vec{L}\right]. \qquad (7.32)$$

Із цього виразу випливає цікава особливість прецесійного обертання осі гіроскопа: зовнішній момент сил визначає не кутове прискорення, як для "звичайного" тіла, а саму кутову швидкість прецесії. Отже, щойно зникає зовнішній момент, одразу припиняється й прецесія. Тому таку прецесію називають вимушеною.

6.7. Контрольні запитання

1. Що називається моментом імпульсу матеріальної точки? Твердого тіла?

2. Частинка маси m, що лежить на осі OZ у точці $x = l$, починає рухатися паралельно до осі OX із заданим сталим прискоренням $\vec{a} = a\vec{i}$. Записати вираз вектора моменту імпульсу $\vec{L}$ частинки відносно початку координат і модуля L цього вектора.

3. Відомо, що сума імпульсів усіх частинок системи відносно центра мас дорівнює нулю. Чи стосується це й моменту імпульсу?

4. Що називається моментом сили відносно точки? Від чого залежить модуль цього вектора?

5. Дві антипаралельні сили однакової величини називаються парою сил. Довести, що момент пари сил не залежить від положення точки, відносно якої він визначається. Від чого залежить модуль моменту пари сил?

6. Сума всіх внутрішніх сил у будь-якій системі дорівнює нулю. Чи це так і для суми моментів цих сил?

7. Що називається моментом імпульсу відносно осі? Якою формулою визначається момент імпульсу відносно осі для твердого тіла?

8. Що таке момент інерції тіла? Який фізичний зміст цієї величини?

9. Чи має задане тіло якийсь момент інерції? Чи має його нерухоме тіло відносно заданої осі ?

10. У чому полягає теорема Штайнера? Яке вона має практичне призначення?

11. Що називається моментом сили відносно осі? Як він обчислюється?

12. Що таке плече сили? Чи можна сказати, що модуль моменту сили відносно осі дорівнює добутку модуля сили на її плече?

13. Чи можна твердити, що під дією заданої сили більш масивне тіло отримає менше кутове прискорення, ніж менш масивне? Чому?

14. Записати вирази кінетичної енергії тіла, що обертається навколо нерухомої осі.

15. Записати вирази кінетичної енергії тіла, що здійснює плоский рух.

16. Як обчислюється робота моменту сили?

17. Коли виконується та як записується закон збереження моменту імпульсу відносно осі? Наведіть прояви цього закону.

18. У чому полягає гіроскопічний ефект? Як він пояснюється?

19. Що таке прецесія гіроскопа? Від чого та як залежить кутова швидкість прецесії?

7. СПЕЦІАЛЬНАЯ ТЕОРІЯ ВІДНОСНОСТІ. КІНЕМАТИКА

Ще на самому початку (див. ВСТУП) говорилося про те, що закони класичної механіки є чинними лише для повільних рухів тіл. При швидких рухах усі співвідношення кінематики й динаміки суттєво змінюються й визначаються законами спеціальної теорії відносності (СТВ), створеної на початку 20-го сторіччя А. Ейнштейном. Причина обмеженості механіки Ньютона виявилася дуже глибокою й пов'язаною із складнішими, ніж уважалося, властивостями самого простору та часу. Тому СТВ по суті трактує про фізичні властивості простору й часу в інерціальних системах відліку.

У цій лекції розглянуті наступні питання:

7.1. Принцип відносності Галілея

7.2. Постулати СТВ, відносність простору та часу

7.3. Перетворення Лоренца

7.4. Перетворення швидкостей в СТВ, гранична швидкість

7.5. Інтервал

7.6. Контрольні запитання

7.1. Принцип відносності Галілея

Простір і час у класичній механіці

Простір і час є фундаментальними і всеосяжними сутностями буття – все існує й усе відбувається в просторі та часі. При всьому тому простір і час самі є фізичним об'єктом, який виявляє себе через взаємне розташування тіл і відстані між ними та через послідовність подій і тривалість процесів і станів. Наукові уявлення про властивості простору та часу сформувалися на основі спостережень і дослідів і найнаочніше відображені в основних поняттях і законах механіки. Зокрема, такими основними властивостями є однорідність й ізотропність простору та однорідність часу, про що свідчить незалежність законів і рівнянь механіки в заданому типі систем відліку від положення початку координат і напрямку координатних осей та від вибору початку відліку часу. Крім того, класична фізика виходить з уявлення про абсолютність і незалежність простору та часу, що відображено в перетвореннях Галілея (1.30), (1.30а) і (1.31). Це може здатися дивним, оскільки галілеєві перетворення говорять про відносність положення тіла в просторі: $\vec{r} \neq \vec{r}'$. Але це свідчить лише про те, що в просторі не існує якоїсь особливої виділеної точки, відносно якої слід фіксувати положення всіх тіл. Але взаємне розташування тіл і відстані між ними, які, власно, й відображають властивості простору, є абсолютними. Справді, з рис. 8.1 видно, що положення точки 1 відносно точки 2 визначається радіусом-вектором $\vec{r}_{12} = \vec{r}_1 - \vec{r}_2$, який, згідно з (1.30) або (1.30а), не змінюється при переході від однієї системи відліку до іншої, тобто, є ***інваріантом*** перетворень Галілея:

$$\vec{r}_{12} = \vec{r}_1 - \vec{r}_2 = (\vec{r}_1' + \vec{V}t) - (\vec{r}_2' + \vec{V}t) = \vec{r}_1' - \vec{r}_2' = \vec{r}_{12}'. \qquad (8.1)$$

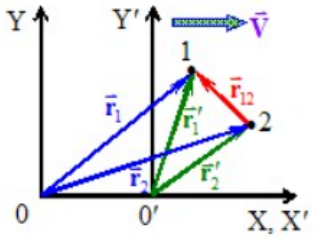


Рис. 8.1

Отож і відстань між двома точками $l = |\vec{r}_{12}|$ є інваріантною, тобто *абсолютною* величиною:

$$l = l'. \qquad (8.1а)$$

Указані інваріанти перетворень Галілея відображають концепцію ***абсолютного простору***, тобто, єдиного для всього Всесвіту й ні від чого не залежного "вмістилища речей".

Так само класична фізика виходить і з концепції єдиного і ні від чого не залежного ***абсолютного часу***, котра безпосередньо закладена в перетворення Галілея (1.31):

$$t = t'.$$

Іншою абсолютною кінематичною величиною (інваріантом перетворень Галілея) в класичній механіці є відносна швидкість (швидкість руху одного тіла відносно іншого), що визначається виразом

$$\vec{v}_{12} = \frac{\mathrm{d}\vec{r}_{12}}{\mathrm{d}t} = \vec{v}_1 - \vec{v}_2. \qquad (8.2)$$

Згідно з (8.1), величина $\mathrm{d}\vec{r}_{12}$ не змінюється при переході від однієї системи відліку до будь-якої іншої, отже, $\mathrm{d}\vec{r}_{12} = \mathrm{d}\vec{r}'_{12}$, і відносна швидкість також є інваріантом перетворень Галілея:

$$\vec{v}_{12} = \vec{v}'_{12} \qquad (8.3)$$

В інерціальних системах відліку (ІСВ) інваріантом перетворень Галілея є також прискорення: $\vec{a} = \vec{a}'$. Це випливає з означення (1.6) і формул (1.33) і (1.33а), оскільки для ІСВ $\vec{V} = \mathrm{const}$.

З іншого боку (див. Лекція 3, п. 3.2) сила взаємодії між тілами, загалом, залежить тільки від їхнього ***взаємного*** положення та ***відносної*** швидкості. Тому сила взаємодії теж є інваріантом перетворень Галілея – вона однакова в усіх інерціальних системах відліку: $\vec{F}\,' = \vec{F}$. Те саме можна сказати й про масу, оскільки вона є *внутрішньою* властивістю тіла: $m' = m$. Як наслідок, інваріантом перетворень Галілея виявляється й основне рівняння класичної механіки – рівняння другого закону Ньютона. Це означає, що в усіх інерціальних системах відліку рух матеріальної точки визначається одними й тими самими рівняннями (3.3) і (3.7). Звідси випливає ***принцип відносності Галілея*** або *принцип відносності класичної механіки*:

за однакових умов у всіх інерціальних системах відліку всі механічні явища відбуваються однаково.

Тому ніякими механічними дослідами *всередині даної ІСВ* неможливо встановити рухається вона, чи перебуває у спокої.

Таким чином, принцип відносності Галілея проголошує повну еквівалентність усіх інерціальних систем відліку *щодо механічних явищ*.

7.2. Постулати СТВ, відносність простору та часу

Закони Ньютона й перетворення Галілея, що складають фундамент всієї класичної механіки, ґрунтуються на уявленні про взаємну незалежність простору і часу та їх абсолютність. Такі уявлення при розгляді руху тіл з малими швидкостями є цілком виправданими, про що свідчить як успішний розвиток самої класичної механіки, так і всіх її інженерних застосувань. Достатньо нагадати блискучі успіхи в галузі космічної навігації.

Однак у другій половині XIX сторіччя після створення Максвеллом класичної електродинаміки й відкриття електромагнітних хвиль, до яких належить і світло, з'ясувалося, що електромагнітні процеси не задовольняють перетворення Галілея. Так у дослідах було неспростовно встановлено, що швидкість світла у вакуумі, всупереч перетворенням Галілея, не залежить від швидкості руху ані джерела, ані приймача. Ці та інші факти в решті решт призвели до радикального перегляду багатьох фундаментальних положень класичної механіки і створення А. Ейнштейном нової механіки - спеціальної теорії відносності.

Незалежність швидкості світла від руху джерела та приймача ("спостерігача") означає, що вона однакова в усіх системах відліку. Звідси випливає висновок, що не лише механічними, а й електромагнітними дослідами неможливо розрізнити рух або спокій будь-якої інерціальної системи відліку. На підставі цього А. Ейнштейн поширив принцип відносності на всі фізичні явища і прийняв його за один із постулатів (вихідних положень) нової механіки – ***принцип відносності Ейнштейна***:

> ***усі фізичні процеси за однакових умов протікають однаково в усіх інерціальних системах відліку; всі закони природи та рівняння, що їх описують, інваріантні, тобто не змінюються при переході від однієї інерціальної системи відліку до іншої.***

Іншими словами,

> ***всі інерціальні системи відліку є повністю рівноправними (еквівалентними): ніякими дослідами неможливо виділити жодну з них як головну.***

Другий постулат СТВ стосується швидкості світла. Доконаний факт її незалежності від руху джерела та приймача не випливає як наслідок з будь-яких логічних міркувань і має розглядатись як фундаментальний закон природи. Тому Ейнштейн прийняв його як другий постулат СТВ – ***принцип сталості швидкості світла***:

> ***швидкість світла у вакуумі не залежить від руху джерела світла і однакова у всіх напрямах.***

Це означає, що

> ***швидкість світла у вакуумі однакова в усіх системах відліку.***

Отже швидкість світла посідає особливе місце в природі. На відміну від інших швидкостей, скажімо, швидкості звуку, що змінюються при переході від однієї системи відліку до іншої, швидкість світла у вакуумі є інваріантною величиною.

Прийняті Ейнштейном постулати СТВ стоять на міцному фундаменті експерименту – на сьогодні немає жодного дослідного факту, який би їм суперечив. У той же час аналіз цих постулатів призводить до дуже несподіваних для класичної фізики висновків про відносність і взаємопов'язаність простору та часу.

Відносність часу

Як уже відмічалося, класична фізика трактує час як абсолютний, тобто єдиний і однаковий для всього Всесвіту. Тому видавалося самоочевидним, що проміжок часу між двома подіями є однаковим у всіх системах відліку. Зокрема, дві одночасні для даного спостерігача події є таким і для будь-якого іншого спостерігача (абсолютність одночасності). Але таке уявлення не сумісне з постулатами СТВ, у чому можна переконатися на наступному позірному експерименті. *Примітка. Позірним експериментом називають уявний дослід, який не суперечить законам фізики, але реально не може бути здійсненим. Позірний експеримент широко використовується у фізиці як один із способів доведення.*

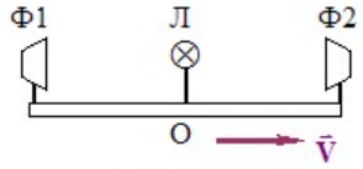


Рис. 8.2

Уявімо довгий стержень, посередині якого розташована лампочка Л, а на кінцях – фотоелементи Ф1 та Ф2 (рис. 8.2). Стержень рухається відносно спостерігача, що знаходиться в точці О. Нехай лампочка, коли порівняється з т. О, на мить спалахує (посилає світловий сигнал). В системі відліку, пов'язаній із стержнем, світловий сигнал доходить до фотоелементів за один і той самий час, оскільки він поширюється в обох напрямках з однаковою швидкістю й проходить однакову відстань. Тому в системі відліку стержня спрацьовування фотоелементів є двома одночасними подіями. Але нерухомий спостерігач в точці О зафіксує зовсім інше. Для нього швидкість світла також однакова в обох напрямках, оскільки вона не залежить від руху джерела. Але відстані, які пройде світло до фотоелементів, будуть різними, оскільки фотоелемент Ф1 рухається назустріч світловому імпульсу, а Ф2 "тікає" від нього. Отже, для нерухомого спостерігача фотоелемент Ф1 спрацює раніше, ніж Ф2. Таким чином, із постулатів СТВ незаперечно випливає висновок про ***відносність одночасності*** віддалених, тобто таких, які відбуваються в різних точках простору, подій:

дві події, одночасні в даній системі відліку, не є такими в інших системах відліку, що рухаються відносно даної.

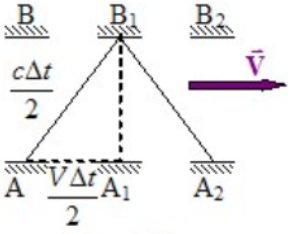


Рис. 8.3

Відносність одночасності означає відносність самого плину часу в різних системах відліку. Це можна побачити із ще одного позірного експерименту, запропонованого Ейнштейном (рис. 8.3). Нехай маємо вертикальний стержень АВ на кінцях якого закріплені два горизонтальні дзеркала, між якими „бігає" короткий світловий імпульс. Уявімо, що стержень рухається горизонтально зі швидкістю *V*, і визначимо час пробігу імпульсу від одного дзеркала до іншого й назад в *системі відліку стержня* К′ і в *нерухомій системі відліку* К, відносно якої стержень рухається з указаною швидкістю *V*. Очевидно, що в К′-системі цей час дорівнює

$$\Delta t_0 = \frac{2l_0}{c}, \qquad (8.4)$$

де l_0 = AB - відстань між дзеркалами в К′-системі відліку, в якій стержень є *нерухомим*, *c* – швидкість світла. Але в К-системі відліку, відносно якої стержень *рухається*, світловий імпульс відбивається від дзеркала В в точці B_1 і повертається до дзеркала А в точці A_2. Отже, за половину часу пробігу він проходить відстань l = AB_1, яка пов'язана з відстанями AA_1 та A_1B_1 очевидним співвідношенням $(AB_1)^2 = (AA_1)^2 + (A_1B_1)^2$. Так само очевидно, що $AA_1 = V\Delta t / 2$, де V - швидкість стержня, Δt – час пробігу світлового імпульсу від нижнього дзеркала до верхнього й назад у К-системі відліку. Що ж до відстані A_1B_1 у К-системі відліку, вона, як це випливає з принципу відносності, та сама, що і в К′ – l_0. Дійсно, аби це було не так, то, зробивши мітку від верхнього кінця стержня, що пролітає, на нерухомій осі AB К-системи і побачивши, що вона не співпадає з точкою В, ми б узнали, яка з двох систем відліку рухається, а яка перебуває в спокої. Але насправді встановити це принципово неможливо.

На основі сказаного можна записати:

$$l^2 = l_0^2 + \left(\frac{V\Delta t}{2}\right)^2. \qquad (8.5)$$

Швидкість світла в К-системі відліку, як і в К′, дорівнює c, отже $l = c\Delta t/2$. Підставивши цей вираз в (8.5), дістанемо:

$$\left(\frac{c\Delta t}{2}\right)^2 = l_0^2 + \left(\frac{V\Delta t}{2}\right)^2 \quad \Rightarrow \quad \Delta t = \frac{2l_0}{c\sqrt{1-(V/c)^2}}. \qquad (8.6)$$

Порівнюючи цей вираз із (8.4), знаходимо для двох ІСВ:

$$\Delta t = \frac{\Delta t_0}{\sqrt{1-(V/c)^2}}. \qquad (8.7)$$

Таким чином, одну подію (відбивання імпульсу від дзеркала А) від іншої події (повернення імпульсу на дзеркало А) в різних системах відліку відділяють різні проміжки часу. Це означає, що

час як такий є відносним: його плин у різних системах відліку є не однаковим.

Важливо чітко усвідомити фізичний зміст виразу (8.7). У К′-системі дзеркала нерухомі, отже, час Δt_0 виміряно нерухомим відносно подій годинником, а час Δt виміряно годинником К-системи відліку, відносно якої події рухаються зі швидкістю V. Як видно з (8.7), $\Delta t > \Delta t_0$, отже спостерігач К-системи відліку дійде висновку, що годинник К′-системи йде повільніше, ніж його власний. У цьому сенсі говорять про ***уповільнення часу в рухомих системах відліку***. При цьому задля однозначності та зручності висловлювань використовують таку термінологію. Систему відліку, в якій дані події відбуваються в нерухомих точках, називають ***власною системою відліку***, а час, який показує нерухомий відносно цієї системи відліку годинник ("власний" годинник) називають ***власним часом***. Будь-яка інша система відліку, відносно якої точки, де відбуваються події, рухаються, називається ***лабораторною системою відліку***, а проміжки часу між такими рухомими подіями, виміряні годинником лабораторної системи відліку, називаються ***лабораторним часом***. Виходячи з цього, можна сказати, що

власний час плине найповільніше

Варто нагадати, що тут слово "годинник" означає не якийсь фізичний прилад, а прийнятий принцип вимірювання проміжків часу. Тому розглянутий ефект уповільнення часу є чисто кінематичним і відображує об'єктивну властивість часу як такого, а саме, його ***відносність***. Це добре видно з такої ситуації. Якщо два спостерігачі в двох різних

ІСВ виконають описаний дослід з дзеркалами, то кожен з них абсолютно обґрунтовано буде вважати, що годинник його колеги з іншої системи відліку йде повільніше, ніж його власний. Тож запитання "який з годинників *насправді* йде повільніше?" позбавлене змісту.

Скорочення поздовжніх розмірів рухомих тіл, відносність простору

Розглянемо ще один позірний дослід. У системі відліку K′ паралельно до осі O′X′ розміщено нерухомий стержень AB довжини $l' = l_0$, який разом із K′-системою рухається із швидкістю V відносно нерухомої K-системи відліку (рис. 8.4). Поставимо завдання визначити довжину стержня AB в K-системі відліку. Спостерігач K-системи може зробити це, вимірявши час прольоту стержня $\Delta t_{ст}$ повз якусь фіксовану точку (наприклад, O) в K-системі відліку. Тоді

$$l = V\Delta t_{пр}. \qquad (8.8)$$

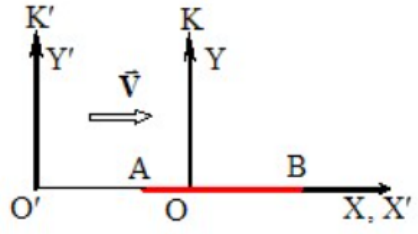


Рис. 8.4

Видається очевидним, що величина l має збігатися з $l_{0,}$ але насправді це не так. Уявімо, що описану процедуру вимірювання контролює й спостерігач K′-системи, для котрого *стержень є нерухомим*, а годинник K-системи рухається повз нього із швидкістю V. Отже, він скаже, що спостерігач K-системи вимірює час переміщення свого годинника $\Delta t_{год}$ на відстань l_0 між кінцями стержня. Тож

$$\Delta t_{\text{год}} = (l_0/V) \quad \Rightarrow \quad l_0 = V\Delta t_{\text{год}}. \qquad (8.8а)$$

А тепер звернімо увагу на те, що у виразах (8.8) і (8.8а) фігурують проміжки часу між тими самими двома подіями: суміщення годинника із одним (подія 1) та другим (подія 2) кінцем стержня. При цьому в K-системі (вираз (8.8)) події відбуваються в одній точці і фіксуються нерухомим відносно неї годинником. Отже $\Delta t_{ст}$ є власним проміжком часу між цими подіями:

$$\Delta t_{ст} = \Delta t_0.$$

Натомість величина $\Delta t_{год}$ вимірюється рухомим годинником, отже вона є лабораторним проміжком часу між подіями:

$$\Delta t_{год} = \Delta t.$$

Так само довжина l_0 нерухомого в K′-системі відліку стержня є його власною довжиною. Відповідно, довжина l стержня в K-системі, в якій він рухається, є лабораторною довжиною. Відтак із зіставлення виразів (8.8) і (8.8а) виходить:

$$\frac{l}{l_0} = \frac{\Delta t_0}{\Delta t},$$

і з урахуванням співвідношення (8.7)

$$l = l_0\sqrt{1-(V/c)^2}. \qquad (8.9)$$

Формула (8.9) виражає так зване ***лоренцове скорочення довжин*:** довжина рухомого стержня (*лабораторна* довжина), у напрямку руху в $\sqrt{1-(V/c)^2}$ раз менша, ніж його *власна* довжина. Отже

власні поздовжні розміри предмета є найбільшими.

Лоренцове скорочення, як і уповільнення часу, є кінематичним ефектом. Менше значення поздовжньої довжини рухомого стержня порівняно з нерухомим пояснюється не виникненням у ньому якихось механічних напруг і деформацій, а відносністю самого простору – в різних системах відліку просторові масштаби (величина одиниці довжини) не збігаються.

Слід іще раз наголосити, що в співвідношенні (8.9) величини l і l_0 - то є ***поздовжні*** розміри тіла (стержня), тобто, розміри в напрямку руху. Що ж до поперечних розмірів тіл, то вони не змінюються, позаяк в обох системах відліку відсутній рух тіла вздовж поперечних осей Y, Y′ та Z, Z′.

Розглянуті радикальні висновки, що випливають із постулатів Ейнштейна, ставлять два очевидних запитання: 1) чому класична механіка, що, як видно, стоїть на принципово хибних уявленнях про простір і час, дозволяє визначати рух тіл із винятково високою точністю, і 2) чи мають експериментальне підтвердження ефекти уповільнення часу (8.7) та скорочення масштабів (8.9)? Відповідь на перше запитання дає масштаб величини $c = 3\cdot10^8$ м/с. Для макроскопічних тіл, а класична механіка тільки такі й розглядає, величина ***c*** є недосяжною. Навіть найшвидші макроскопічні тіла – космічні тіла і ракети – рухаються із швидкостями не більшими, ніж ~ 10^4 м/с. При таких швидкостях $(V/c)^2 \sim 10^{-8}$, і формули (8.7) і (8.9) з величезною точністю дають $\Delta t = \Delta t_0$ і $l = l_0$. Тому відносність простору та часу й численні пов'язані з цим ***релятивістські*** ефекти стають відчутними лише при гранично високих швидкостях, які, крім світла, притаманні тільки елементарним частинкам високих енергій. Що ж до другого запитання, то відповідь ствердна, причому формула (8.7) підтверджена на тільки опосередковано, а й безпосередніми вимірами.

7.3. Перетворення Лоренца

Відносність простору та часу означає, що перетворення Галілея (1.30), (1.31), (1.32), які відображають концепцію абсолютного простору й абсолютного часу, принципово є помилковими. Треба мати перетворення координат і часу, які задовольняють і постулати Ейнштейна. Їх можна знайти різними способами, наприклад, так.

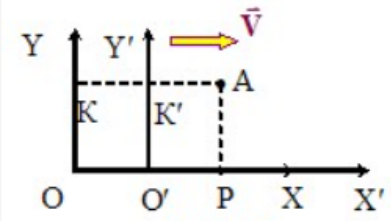


Рис. 8.5

Розглянемо дві "стандартні" інерціальні системи відліку K і K′, осі X та X′ яких співпадають, а осі Y, Y′ і Z, Z′ мають однакові напрямки (рис. 8.5). За початок відліку часу в обох системах приймаємо момент, коли початки координат O і O′ й відповідні координатні осі обох систем збігаються. Нехай в K-системі в якійсь точці A з координатами (x, y, z) в момент часу t відбулася певна подія, наприклад, спалахнула лампочка. Постає питання, які координати (x', y', z') має ця точка і в який момент часу t' відбувається ця подія для спостерігача K′-системи відліку, що рухається відносно K із швидкістю $\vec{V}$. Вище було показано, що при малих швидкостях перетворення Галілея (1.32) і (1.32а) є практично точними. Це дає підстави шукати нові перетворення координат x, x' у вигляді (дві інші координати в обох системах однакові):

$$x = \gamma(x' + Vt'), \tag{8.10}$$

$$x = \gamma(x' - Vt'), \tag{8.10a}$$

де γ – функціональний множник, що підлягає визначенню і має наближатися до одиниці при зменшенні швидкості V. В обох співвідношеннях фігурує одна й та сама величина γ, тим самим вони задовольняють принципу відносності (I постулат СТВ), адже обидві системи відліку є повністю рівноправними й відрізняються тільки напрямком відносного руху (знаком проекції $\vec{V}$). Через відносність часу перетворення (t, t') теж не можуть бути тривіальними, оскільки при $t = t'$ з перетворень (8.10) і (8.10а) виходить $\gamma = 1$, і вони переходять у перетворення Галілея. Загальний вигляд перетворень часу можна встановити, якщо підставити в (8.10а) замість координати x її вираз (8.10) і розв'язати отримане рівняння відносно t:

$$x' = \gamma(\gamma(x' + Vt') - Vt) \quad \Rightarrow \quad t = \gamma t' - \frac{(1-\gamma^2)x'}{V} \tag{8.11}$$

З урахуванням принципу відносності, зворотне перетворення має вигляд:

$$t' = \gamma t + \frac{(1-\gamma^2)x}{V}. \tag{8.11a}$$

У цьому можна переконатися, визначивши t' безпосередньо з (8.11).

Шукані перетворення мають задовольняти ще й умову сталості швидкості світла c (II постулат СТВ). Аби накласти цю умови розглянемо ще один позірний експеримент. Нехай в момент часу $t = t' = 0$, коли точки O і O′ збігалися, спалахнула лампочка, закріплена в початку координат якоїсь із даних двох систем відліку. Оскільки швидкість світла в обох системах дорівнює c, поширення світлового сигналу вздовж осей X, X′ описується рівнянням $x = ct$, і $x' = ct'$. Підставивши ці вирази в (8.10) і (8.10а), отримаємо: $ct = \gamma(c + V)t'$ і $ct' = \gamma(c - V)t$. Нарешті, перемноживши ліві та праві частини цих виразів, знайдемо, що

$$\gamma = \frac{1}{\sqrt{1-(V/c)^2}}, \text{ або } \gamma = \frac{1}{\sqrt{1-\beta^2}}, \text{ де } \beta = V/c. \qquad (8.12)$$

Підставляючи знайдену величину γ у (8.10), (8.10а) і (8.11), (8.11а) та враховуючи збіг поперечних координат, дістанемо перетворення координат і часу при переході від інерціальної К′-системи відліку до інерціальної К-системи відліку, які задовольняють постулати СТВ і називаються ***перетвореннями Лоренца***:

$$x = \frac{x' + Vt'}{\sqrt{1-(V/c)^2}} = \frac{x' + Vt'}{\sqrt{1-\beta^2}}, \quad y = y', \quad z = z';$$

$$t = \frac{t' + x'V/c^2}{\sqrt{1-(V/c)^2}} = \frac{t' + x'V/c^2}{\sqrt{1-\beta^2}}. \qquad (8.13)$$

Зворотні перетворення мають вигляд:

$$x' = \frac{x - Vt}{\sqrt{1-(V/c)^2}} = \frac{x - Vt}{\sqrt{1-\beta^2}}, \quad y' = y, \quad z' = z;$$

$$t' = \frac{t - xV/c^2}{\sqrt{1-(V/c)^2}} = \frac{t - xV/c^2}{\sqrt{1-\beta^2}}. \qquad (8.13а)$$

Одразу зауважимо таке. При малих швидкостях $\beta << 1$, перетворення Лоренца “автоматично” переходять у перетворення Галілея і вступають у силу закони класичної механіки. А при $\beta \geq 1$, ($V \geq c$) перетворення Лоренца втрачають фізичний зміст.

7.4. Перетворення швидкостей в СТВ, гранична швидкість

Установимо тепер ***формули перетворення швидкостей*** в СТВ, тобто, формули, котрі пов'язують між собою швидкості руху частинки в двох різних системах відліку - нерухомій К-системі та в К′-системі, що рухається відносно К із швидкістю V уздовж осі ОХ. Нехай частинка в К-системі відліку рухається із швидкістю $\vec{v}$. Її компоненти визначаються, як

$$v_x = \frac{\mathrm{d}x}{\mathrm{d}t} \qquad v_y = \frac{\mathrm{d}y}{\mathrm{d}t} \qquad v_z = \frac{\mathrm{d}z}{\mathrm{d}t} \tag{8.15}$$

Так само швидкість частинки в К′-системі $\vec{v}'$ визначається компонентами

$$v'_x = \frac{\mathrm{d}x'}{\mathrm{d}t}, \quad v'_y = \frac{\mathrm{d}y'}{\mathrm{d}t}, \quad v'_z = \frac{\mathrm{d}z'}{\mathrm{d}t} \tag{8.15а}$$

Аби знайти зв'язок між векторами $\vec{v}'$ і $\vec{v}$, виразимо переміщення й проміжок часу в (8.15а) через відповідні величини в К-системі відліку. Для цього продиференціюємо співвідношення (8.13)

$$\mathrm{d}x' = \frac{\mathrm{d}x - V\mathrm{d}t}{\sqrt{1-(V/c)^2}}, \quad \mathrm{d}y' = \mathrm{d}y, \quad \mathrm{d}z' = \mathrm{d}z, \quad \mathrm{d}t' = \frac{\mathrm{d}t - \frac{V}{c^2}\mathrm{d}x}{\sqrt{1-(V/c)^2}}.$$

Після підстановки цих виразів у (8.15а) і заміни $(V/c) = \beta$ матимаємо:

$$v'_x = \frac{\mathrm{d}x - V\mathrm{d}t}{\mathrm{d}t - \frac{V}{c^2}\mathrm{d}x} = \frac{\frac{\mathrm{d}x}{\mathrm{d}t} - V}{1 - \frac{V}{c^2}\cdot\frac{\mathrm{d}x}{\mathrm{d}t}},$$

$$v'_y = \frac{\mathrm{d}y\sqrt{1-\beta^2}}{\mathrm{d}t - \frac{V}{c^2}\mathrm{d}x} = \frac{\frac{\mathrm{d}y}{\mathrm{d}t}\cdot\sqrt{1-\beta^2}}{1 - \frac{V}{c^2}\cdot\frac{\mathrm{d}x}{\mathrm{d}t}}$$

$$v'_z = \frac{\mathrm{d}z\sqrt{1-\beta^2}}{\mathrm{d}t - \frac{V}{c^2}\mathrm{d}x} = \frac{\frac{\mathrm{d}z}{\mathrm{d}t}\cdot\sqrt{1-\beta^2}}{1 - \frac{V}{c^2}\cdot\frac{\mathrm{d}x}{\mathrm{d}t}}.$$

Урахувавши (8.15), отримаємо остаточно:

$$v'_x = \frac{v_x - V}{1 - \frac{v_x V}{c^2}}, \quad v'_y = \frac{v_y\sqrt{1-\beta^2}}{1 - \frac{Vv_x}{c^2}}, \quad v'_z = \frac{v_z\sqrt{1-\beta^2}}{1 - \frac{Vv_x}{c^2}}. \tag{8.16}$$

$$v_x = \frac{v'_x + V}{1 + \frac{Vv'_x}{c^2}}, \quad v_y = \frac{v'_y\sqrt{1-\beta^2}}{1 + \frac{Vv'_x}{c^2}}, \quad v_z = \frac{v'_z\sqrt{1-\beta^2}}{1 + \frac{Vv'_x}{c^2}}.. \tag{8.16а}$$

Ці формули виражають ***релятивістський закон перетворення швидкостей***. При малих швидкостях ($V << c$ і $v << c$) вони трансформуються в класичний закон додавання швидкостей (1.33) і (1.33а), виражений в координатній формі. В цьому зв'язку відмітимо, що y- і z- проекції швидкості частинки в одній системі відліку залежать не тільки від однойменних проекцій в іншій, а й від x-проекції швидкості. Тому для релятивістської частинки неможливо виразити зв'язок між швидкостями безпосередньо у векторній формі.

Гранична швидкість

Визначення за формулами (8.16), (8.16а) величини (модуля) та напрямку швидкості частинки в одній системі відліку через ці характеристики, задані в іншій системі відліку, потребує громіздких викладок. Тому при аналізі загальних наслідків з релятивістського закону перетворення швидкостей для більшої прозорості розглядають випадок руху частинки паралельно до осей X, X′ систем відліку. В такому разі поперечні компоненти швидкості дорівнюють нулю, і величина та напрям швидкості частинки в обох системах відліку визначаються тільки однією проекцією v_x або v'_x (у таких випадках у формулах перетворень швидкості часто опускають індекс проекції). Розглянемо приклади застосування формул (8.16) і (8.16а) за вказаних умов.

Приклад 1. Нехай спостерігач у рухомій K′-системі відліку вимірює швидкість c' світлового променя, що йде вздовж осі O′X′ від джерела, закріпленого в початку відліку O нерухомої K-системи відліку. В K-системі швидкість променя дорівнює c, тому, згідно з (8.16),

$$c' = \frac{c - V}{1 - \frac{Vc}{c^2}} = c$$

Отже, в обох системах відліку світловий промінь поширюється з однаковою швидкістю. Взагалі, в цьому немає нічого несподіваного, адже в самі перетворення Лоренца, з яких і випливають формули (8.16), вже закладено постулат сталості швидкості світла.

Приклад 2. Із космічної ракети, що віддаляється від Землі зі швидкістю 0,8c, випустили в напрямку руху зонд із швидкістю 0,6c відносно ракети. Визначимо швидкість зонда v відносно Землі. Для цього пов'яжемо із Землею нерухому K-систему відліку, а з ракетою – рухому K′-систему відліку так, щоб осі X, X′ мали напрям руху ракети та зонда. Тоді V = 0,8c і v' = 0,6c. Підставляємо ці значення у відповідну формулу (8.16а) і отримуємо:

$$v = \frac{0,6c + 0,8c}{1 + 0,48} \approx 0,95c.$$

Отже, виходить $v < c$, тоді як згідно з класичним законом (1.33а) і "здоровим глуздом" мало би вийти 1,4 c.

Приклад 3. Нехай дві елементарні частинки 1 і 2 рухаються назустріч одна одній паралельно осі OX K-системи відліку з однаковою швидкістю $v_1 = v_2 = nc$. Визначимо відносну швидкість частинок v, наприклад, швидкість руху другої відносно першої. Для цього перейдемо в K′-систему відліку, котра рухається відносно K із швидкістю першої частинки $V = v_1$. Тоді перша частинка є нерухомою, і швидкість другої частинки в цій системі відліку є шуканою її швидкістю відносно першої: $v'_2 = v$. Підставляючи значення проекцій швидкостей в (8.16), знаходимо:

$$v = v'_{2x} = \frac{-nc - nc}{1 - \frac{nc \cdot (-nc)}{c^2}} = -\frac{2n}{1 + n^2}c = -\frac{2c}{\frac{1}{n} + n}\,. \qquad (8.17)$$

Знак "–" відображає те, що частинка 2 рухається назустріч частинці 1. Із цього виразу випливає важливий висновок про те, що за будь-яких умов $v \le c$. У цьому переконує формальний аналіз виразу (8.17). Справді, при малих швидкостях ($n << 1$) модуль відносної швидкості $v \approx 2nc$ зростає при збільшенні n. Але при $n >> 1$ $v \approx 2c/n$ і зменшується при зростанні n. Отже, при деякому значенні $n = n_m$ відносна швидкість має максимум $v = v_{max}$. Значення n_m і величину v_{max} не важко знайти методами вищої математики. А саме, треба взяти похідну від виразу в знаменнику (8.17) і прирівняти її до нуля:

$$\left(\frac{1}{n}+n\right)' = -\frac{1}{n^2}+1 \quad \Rightarrow \quad \frac{1}{n_m}=1 \quad n_m = 1.$$

Підставляємо це значення в (8.17) і отримуємо для модуля $v_{\max} = c$.

Таким чином, максимальна можлива швидкість руху однієї частинки (одного тіла) відносно іншої частинки (іншого тіла) дорівнює величині c. Але ж саме поняття руху визначене (тобто, має зміст) тільки по відношенню до якогось іншого тіла (тіла відліку). Тому за будь-яких умов швидкість руху будь-якого тіла

$$v \le c. \qquad (8.18)$$

Інакше говорячи, в природі діє фундаментальний ***закон граничності швидкості c***, який гласить:

> ***швидкість будь-якого матеріального об'єкта ні за яких умов не може бути більшою, ніж величина c = 3·10⁸ м/с .***

На цьому слід спеціально наголосити: йдеться про рух матеріальних частинок або поширення випромінювання, тобто, про процеси, що супроводжуються перенесенням енергії і, тим самим, можуть бути використані для передачі сигналів (інформації). На інші швидкості це обмеження не поширюється

Те, що швидкість матеріальних об'єктів не може бути більшою за $\approx 3 \cdot 10^8$ м/с, зокрема, пояснює інваріантність швидкості світла у вакуумі: оскільки вона є граничною, то, зрозуміло, не може залежати від системи відліку.

7.5. Інтервал

На основі виразів (8.14) і (8.14а) можна зробити ще один важливий висновок: проміжок часу між двома подіями в одній системі відліку залежить не тільки від відповідного проміжку в іншій системі відліку, а й від відстані між точками, в яких відбулися події. Це свідчить про те, що простір і час є не тільки відносними, а й *органічно взаємопов'язаними.* Тому слід говорити не про "простір і час", а про єдиний "*простір-час*". Про це свідчить існування кінематичної величини, яка визначається і координатами, і часом, та є ***інваріантом перетворень Лоренца***, тобто, не змінюється при переході від однієї ІСВ до іншої. Ця величина називається *просторово-часовим інтервалом s* між двома подіями (коротко - ***інтервалом***) і задається співвідношенням

$$s_{12}^2 = c^2\Delta t_{12}^2 - l_{12}^2 = c^2(t_1 - t_2)^2 - \left((x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2\right), \quad (8.19)$$

де (t_1, x_1, y_1, z_1) і (t_2, x_2, y_2, z_2) - просторово-часові координати подій.

Інваріантність інтервалу, тобто, що $s_{12}^2 = s'^2_{12}$, можна довести прямими обчисленнями після заміни у виразі інтервалу нештрихованих величин штрихованими на основі (8.14). Інваріантні величини позначають як "inv", отже можна записати:

$$s_{12}^2 = c^2\Delta t_{12}^2 - l_{12}^2 = \text{inv}.$$

Таким чином, твердження, що дві події розділені заданим інтервалом має абсолютний характер – воно справедливе для будь-якої інерціальної системи відліку.

Інваріантні величини в теорії відносності мають велике теоретичне та практичне значення. Зауважимо також, що, крім інтервалу, *кінематичними інваріантами* перетворень Лоренца є *гранична швидкість c*, а також *власний проміжок часу* Δt_o *і власна довжина* l_o.

В залежності від того, яка складова інтервалу переважає, інтервали поділяють на *простороподібні* ($l_{12} > c\Delta t_{12}$), *часоподібні* ($c\Delta t_{12} > l_{12}$) та *світлоподібні* ($l_{12} = c\Delta t_{12}$).

Якщо інтервал між двома подіями є простороподібним, то завжди можна знайти таку систему відліку K′, в якій обидві події відбуваються одночасно ($\Delta t_{12} = 0$):

$$c^2\Delta t_{12}^2 - l_{12}^2 = l'^2_{12}.$$

Якщо ж інтервал часоподібний, то завжди можна знайти таку систему відліку K', в якій обидві події відбуваються в одній точці ($l'_{12} = 0$):

$$c^2\Delta t_{12}^2 - l_{12}^2 = c^2\Delta t'^2_{12}$$

У випадку простровоподібних інтервалів, коли $l_{12} > c\Delta t_{12}$, ні в якій системі відліку події не можуть впливати одна на одну, навіть якщо зв'язок між ними здійснювався з граничною швидкістю *c*. На відміну від цього, у випадку часоподібних або світлоподібних інтервалів, коли $c\Delta t_{12} \geq l_{12}$, події можуть бути пов'язані одна з одною причинно-наслідковим зв'язком.

7.6. Контрольні запитання

1. Що спонукало створення спеціальної теорії відносності?
2. Сформулюйте постулати Ейнштейна.
3. У чому полягає відміна між принципом відносності Ейнштейна та принципом відносності Галілея?
4. Які висновки про властивості простору й часу випливають із постулатів Ейнштейна?
5. На підставі яких міркувань можна твердити, що поперечні координати точки в різних системах відліку однакові?
6. Що таке власна та лабораторна довжина стержня? Як вони пов'язані між собою?
7. Що таке власний та лабораторний проміжок часу? Як вони пов'язані між собою?
8. На основі чого можна зробити висновок про те, що простір і час не є незалежними одне від одного ?
9. Чому говорять, що скорочення поздовжньої довжини тіла та скорочення проміжку часу між подіями є взаємними ефектами?
10. Що таке релятивістський інтервал між подіями? Які бувають типи інтервалів?
11. При якому типі інтервалу між подіями не може бути причинно-наслідкового зв'язку?
12. Згідно з перетвореннями Лоренца, поперечні координати (y; z) рухомої точки в різних ІСВ однакові. Чому ж тоді не однакові поперечні проекції швидкості (v_y, v_z)?
13. Який зв'язок існує між формулами СТВ і відповідними формулами класичної механіки?
14. У чому полягає закон про граничну швидкість?
15. Що ми спостерігали б навколо себе, якби гранична швидкість була рівною с = 50 км/год?
16. Чи існували би релятивістські ефекти за відсутності в природі будь-яких обмежень на величину швидкості?
17. Виробник твердить, що в його телевізорах електронний промінь "біжить" по екрану кінескопа із швидкістю $5 \cdot 10^8$ м/с. Чи таке можливо?
18. Із молекулярної фізики відомо, що температура є прямо пропорційна середній кінетичній енергії теплового руху молекул речовини. А в СТВ показано, що існує верхня межа швидкості – c. Чи означає це, що існує й верхня межа можливих значень температури?

8. СПЕЦІАЛЬНА ТЕОРІЯ ВІДНОСНОСТІ. ДИНАМІКА

Зміна в СТВ поглядів практично на всі основні поняття й величини кінематики вимагає аналізу та ревізії й основних понять і положень класичної динаміки. Це, зокрема стосується поняття імпульсу, зв'язку між силою та прискоренням, поняття енергії, тощо.

Нижче мова піде про таке:

8.1. Релятивістський імпульс і релятивістська маса

8.2. Основне рівняння релятивістської динаміки

8.3. Енергія та маса

8.4. Контрольні запитання

8.1. Релятивістський імпульс і релятивістська маса

Одна з основних величин динаміки – мпульс частинки – в класичній механіці визначається як добуток маси на швидкість (3.2). При цьому маса виражає внутрішню властивість – інертність частинки і є інваріантною (незалежною від системи відліку) величиною. Велике значення поняття імпульсу в теорії зумовлене тим, що для нього виконується універсальний закон збереження (5.2). Цей закон трактує про поведінку імпульсу замкненої сукупності тіл всередині даної інерціальної системи відліку й ніяк не пов'язаний із переходом від однієї системи відліку до іншої. Тому він виконується і в релятивістській механіці. Однак, як показує аналіз, імпульс більш складно, ніж уважалося, залежить від швидкості й визначається формулою:

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1-(V/c)^2}}. \qquad (9.1)$$

Отже, імпульс нелінійно залежить від швидкості частинки, але, як видно з (9.1), це стає помітним лише при релятивістських (сумірних із c) швидкостях. Тому величину (9.1) називають ***релятивістським імпульсом***. При малих швидкостях цей вираз переходить у класичну формулу (3.2).

Інколи буває зручно записувати вираз релятивістського імпульсу згорнуто у вигляді

$$\vec{p} = m_r\vec{v}, \qquad (9.1a)$$

де величина

$$m_r = \frac{m}{\sqrt{1-(V/c)^2}} \qquad (9.2)$$

називається ***релятивістською масою*** частинки. В такому контексті власну масу частинки m називають ***масою спокою***. Слід одразу зазначити, що поняття релятивістської маси, яке збереглося з початкового періоду становлення СТВ, є формальним і не має визначеного фізичного змісту. Зокрема, величину m_r не можна розглядати як міру інертності рухомого тіла.

8.2. Основне рівняння релятивістської динаміки

Із виразу релятивістського імпульсу (9.1) зрозуміло, що $\mathrm{d}\vec{p}/\mathrm{d}y \neq m\vec{a}$, отже, для релятивістської частинки другий закон Ньютона у формі $m\vec{a} = \vec{F}$ не виконується. Але, як доведено, основне рівняння динаміки в більш загальній формі (3.3) зберігає чинність і в СТВ:

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = \vec{F}. \qquad (9.3)$$

Само собою, в ***основному рівнянні релятивістської динаміки*** (9.3) фігурує релятивістський імпульс. Розгорнуто воно записується так:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m_0\vec{v}}{\sqrt{1-(v/c)^2}}\right) = \vec{F}. \qquad (9.3а)$$

Рівняння (9.3) можна подати й через релятивістську масу:

$$\frac{\mathrm{d}(m_r\vec{v})}{\mathrm{d}t}\vec{F}. \qquad (9.3б)$$

Рівняння (9.3) і (9.3б), начебто, такі самі, як і в класичній механіці, але це лише видимість. Насправді для релятивістської частинки немає прямого зв'язку між величиною сили та прискоренням. В цьому легко переконатись, якщо виконати диференціювання в (9.3б), взявши до уваги, що m_r теж залежить від часу:

$$m_r\vec{a} + \vec{v}\frac{\mathrm{d}m_r}{\mathrm{d}t} = \vec{F}, \qquad (9.4)$$

де $\vec{a} = \dfrac{\mathrm{d}\vec{v}}{\mathrm{d}t}$ – прискорення частинки. При цьому, позаяк другий доданок у лівій частині рівняння (9.4) за напрямом не збігається з $\vec{F}$ (див. рис 9.1), то й напрям прискорення частинки $\vec{a}$ теж не збігається з напрямом сили, що діє на неї. Є лише два винятки з такої неординарної поведінки релятивістської частинки: 1) діюча на частинку сила поперечна $\vec{F}\perp\vec{v}$ і 2) сила поздовжня $\vec{F} \parallel \vec{v}$.

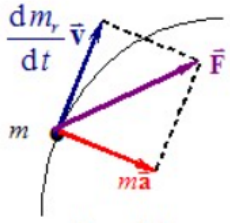


Рис. 9.1

У першому випадку сила не виконує роботи, отже, рух є рівномірним і m_r = const. Тому

$$m_r\vec{a} = \vec{F} \quad \Rightarrow \quad \frac{m\vec{a}}{\sqrt{1-(v/c)^2}} = \vec{F} \qquad (9.4а)$$

У другому випадку рух частинки є прямолінійним і, як можна показати, виконавши диференціювання в (9.3а),

$$\frac{m\vec{a}}{\left(1-(v/c)^2\right)^{3/2}} = \vec{F} \qquad (9.4б)$$

Отже, в цих двох випадках прискорення частинки має напрям діючої на неї сили, а рівняння (9.4а) і (9.4б) нагадують рівняння другого закону Ньютона (3.5). Тому в них множники при прискоренні іноді називають поперечною та поздовжньою масою, відповідно. В такому сенсі можна говорити, що сила дорівнює добутку маси на прискорення. Проте це не означає, що виконується другий закон Ньютона, оскільки прискорення ***не є*** *прямо пропорційним* діючій на частинку силі. Зокрема, з рівняння (9.4б) видно, що при дії сталої сили прискорення частинки, через зростання швидкості, буде невпинно зменшуватися, простуючи до нуля при наближенні швидкості до значення $v = c$. Отже, з основного рівняння релятивістської динаміки теж випливає висновок про граничний характер швидкості c, який раніше було отримано в кінематиці СТВ (див. Лекція 8, п. 8.4).

Наостанок відмітимо, що при малих швидкостях руху релятивістські рівняння (9.3а), (9.4а) і (9.4б), як завжди, автоматично переходять у відповідне класичне рівняння (3.5).

8.3. Енергія та маса

Усі розглянуті досі результати СТВ стосуються понять і величин, які відомі й у класичній механіці. Але, створюючи СТВ, Ейнштейн відкрив і знаменитий закон взаємозв'язку між енергією та масою, котрому немає аналогів у ньютонівській механіці. Про існування такого зв'язку можна здогадатись і встановити відповідні співвідношення, аналізуючи зв'язок між механічною енергією та роботою сил для релятивістської частинки.

Кінетична енергія релятивістської частинки

Одне з основних положень механіки – теорема про кінетичну енергію (Л1-6, п. 6.2), що встановлює зв'язок між нею та роботою сил (6.8), (6.8а), зберігає чинність і в релятивістській механіці. Але кінетична енергія визначається іншою, ніж у класичній механіці, формулою. Для її встановлення розглянемо роботу, яку виконує якась сила $\vec{F}$ над релятивістською частинкою на переміщенні $\mathrm{d}\vec{r}$ (рис 9.2). Згідно з (6.1) і (1.3)

$$\delta A = \vec{F} d\vec{r} = \vec{F}\vec{v}dt$$

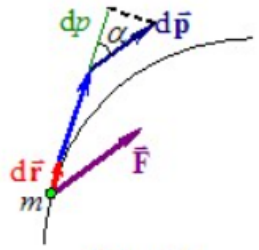


Рис. 9.2

Відповідно до (9.3б), $\vec{F}dt = d(m_r\vec{v}) = m_r d\vec{v} + \vec{v}dm_r$, де m_r – релятивістська маса. Тому

$$\delta A = \vec{v}\left(\vec{v}dm_r + m_r d\vec{v}\right) = v^2 dm_r + m_r v dv \qquad (9.5)$$

де враховано, що $\vec{v}^2 = v^2$ і $\delta A = \vec{v}\left(\vec{v}dm_r + m_r d\vec{v}\right) = v^2 dm_r + m_r v dv$, (див. рис 9.2).

Для визначення $\mathrm{d}m_r$ спочатку перетворимо (9.2):

$$m_r^2 = \frac{m^2}{1-(v/c)^2} \quad \Rightarrow \quad m_r^2 - m_r^2 v^2/c^2 \quad \Rightarrow \quad m_r^2 c^2 - m_r^2 v^2 = m^2 c^2$$

Тепер продиференцюємо отриманий вираз:

$$2m_r c^2 \mathrm{d}m_r - 2m_r v^2 \mathrm{m_r} - 2m_r^2 v \mathrm{d}v = 0,$$

звідки, після ділення на $2m$, одразу одержимо вираз усієї правої частини (9.5):

$$v^2 \mathrm{d}m_r + m_r v \mathrm{d}v = c^2 \mathrm{d}m_r. \qquad (9.6)$$

Отже, з урахуванням (6.8), можна записати:

$$\mathrm{d}K = \delta A = c^2 \mathrm{d}m_r. \qquad (9.7)$$

Звідси випливає, що приріст кінетичної енергії частинки зумовлений приростом її релятивістської маси. Якщо частинка перебуває у спокої, то K = 0 і m_r = m. Тому, інтегруючи (9.7) у границях від 0 до K і від m до m_r, знаходимо:

$$\int_0^K \mathrm{d}K = \int_m^{m_r} c^2 \mathrm{d}m \quad \Rightarrow \quad K = c^2(m_r - m). \tag{9.7a}$$

Нарешті, врахувавши (9.2), одержимо формулу кінетичної енергії релятивістської частинки:

$$K = mc^2 \left(\frac{1}{\sqrt{1-(v/c)^2}} - 1 \right). \tag{9.8}$$

При малих швидкостях ($v << c$) ця формула теж переходить у формулу кінетичної енергії ньютонівської механіки (6.7). Щоб у цьому переконатися, розкладемо перший доданок (9.8) у степеневий ряд за формулою бінома Ньютона:

$$\frac{1}{\sqrt{1-(v/c)^2}} = \left(1+(v/c)^2\right)^{-\frac{1}{2}} = 1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \ldots.$$

При $v \ll c$ можна обмежитися першими двома доданками, і тоді (9.8) трансформується в класичну формулу

$$K = \frac{mv^2}{2}.$$

Зауважимо, що останній вираз є чинним тільки при малих швидкостях. Для швидких рухів він не придатний, навіть, якщо в нього замість маси спокою підставити релятивістську масу частинки.

Взаємозв'язок енергії і маси

Якщо розкрити дужки в (9.7а), то кінетична енергія релятивістської частинки постане у вигляді різниці двох величин:

$$K = m_r c^2 - mc^2. \tag{9.9}$$

Перша з них

$$E = m_r c^2 \tag{9.10}$$

стосується частинки, що рухається, і називається ***повною (релятивістською) енергією***. Друга

$$E_0 = mc^2 \tag{9.11}$$

відноситься до тієї ж частинки в стані спокою і, відповідно, називається ***енергією спокою***.

За допомогою цих понять кінетична енергія (енергія руху) релятивістської частинки відповідно до (9.9) виражається дуже природньо: кінетична енергія – то є різниця енергії частинки в стані руху E та в стані спокою E_0:

$$K = E - E_0. \tag{9.12}$$

У зв'язку з цим виникає питання про природу енергії спокою, адже можна подумати, що то є, подібно до релятивістської маси, лише зручна формальна величина. Але Ейнштейн дійшов висновку, що енергія спокою є реальною фізичною величиною, що визначає узагальнену внутрішня енергію частинки (тіла). Якщо частинка є складною, приміром, як ядро атома, то енергія спокою включає енергію взаємодії складових

частинок, але не тільки! Згідно з Ейнштейном, будь-яка елементарна частинка, котра “ні з чого не складається”, як от електрон, теж має відповідну енергію спокою. Реальність енергії спокою доводиться тим, що вона здатна перетворюватися на інші види енергії. При цьому будь-яка зміна енергії спокою супроводжується еквівалентною зміною маси спокою частинки:

$$\Delta m = \frac{\Delta E_0}{c^2}. \qquad (9.13)$$

Наприклад, при окисленні вуглецю (горінні) виділяється енергія, тож енергія молекули вуглекислого газу менша за сумарну енергію частинок, які беруть участь у реакції. Тому й маса спокою молекули вуглекислоти повинна бути менша, ніж сума мас спокою атома вуглецю та молекули кисню. Але в хімічних реакціях і взаємодії макроскопічних тіл зміна маси спокою настільки мізерна (див. Задача 9.8), що її неможливо виявити. Інша справа ядерні реакції, в яких взаємодіють ядра атомів або елементарні частинки надвисоких енергій. В таких процесах виділяються й поглинаються настільки великі енергії, що зміну маси спокою можна не тільки виявити, а й визначити на досліді. Зокрема, встановлено, що маса ядра будь-якого хімічного елемента менша за сумарну масу складових нуклонів у згоді з (9.13). Більше за те, завдяки цьому отримують промислову енергію на атомних електростанціях. Таким чином, формула (9.11) виражає один з найбільш фундаментальних законів природи – ***взаємозв’язок маси та енергії***. Його глибинний зміст полягає в тому, що енергія є невід’ємною характеристикою матерії – будь-яка частинка чи тіло має відповідний (і дуже великий) запас енергії завдяки самому факту свого існування.

Зв’язок між енергією та імпульсом

Імпульс (9.1) і повна енергія (9.10) релятивістської частинки залежать від швидкості, отже й від системи відліку. Однак існує певна комбінація цих величин, яка є релятивістським інваріантом, тобто не залежить від системи відліку. З формули (9.2) маємо

$$m_r^2 = \frac{m^2}{1-(v^2/c^2)} \quad \Rightarrow \quad m_r^2 c^2 = m_r^2 v^2 = m^2 c^2.$$

Домноживши цей вираз на c^2 і врахувавши (9.10) і (9.1а), одержимо

$$E^2 - p^2c^2 = m^2c^4. \qquad (9.14)$$

Оскільки в правій частині фігурують тільки інваріантні величини m і c, величина, що стоїть у лівій частині, теж є релятивістським інваріантом, тобто, не залежить від системи відліку. Отже,

$$E^2 - p^2c^2 = \text{inv}. \qquad (9.15)$$

Запишемо ще деякі корисні співвідношення. По-перше, вираз (9.14) зручно переписати у вигляді:

$$E = c\sqrt{p^2 + m^2c^2}. \qquad (9.16)$$

Окрім того, безпосередньо з (9.1а) і (9.10) випливає співвідношення:

$$\vec{p} = \frac{E\vec{v}}{c^2}. \qquad (9.17)$$

Нарешті, з (9.14) і (9.12) можна отримати зв’язок між імпульсом і кінетичною енергією релятивістської частинки:

$$pc = \sqrt{E^2 - m^2c^4} = \sqrt{\left(K + mc^2\right)^2 - m^2c^4} = \sqrt{K\left(K + 2mc^2\right)}$$

звідки

$$p = \frac{1}{c}\sqrt{K\left(K + 2mc^2\right)}. \qquad (9.18)$$

Зауважимо, що, відповідно до цього виразу, в ядерній фізиці та фізиці елементарних частинок імпульс прийнято вимірювати в одиницях енергії (МеВ), поділених на швидкість світла c, тобто – в МеВ/c.

Безмасові частинки

Сучасною наукою незаперечно встановлено, що в природі існують так звані ***безмасові частинки*** - частинки, які не мають маси спокою. Такими є фотони - елементарні частинки, що переносять електромагнітне випромінювання, зокрема, світло. З деякою імовірністю безмасовими є й різні види нейтрино – частинки, що народжуються в певних ядерних реакціях.

Для безмасової частинки m = 0, і з (9.16) випливає, що

$$E = pc \text{ або ж } p = \frac{E}{c}. \qquad (9.19)$$

Ця формула узгоджується з (9.17) тільки за умови, що $v = c$. Отже, єдиним можливим станом безмасової частинки є рух із граничною швидкістю. Тому, стикаючись з іншими частинками, безмасові частинки або відбиваються без зміни величини швидкості, або поглинаються, тобто, припиняють своє існування. При цьому вони тілу, з яким стикаються, передають імпульс і, згідно з (9.3), створюють відповідну силу тиску. Тиск світла вперше виміряв учений Лебедєв. Отримані ним у дослідах значення тиску світла узгоджуються з теоретичними величинами, розрахованими на основі механізму зіткнень фотонів з частинками опромінюваного тіла. Це є одним із експериментальних підтверджень розглянутої концепції безмасових частинок і їхніх властивостей.

8.4. Контрольні запитання

1. Частинка рухається із швидкістю υ = 0,6 c. У скільки разів її релятивістська маса більша за масу спокою?

2. З якою швидкістю рухається частинка, якщо її релятивістська маса в три рази більша за масу спокою?

3. На скільки відсотків релятивістська маса частинки більша за масу спокою при швидкості 30 Мм/с?

4. Релятивістська частинка має імпульс mc (m — маса спокою). Знайти швидкість частинки.

5. Знайти швидкість частинки, при якій її релятивістський імпульс у η = 2 рази перевищує ньютонівський імпульс.

6. Повна енергія тіла зросла на 1 Дж На скільки при цьому змінилася маса тіла?

7. На скільки повинна збільшитися повна енергія тіла, щоб його релятивістська маса зросла на 1 г?

8. Швидкість електрона υ = 0,6 с. Знайти кінетичну енергію електрона в МеВ. Енергію спокою прийняти 0,5 МеВ.

9. При якій швидкості кінетична енергія будь-якої частинки дорівнює її енергії спокою?

10. Визначити швидкість електрона, якщо його кінетична енергія T = 2,04 МеВ. Енергія спокою електрона E_0 = 0,51 МеВ .

11. Показати, що релятивістський вираз кінетичної енергії $T = (m_r - m)c^2$ при $V << c$ переходить у відповідний вираз класичної механіки.

12. Показати, що вираз релятивістського імпульсу через кінетичну енергію $p = \frac{1}{c}\sqrt{T(2E_0 + T)}$ при $V << c$ переходить у відповідний вираз класичної механіки .

МОЛЕКУЛЯРНА ФІЗИКА І ТЕРМОДИНАМІКА

ЗМІСТ

ВСТУП

I. МОЛЕКУЛЯРНА ФІЗИКА

1. Основні положення МКТ
2. Властивості газів
3. Розподіл Максвелла-Болгьцмана

II. ЕЛЕМЕНТИ ТЕРМОДИНАМІКИ

4. Перше начало термодинаміки

5. Теплоємність. Політропні (політропічні) процеси

6. Явища переносу

7. Друге начало термодинаміки

8. Ентропія

ВСТУП

Розділ «Молекулярна фізика і термодинаміка» є другою частиною курсу загальної фізики за програмою підготовки бакалаврів у якому вивчаються фізичні властивості речовин у зв'язку з їх внутрішньою будовою та поведінкою систем, які складаються з величезної кількості частинок.

У молекулярній фізиці вивчають властивості речовини та явища, що зумовлені її внутрішньою будовою. Загальновідомо, що речовина складається з мікроскопічних частинок - молекул і атомів,- які невпинно рухаються та взаємодіють між собою. Кількість молекул, навіть у мізерній крупинці речовини, є дуже великою (наприклад, навіть порошинка розміром 1 мкм, яку можна побачити тільки в потужний мікроскоп, містить $\sim 10^{12}$ молекул), тобто, вона є ***макроскопічною системою***. Через це та через специфічні квантові властивості атомів і молекул, досліджувати властивості речовини, розглядаючи рух окремих частинок на основі законів механіки, принципово неможливо. Але в макроскопічних системах виявляються якісно нові закономірності, котрі досліджують ***статистичним*** та ***термодинамічним методами***. Перший використовується в ***молекулярно-кінетичній теорії*** (МКТ), де характеристики й поведінка макроскопічної системи встановлюються через усереднені характеристики руху окремих молекул і їхньої взаємодії між собою. Математичною основою МКТ є теорія ймовірностей та математична статистика. Термодинамічний метод, або просто - ***термодинаміка,*** ґрунтується на декількох основних законах, або ***началах термодинаміки***, які встановлені з досліду. Закони термодинаміки є універсальними в тому сенсі, що вони є чинні для будь-якої макроскопічної системи незалежно від її внутрішньої будови та індивідуальних властивостей частинок. Через це термодинамічний метод є незамінним, коли внутрішня будова системи та взаємодія між частинками є дуже складною чи недостатньо вивченою. З іншого боку, загальні принципи та положення МКТ дозволяють зрозуміти походження законів термодинаміки та механізми термодинамічних процесів. Отже, статистичний та термодинамічний методи органічно є взаємопов'язані та доповнюють один одного.

Властивості тіл і процеси в них ми спостерігаємо і вивчаємо через величини, які характеризують систему як ціле, і які можна вимірювати приладами. Такі величини називаються ***макроскопічними*** або ***термодинамічними***. Для будь-якої системи існує відповідний набір ***параметрів стану*** – термодинамічних величин, які визначають макроскопічний (термодинамічний) стан системи, тобто стан системи як цілого. Наприклад, основними параметрами стану газу є об'єм, тиск і температура. Наголосимо ще раз, що макроскопічні (термодинамічні) величини для окремої молекули позбавлені змісту, вони мають сенс тільки для всієї системи чи якоїсь її макроскопічної частини. Сукупність макроскопічних параметрів системи визначає її ***макроскопічний*** (***термодинамічний***) ***стан***. При цьому, якщо всі параметри мають однакові значення по всій системі, її стан називається ***рівноважним***, якщо ні, то – ***нерівноважним***.

Очевидно, що макроскопічний стан системи визначається рухом та взаємодією молекул – мікроскопічних частинок з яких складається система. Відповідно, величини, що характеризують окремі частинки системи називаються ***мікроскопічними параметрами***, а сукупність мікроскопічних параметрів усіх молекул визначає ***мікроскопічний стан*** системи. Основним завданням МКТ якраз і є встановлення зв'язку між мікроскопічними та макроскопічними (термодинамічними) станами в різних рівноважних молекулярних системах.

I. МОЛЕКУЛЯРНА ФІЗИКА

У цьому розділі наступні питання

1. Основні положення МКТ

2. Властивості газів

3. Розподіл Максвулла-Больцмана

1. Основні положення МКТ

У наш час дискретна молекулярна будова речовини не потребує доказів. Достатньо лише вказати, що сучасна техніка дозволяє отримувати зображення не тільки окремих молекул, а й атомів. Нагадаємо деякі основні відомості про характеристики молекул та їхній рух і взаємодію.

1. Характеристики молекул

2. Рух молекул. Теплова рівновага та температура

3. Взаємодія молекул

Контрольні запитання

1. Характеристики молекул

Маса молекул. Сучасні методи дозволяють досить точно визначати масу молекул і атомів, попри їхню гранично малу величину. Наприклад, маса атома Гідрогену складає $1{,}673\cdot10^{-27}$ кг, атома Карбону – $1{,}994\cdot10^{-26}$ кг, молекули кисню – $5{,}310\cdot10^{-26}$ кг, тощо. Такі числа є незручними, тому в молекулярній фізиці та хімії користуються відносними ***атомними одиницями маси*** (а.о.м.). За означенням 1 а.о.м. = 1/12 маси атома найпоширенішого в природі ізотопу[1] Карбону ^{12}C, тобто

$$1 \text{ а.о.м.} = 1{,}660\cdot10^{-27} \text{ кг}.$$

Маса молекули (атома), виражена в а.о.м., називається відносною молекулярною (атомною) масою M_r. Атомні маси хімічних елементів указані в періодичній таблиці Менделєєва. Для наведених вище речовин вони складають, відповідно, 1,008; 12,011 та 31,988[2]. Атомні та молекулярні маси для більшості речовин близькі до цілих чисел і в практичних розрахунках заокруглюються.

Кількість речовини. Молярна маса. Деякі характеристики речовини та процеси визначаються не масою, а кількістю частинок. Це, зокрема, стосується тиску газу або хімічних реакцій. Тому в молекулярній фізиці та хімії поряд з масою широко використовують поняття ***кількості речовини***. Мірою кількості речовини є кількість структурних частинок (атомів, молекул, іонів, тощо), яка в ній міститься. Але через величезну кількість таких частинок для вимірювання кількості речовини використовують, знов таки, відносну одиницю - ***моль***. За означенням 1 моль - це кількість речовини, в якій міститься стільки ж частинок, як у 0,012 кг ізотопу Карбону ^{12}C. Така кількість частинок називається сталою Авогадро й дорівнює

$$N_A = 6{,}02\cdot10^{23} \text{ 1/моль} = 6{,}02\cdot10^{26} \text{ 1/кмоль}.$$

Кількість речовини ν виражається через кількість частинок N очевидним співвідношенням:

$$\nu = \frac{N}{N_A} \qquad (1.1)$$

або, через масу m,

$$\nu = \frac{m}{M}, \qquad (1.1а)$$

де M – маса одного моля речовини, яка називається ***молярною масою*** і вимірюється в кг/моль. Зрозуміло, що

$$M = m_0 N_A, \qquad (1.1б)$$

де m_0 – маса одного атома (молекули).

Між молярною масою речовини M та відносною масою її молекул M_r існує простий числовий зв'язок:

$$M(\text{кг/моль}) = 10^{-3} \cdot M_r \quad \text{або} \quad M(\text{кг/кмоль}) = M_r.$$

Розміри молекул. Уявлення про розміри атомів і молекул можна отримати з таких простих міркувань. Вода (H_2O) масою 1 г, що складає 1/18 моля (M = 18 г/моль), займає об'єм $V = 1$ см^3 і містить $N \approx 3 \cdot 10^{22}$ молекул. Отже, на одну молекулу припадає об'єм $V_0 = V/N \approx 3 \cdot 10^{-26}$ м^3. Оскільки в рідинах молекули розміщуються впритул одна до одної, поперечний розмір молекули можна оцінити як $d \approx (V_0)^{1/3} \approx 10^{-10}$ м. Існують і точніші методи визначення розмірів молекул, які дають величини, близькі до наведеної оцінки, причому розміри молекул різних неорганічних речовин виявляються приблизно однаковими[3].

2. Рух молекул. Теплова рівновага та температура

Рух молекул. Молекули речовини завжди перебувають у стані хаотичного руху. Про рух молекул, зокрема, свідчить явище дифузії – взаємопроникнення молекул однієї речовини в іншу. Хаотичність руху молекул наочно відображає броунівський рух - безперервний невпорядкований рух дуже дрібних частинок твердої речовини, що зважені в рідині. Рух молекул як такий не зумовлений якимось причинами або процесами в речовині, він є природнім станом частинок матерії. Водночас хаотичність руху молекул має своє пояснення. Молекула, стикаючись із іншою, змінює величину та напрям своєї швидкості, причому такі зіткнення відбуваються дуже часто. До прикладу, в газі за звичайних умов відбувається ~10^9 зіткнень за 1 с. Тому величини швидкості та напрямки руху окремих молекул є неконтрольованими, а їхній рух - невпорядкованим (хаотичним).

Рівноважні та нерівноважні стани й процеси. Із хаотичним рухом молекул пов'язані такі фундаментальні поняття, як ***теплова (термодинамічна) рівновага*** та ***температура.*** Попри невизначеність характеристик руху (мікроскопічних параметрів) окремої молекули, їхні середні значення для всієї системи або окремих її частин, а отже й макроскопічні параметри, у кожен момент часу мають визначену величину[4]. У загальному випадку макроскопічні параметри можуть змінюватися з часом, тоді макроскопічний стан системи називається *нерівноважним*. Але, як свідчить дослід, якщо система є замкненою, то її макроскопічні параметри в решті-решт обов'язково набувають значень, які надалі лишаються незмінними. Такий стан системи називається ***рівноважним***, а її перехід в рівноважний стан називається ***релаксацією***. Те, що будь-яка ізольована система спонтанно (сама по собі) переходить у рівноважний стан, є фундаментальним законом природи.

Зміна параметрів стану системи із плином часу називається ***термодинамічним процесом***. Як і стани, процеси можуть бути нерівноважними або рівноважними. При нерівноважному процесі всі або деякі термодинамічні параметри не мають визначеної величини для всієї системи. Наприклад, при швидкому стисканні газу в циліндрі внаслідок руху поршня тиск і температура газу будуть різними в різних місцях - найвищими біля поршня й найнижчими біля основи циліндра. Отже, такий процес є ***нерівноважним*** або ***нестатичним***. Але, якщо поршень переміщувати дуже повільно, то внаслідок взаємодії між молекулами газу їх концентрація та середня швидкість хаотичного руху будуть устигати вирівнюватися так, що в кожен момент часу термодинамічні параметри (тиск і температура) будуть мати практично однакові значення по всьому об'єму газу. Такий процес називається ***рівноважним*** або ***квазістатичним***. Рівноважний процес можна розглядати як неперервну послідовність рівноважних

станів системи. Строго рівноважним може бути лише нескінченно повільний процес. Наприклад, для того, щоби зміна об'єму газу в циліндрі під поршнем була рівноважною (квазістатичною), треба щоб у кожен момент часу тиск газу був лише на нескінченно малу величину більшим за зовнішній тиск при розширенні і так само меншим при стисканні газу[5]. Тому напрям процесу можна змінити, змінивши в потрібний бік зовнішній тиск знов-таки на нескінченно малу величину. Як наслідок, провівши такий процес в одному, а потім у зворотному напрямі, ми повернемо газ у вихідний стан так, що в навколишньому середовищі не відбудеться ніяких змін. Процес, який має таку властивість, називається ***оборотним***. Отже, ***квазістатичні процеси є оборотними***. Зрозуміло, що такі процеси є ідеалізацією - усі ***реальні процеси є необоротні***. Проте, не рідко, вирівнювання параметрів у системі відбувається швидко. Наприклад, у газах релаксація йде приблизно із швидкістю звуку. Тому реальні процеси часто можна з достатньою точністю розглядати як оборотні (квазистатичні).

Температура. Індикатором стану термодинамічної рівноваги системи є ***температура***[6]. Якщо система не є рівноважна, то різні її частини (тіла) мають не однакову температуру, але в процесі релаксації більш гарячі тіла охолоджуються, а холодні нагріваються. Відтак, у стані термодинамічної рівноваги температура всіх тіл будь-якої системи або різних частин одного тіла стає однаковою. Слід наголосити, що температура є єдиним термодинамічним параметром, який має вказану властивість. Наприклад, у системі, що складається з декількох гумових кульок із газом і перебуває в стані термодинамічної рівноваги, тиск газу та концентрація молекул у різних кульках може бути різною, але температура - обов'язково є однаковою й рівною температурі навколишнього повітря. Для практичного вимірювання температури використовують якусь залежну від температури властивість певної речовини ("термометричного тіла") і відповідну температурну шкалу. Так, у побуті найчастіше користуються рідинними термометрами, дія яких ґрунтується на тому, що об'єм термометричної рідини (спирт або ртуть) лінійно залежить від температури в необхідному для практичних потреб інтервалі значень. При цьому найпоширенішою є шкала Цельсія, в якій початок відліку (0 °C) відповідає температурі твердіння води при нормальному атмосферному тиску, а температура кипіння води приймається рівною 100 °C. Є й інші температурні шкали такого типу (шкала Реомюра, шкала Фаренгейта), які відрізняються вибором нульової точки (початку відліку) та розміром одиниці температури.

Згадані температурні шкали не відображають фізичної природи температури, але виражають її зручними для практики умовними числами. Але в науці найчастіше використовують так звану ***абсолютну термодинамічну шкалу температур*** або ***шкалу Кельвіна***. Одиниця температури в цій шкалі називається ***кельвін*** **(К)**, а температура, виміряна в кельвінах, називається ***абсолютною температурою***. Абсолютна температура T є пов'язана із температурою t в шкалі Цельсія співвідношенням $T = 273{,}15 + t$. На практиці перший доданок округлюють до цілого і приймають

$$T = 273 + t.$$

Звідси випливає, що розмір одиниці в обох шкалах однаковий і температурі $T = 0$ К, яка називається ***абсолютним нулем***, у шкалі Цельсія відповідає температура -273 °C.

Шкала Кельвіна має глибокий фізичний зміст, який розкриває молекулярно-кінетична теорія. Розглянемо молекулярний механізм установлення термодинамічної рівноваги в теплоізольованому циліндрі, що розділений тонкою теплопровідною перегородкою та заповнений газом. Нехай початковий стан газу є нерівноважним, і молекули по обидва боки від перегородки мають різні середні швидкості хаотичного руху. Молекули обох газів стикаються з молекулами перегородки, через що кінетична

енергія більш швидких молекул зменшується, а повільних - збільшується. Позаяк в одній частині циліндра кількість швидких молекул більша, ніж в іншій, то через перегородку встановиться потік енергії від газу з у середньому більш швидкими молекулами до газу з повільнішими молекулами. Очевидно, що цей процес триватиме доти, доки не зрівняються середні кінетичні енергії поступального руху молекул газу в одній та в інший частині циліндра. Після цього макроскопічний стан газу в обох частинах буде лишатися незмінним, тобто система перейде в рівноважний стан. Але, як відомо з досліду, так само поводиться й температура. При встановленні рівноваги температура газу в одній частині циліндра знижується, а в іншій - підвищується, доки в усій системі не встановиться однакова температура. Отже, можна зробити висновок про те, що температура пов'язана з хаотичним рухом молекул так, що тіло з більш високою температурою має більшу середню кінетичну енергію молекул [7]. У молекулярно-кінетичній теорії доводиться, що цей зв'язок є універсальним і виражається формулою

$$\langle E \rangle = \frac{3}{2} kT, \qquad (1.2)$$

де $\langle E \rangle$ - середня кінетична енергія *поступального* теплового руху однієї молекули, k = $1{,}38 \cdot 10^{-23}$ Дж/К - універсальна стала, що називається ***сталою Больцмана***, T - абсолютна температура. Це співвідношення розкриває зміст абсолютного нуля температури:

> ***T* = 0 К - *то є температура, при якій припиняється тепловий рух частинок речовини*** [8].

3. Взаємодія молекул

Агрегатні стани. Хоча атоми й молекули в цілому є електрично нейтральними, вони складаються із заряджених частинок - позитивних ядер і негативних електронів. Якщо відстань між молекулами велика, сили що діють на заряди однієї молекули з боку зарядів іншої є компенсованими, і взаємодія між молекулами відсутня. Але на відстанях, сумірних із розмірами самих молекул, такої компенсації немає й між молекулами виникають сили взаємодії електромагнітної природи. Механізми виникнення та властивості цих сил пояснюються в квантовій механіці і є досить складними. Зокрема відомо, що між зарядженими частинками молекул діють і сили притягання [9] $F_п$ і сили відштовхування $F_в$, причому із зміною відстані між молекулами сили відштовхування змінюються набагато швидше, ніж сили притягання, як показано на схематичному рис. 1.1 а. Тому результуюча сила взаємодії F на порівняно великих відстанях r між молекулами є силою притягання, а на малих - силою відштовхування. На певній відстані $r = r_o$ сили притягання та відштовхування є компенсованими (F = 0). При зміщенні однієї молекули відносно іншої з положення $r = r_o$ виникає сила, що напрямлена проти переміщення й намагається повернути молекули у вихідне положення. Отже, відстань між молекулами $r = r_o$ відповідає положенню стійкої рівноваги. При будь-якому зміщенні молекули з цього положення сила F виконує від'ємну роботу і потенціальна енергія взаємодії молекул збільшується [10]. Тому при $r = r_o$ потенціальна енергія взаємодії двох молекул мінімальна і її залежність від відстані r (рис. 1.1б) має вигляд характерної кривої з мінімумом. Її називають ***потенціальною ямою***, а величину $E_{по}$ – ***глибиною ями*** й вимірюють в електрон-вольтах (еВ). Нагадаємо, що 1 еВ = $1{,}6 \cdot 10^{-19}$ Дж.

Окрім потенціальної енергії взаємодії $E_п$ кожна молекула має також кінетичну енергію теплового руху $E_к$ та відповідну повну енергію $E = E_п + E_к$, яка не змінюється при русі молекули без зіткнення з іншою (горизонтальні лінії E_1, E_2, E_3 на рис 1.1б). Глибина потенціальної ями $E_{по}$ визначається тільки природою молекул і для різних

речовин суттєво відрізняється. Кінетична енергія теплового руху молекул $E_{к0}$ не залежить від речовини і визначається тільки її температурою. Тому спостерігаються три характерні ситуації.

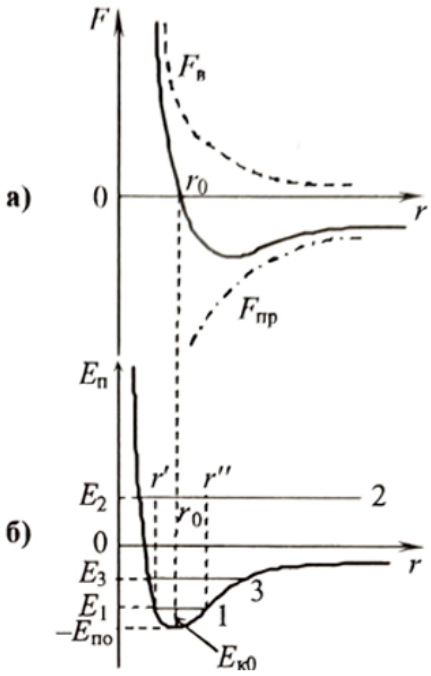


Рис. 1.1

1) Сили взаємодії між молекулами великі, тоді $E_{к0} << E_{п0}$ і повна енергія на рис. 1.1б зображується горизонтальною лінією 1. У такому випадку середня відстань між молекулами має певну величину $r_о$, і вони лише здійснюють малі коливання навколо положень рівноваги. Така поведінка молекул спостерігається в твердих тілах, для яких є характерними здатність зберігати свою форму, твердість і велика механічна міцність.

2) Якщо сили взаємодії між молекулами малі, так що $E_{к0} > E_{п0}$, то за рахунок своєї кінетичної енергії одна молекула може віддалитися від іншої на необмежену відстань. Така ситуація спостерігається в газах і зумовлює їхні характерні властивості – здатність займати весь наданий об'єм і порівняно легко стискатись.

3) Найскладнішим є молекулярний рух у рідинах. У рідинах глибина потенціальної ями лише трохи більша за середню кінетичну енергію теплового руху молекул і лінія повної енергії займає положення 3 на рис. 1.1б. За таких умов молекули рідини, в основному, здійснюють коливання навколо положень рівноваги як і в твердому тілі. Цим пояснюється те, що рідини зберігають свій об'єм і є майже нестисливими. Але, на відміну від твердого тіла, в рідині завжди присутня помітна кількість "гарячих" молекул, які мають більшу, ніж середня, кінетичну енергію і здатні подолати силу притягання до сусідніх молекул. Такі молекули час від часу переміщуються на декілька міжатомних відстаней ("кочують") і займають нові рівноважні положення. Завдяки цьому рідина має текучість, і набуває форму посудини, в якій знаходиться.

Наявністю "гарячих" молекул пояснюється й пароутворення, тобто перехід рідини в газоподібний стан - пару[11]. Так само при нагріванні твердого тіла, коли кінетична енергія молекул стає достатньою для подолання жорсткого зв'язку між молекулами, спостерігається плавлення - перехід речовини з твердого стану в рідкий.

Демонстрація теплового руху та взаємодії молекул.

Теплове розширення. Особливості взаємодії між молекулами пояснюють і добре відомий ефект теплового розширення. Він спостерігається через неоднакову залежність сил притягання та відштовхування від відстані між молекулами, крива потенціальної енергії (потенціальна яма) на рис. 1.1б є асиметричною – її права стінка йде більш полого, ніж ліва. Тому при зростанні температури та кінетичної енергії збільшується не тільки амплітуда коливань молекул навколо положень рівноваги, а й середня відстань між ними $r = (r' + r'')/2$ (штрихова крива на рис. 1.1в). Саме це є причиною збільшення лінійних розмірів і об'єму рідин та твердих тіл. При порівняно невисоких температурах середня відстань між молекулами змінюється пропорційно до температури, тому лінійні розміри l і об'єм V тіл лінійно залежать від температури:

$$l = l_0\,(1 + \alpha t^\circ),$$

$$V = V_0\,(1 + \beta t^\circ),$$

де l_0, V_0 – розміри та об'єм при температурі t = 0 °C, α, β – коефіцієнти лінійного та об'ємного розширення, відповідно, причому $\beta = 3\alpha$.

Контрольні запитання

1. Які методи дослідження використовуються в молекулярній фізиці? В чому полягає їхня сутність?
2. Що означають терміни "мікроскопічні параметри" та "макроскопічні параметри"?
3. Які одиниці використовують для визначення мас атомів і молекул? Як вони пов'язані муж собою?
4. У чому полягає відмінність між кількістю та масою речовини та як вони пов'язані між собою?
5. Який зв'язок існує між молярною масою речовини та масою однієї молекули, якщо остання виражена в кг і в а.о.м.?
6. Як можна оцінити розміри молекул у рідині через її молярну масу та густину? Чи придатний такий спосіб для твердих тіл? газів?
7. Який фізичний зміст мають поняття "температура" та "абсолютний нуль" температури?
8. У посудині з газом, яка рухається, середня кінетична енергія молекул більша, ніж у нерухомій. Чи означає це, що й температура газу в рухомій посудині вища? Чому?
9. Що таке "термодинамічна рівновага"? Чи є зв'язок між нею та механічною рівновагою?
10. Який характер мають сили взаємодії між атомами та молекулами? Наведіть наочні свідчення існування сил притягання та сил відштовхування між молекулами.
11. Відомо, що дві зіставлені поверхнями гладенькі плоскі пластинки не пристають одна до одної, тоді як дві відполіровані до дзеркального блиску пластинки "склеюються" так, що їх важко відірвати одна від одної. Чому?
12. Чому існують різні агрегатні стани речовини – твердий, рідкий та газоподібний?
13. Чому одна й та сама речовина може перебувати в різних агрегатних станах?
14. Як змінюються розміри тіл при зміні температури? З якою особливістю сил взаємодії між молекулами це пов'язано?

---

[1] Ізотопи - то є різновиди атомів даного хімічного елемента, що відрізняються кількістю нейтронів у ядрі.

[2] Найменування "а.о.м."не пишуть.

[3] Це означає, що молекули не можна розглядати просто як зіставлені докупи декілька кульок-атомів. Чому так, пояснює квантова механіка.

[4] Детальніше про це буде сказано при розгляді розподілу Максвелла-Больцмана

[5] Зрозуміло, що це можливо тільки за відсутності будь-якого тертя між поршнем і циліндром.

[6] Суб'єктивно температура пов'язується із відчуттям нагрітості - чим гарячіше тіло, тим вища його температура.

[7] Тому хаотичний рух частинок речовини називають *тепловим рухом*.

[8] Температура 0 К є граничною величиною: температура реальної фізичної системи може лише наближатися до абсолютного нуля, ніколи не досягаючи його точно. Крім того абсолютний нуль температури означає припинення не будь-якого , а лише теплового руху частинок. Зокрема, при 0 К зберігається рух електронів в атомах і коливання атомів у молекулах (так звані нульові коливання).

[9] Їх називають силами Ван-дер-Ваальса.
[10] Так відбувається тому, що ця робота дорівнює спаду потенціальної енергії (див. МЕХАНІКА, розділ IV).

[11] Теоретично пароутворення відбувається і в твердих тілах, але в твердому тілі кількість молекул, кінетична енергія котрих достатня для подолання зв'язку з іншими та виходу за межі тіла, є гранично малою.

2. Властивості газів

У газах сили притягання між молекулами малі, тому молекули не утримуються одна біля одної, а рухаються по всьому наданому об'єму. При не дуже низьких температурах середня кінетична енергія теплового руху молекул у газі істотно перевищує енергію зчеплення між ними. За таких умов можна вважати, що молекули в газах рухаються як вільні частинки, взаємодіючи лише при безпосередніх зіткненнях між собою та зі стінками посудини.

З іншого боку, при не дуже високих тисках концентрація молекул у газах є порівняно невелика, через що власний об'єм молекул виявляється набагато меншим, ніж об'єм посудини, в якій знаходиться газ. Наприклад, при кімнатній температурі й нормальному атмосферному тискові об'єм молекул складає ≈ 0,1% об'єму посудини, і на практиці власні розміри молекул можна не враховувати. Тому, розглядаючи гази, широко використовують спрощену модель "***ідеальний газ***" - систему, що складається з частинок несуттєвих розмірів, які взаємодіють лише при випадкових пружних зіткненнях між собою та зі стінками посудини. Інакше говорячи молекули ідеального газу не мають власного об'єму та потенціальної енергії взаємодії між собою[1]. Така модель добре відображає властивості справжніх газів за звичайних умов. При екстремальних умовах (наднизькі температури та надвисокі тиски) наявність власного об'єму та взаємодії між молекулами даються взнаки і властивості реальних газів стають більш складними. Але властивості ідеального газу можна теоретично встановити на основі лише простих молекулярно-кінетичних уявлень.

Нижче розглянуті наступні питання

2.1. Тиск ідеального газу

2.2. Рівняння Клапейрона. Ізопроцеси в газах

2.3. Реальний газ

Контрольні запитання

2.1. Тиск ідеального газу

Різноманітні технічні застосування газів визначаються їхньою здатністю створювати тиск на стінки посудини. Нагадаємо, що тиском називається відношення розподіленої сили, яка діє нормально на дану поверхню, до площі цієї поверхні. Тиск вимірюється в *паскалях*[2], 1 Па = 1 Н/м^2.

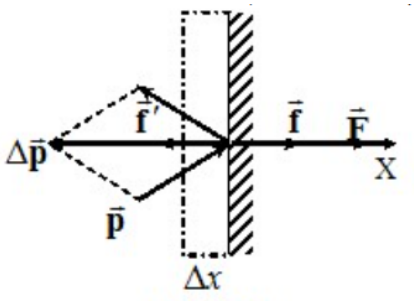


Рис. 2.1

Газ створює тиск на стінки посудини завдяки ударам молекул, причому в ідеальному газі ці удари є абсолютно пружними. Розглянемо дію молекул газу на малу ділянку стінки посудини, рис. 2.1.

Під час зіткнення на молекулу з боку стінки діє мікроскопічна сила, середня величина якої дорівнює відношенню зміни імпульсу молекули до тривалості зіткнення (другий закон Ньютона): $\vec{f}' = \Delta\vec{p}/\tau$. За третім законом Ньютона з боку молекули на стінку діє сила $\vec{f} = -\vec{f}'$. Можна сказати, що ця сила дорівнює імпульсу, переданому

молекулою стінці за одиницю часу. Величина Δp є дуже малою, тому дія молекул газу на стінку посудини має характер мікроскопічних миттєвих "уколів", урахувати які безпосередньо неможливо. Але у тому й немає потреби - через макроскопічну кількість молекул зіткнення відбуваються так часто та щільно, що створюють рівномірно розподілену по поверхні стінки сталу силу тиску F. Ця сила дорівнює сумарному імпульсу, що передається стінці всіма молекулами, котрі стикаються з нею протягом 1 с. Величину F можна визначити за допомогою таких міркувань. Пружне зіткнення молекули зі стінкою відбувається без утрати кінетичної енергії та величини швидкості. Тому молекула відлітає від стінки з таким самим за модулем імпульсу й під таким самим кутом, як і налітає. При цьому модуль переданого стінці імпульсу складає (див. рис. 2.1):

$$|\Delta\vec{p}| = 2p\cos\alpha = 2mv\cos\alpha \qquad \Rightarrow \qquad |\Delta\vec{p}| = 2mv_x,$$

де m - маса молекули, α – кут між вектором імпульсу молекули та нормаллю до стінки, v_x - проекція на вісь X швидкості, з якою молекула налітає на стінку. Для визначення сили тиску припустимо спочатку, що всі молекули, котрі налітають на стінку, мають однакову величину v_x. Тоді сила тиску

$$F = \nu \cdot 2mv_x.$$

де ν - кількість зіткнень молекул зі стінкою за 1 с. Щоб знайти ν, виділимо біля стінки шар молекул настільки малої товщини Δx, що в його межах молекули не стикаються між собою. Тоді всі молекули в шарі, що рухаються до стінки, зіткнуться з нею протягом часу $\Delta t = \Delta x/v_x$. Кількість таких молекул $\Delta N'$ дорівнює половині загальної кількості молекул у шарі. Тоді, якщо концентрація молекул газу (кількість в одиниці об'єму) дорівнює n і площа ділянки стінки S, то $\Delta N' = nS\Delta x/2$, і кількість зіткнень за одиницю часу

$$\nu = \frac{\Delta N'}{\Delta t} = \frac{1}{2}nx_xS \quad \text{і} \quad F = nmv_v^2S.$$

Насправді молекули мають різні швидкості, тому реальна сила тиску визначається середнім значенням квадрата проекції швидкості $\langle v_x^2\rangle$:

$$F = nm\langle v_x^2\rangle S. \tag{2.1}$$

Усі напрямки руху молекул є рівноймовірними й незалежними, тому можна записати

$$\langle v_x^2\rangle = \langle v_y^2\rangle = \langle v_z^2\rangle \text{ і } \langle v^2\rangle = \langle v_x^2\rangle + \langle v_y^2\rangle + \langle v_z^2\rangle = 3\langle v_x^2\rangle \quad \Rightarrow \quad \langle v_x^2\rangle = \frac{1}{3}\langle v^2\rangle. \quad .$$

Підставивши цей вираз у (2.1) і поділивши на площу S, отримуємо для тиску ідеального газу:

$$P = \frac{1}{3}nm\langle v^2\rangle \tag{2.2}$$

або

$$P = \frac{2}{3}n\langle E\rangle. \tag{2.2а}$$

Ці рівняння є принципово важливими, оскільки вони встановлюють теоретичний зв'язок між макроскопічним параметром газу - тиском і мікроскопічними параметрами - масою та швидкістю теплового руху молекул, або їх кінетичною енергією. Тому рівняння (2.2) або (2.2а) часто називають ***основним рівнянням молекулярно-кінетичної теорії газу***. З них випливають важливі наслідки.

Перш за все зауважимо, що в (2.2а) фігурує середня енергія *поступального* руху молекул, яка визначається тільки температурою газу, згідно з (1.2). Тому для тиску можна записати:

$$P = nkT. \qquad (2.3)$$

Отже, тиск ідеального газу не залежить від його природи й визначається тільки концентрацією молекул та температурою газу. Зокрема, це означає, що при однакових температурі та тиску всі гази мають однакову концентрацію молекул (закон Авогадро). За нормальних умов (температура 0 ^{0}C (273 К) і тиск 760 мм. рт. ст (1,01·10^5 Па)) вона складає n_0 = 2,69·10^{25} 1/м^3 і називається числом Лошмідта. При цьому 1 моль будь-якого газу за нормальних умов займає однаковий об'єм V_0 = 22,4·10^{-3} м^3. Із незалежності тиску газу від природи молекул випливає, що вираз (2.3) є чинним не лише для однорідних газів, а й для суміші газів, які не вступають між собою в хімічні реакції. У цьому випадку

$$n = n_1 + n_2 + \ldots + n_i + \ldots = \sum_i n_i,$$

де n_i - концентрації компонентів суміші. Відповідно, тиск суміші

$$P = \left(\sum_i n_i\right) kT = \sum_i (n_i kT) \quad \Rightarrow \quad P = \sum_i P_i, \qquad (2.4)$$

де $P_i = n_i kT$ - ***парціальні тиски*** компонентів, тобто тиски, які створювала би в посудині дана компонента суміші за відсутності всіх інших.

2.2. Рівняння Клапейрона. Ізопроцеси в газах

Рівняння стану. Кожному набору макроскопічних (термодинамічних) параметрів відповідає певний макроскопічний (термодинамічний) стан системи. Але термодинамічні параметри системи не можуть набувати тих або інших значень незалежно, вони пов'язані певними співвідношеннями, що називаються ***рівняннями стану*** системи. Прикладом може бути рівняння (2.3), яке по суті є одним з рівнянь стану ідеального газу. Але вказане рівняння не включає явно третій параметр стану газу - об'єм. Тому основним рівнянням стану ідеального газу вважається рівняння Клапейрона, яке можна легко отримати з (2.3). Для цього врахуємо, що концентрація молекул газу *n* дорівнює відношенню їхньої повної кількості *N* до об'єму посудини *V*: n =*N*/*V*. Тоді з (2.3) маємо *PV* = *NkT*. З іншого боку, згідно з (1.1), $N = \nu N_A$, де ν - кількість газу, N_A - стала Авогадро. Отже, $PV = \nu N_A kT$. Константа

$$R = N_A k = 8{,}31 \text{ Дж/(моль·К)} \qquad (2.5)$$

називається ***універсальною газовою сталою.*** Врахувавши це, отримуємо рівняння стану ідеального газу, яке називається ***рівнянням Клапейрона (Клапейрона-Менделєєва)***:

$$PV = \nu RT = \frac{m}{M} RT, \qquad (2.6)$$

де ν - кількість молів, *m* - маса і *M* - молярна маса газу. Для одного моля газу.

$$PV = RT.$$

Ізопроцеси в газах. Рівняння Клапейрона (2.6) дозволяє аналізувати не лише рівноважні стани, а й рівноважні процеси в газах, тобто з'ясовувати характер зміни термодинамічних параметрів газу за тих чи інших умов. Для теорії найбільший інтерес викликають так звані ***ізопроцеси***, тобто, процеси, що відбуваються з

незмінною кількістю газу і при незмінному значенні одного з параметрів стану. Розрізняють ***ізотермічний*** (ν = const, T = const), ***ізобарний*** (ν = const, P = const) та ***ізохорний*** (ν = const, V = const) процеси.

В ***ізотермічному процесі*** права частина (2.6) не змінюється, отже

$$PV = \text{const, де const} = \nu RT. \qquad (2.7)$$

Це рівняння виражає закон Бойля-Маріотта:

при постійній температурі добуток тиску заданої кількості газу на об'єм є сталим .

Графік, що показує зв'язок між параметрами стану в ізотермічному процесі, називається ізотермою і на діаграмі (P, V) для ідеального газу має форму гіперболи (рис. 2.2а).

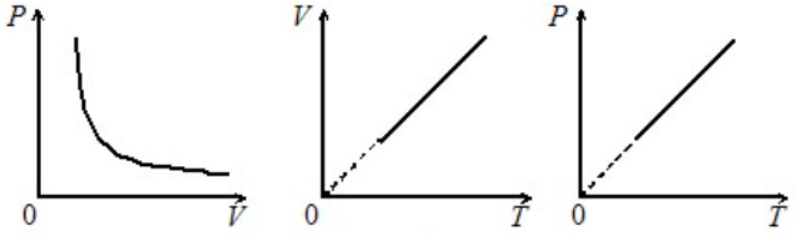


Рис.2.2

Ізобарний (ізобаричний) процес відбувається при сталому тиску газу, тому на основі (2.6) можна записати:

$$V = aT, \text{ де } a = \frac{mR}{MP} = \text{const}. \qquad (2.8)$$

Це є вираз закоу Гей-Люссака:

при сталому тиску об'єм заданої кількості газу змінюється прямо пропорційно абсолютній температурі.

Графік цього процесу (рис. 2.2*б* називається ***ізобарою***. на діаграмі (V, T).

В ***ізохорному (ізохоричному) процесі*** подібним чином змінюється тиск:

$$P = bT, \text{ де } b = \frac{mR}{MV} = \text{const}. \qquad (2.9)$$

Це є аналітичний вираз закону Шарля:

тиск заданої кількості газу при сталому об'ємі є прямо пропорційним температурі. Вигляд ***ізохори*** на діаграмі (P, T) показано на рис. 2.2*в*.

Рівняння (2.8) і (2.9) інколи записують через температуру в шкалі Цельсія, роблячи в (2.8) і (2.9) заміну $T = 273 + t$. Тоді

$$V = V_0(1 + \alpha t) \text{ і } P = P_o(1 + \beta t), \qquad (2.10)$$

де V_0 = 273 та P_0 = 273 – об'єм і тиск при 0 °С (273 К); α - термічний коефіцієнт об'єму (коефіцієнт об'ємного розширення) і β - термічний коефіцієнт тиску. Очевидно, що

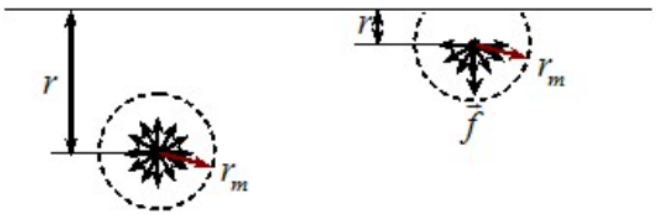
$$\alpha = \beta = \frac{1}{273}\ K^{-1}.$$

2.3. Реальний газ

Розглянуті рівняння стану для ідеального газу добре відображають властивості реальних газів за звичайних умов. Але при великих тисках і низьких температурах спостерігаються істотні відхилення в поведінці реальних газів порівняно з ідеальним. Тому є пояснення. При високих тисках концентрація молекул у газі стає настільки великою (див. (2.3)), що вже не можна нехтувати власним об'ємом молекул, тобто їхніми розмірами. Відповідно, середні відстані між молекулами стають настільки малими, що більше не можна нехтувати силами притягання між молекулами. Це особливо суттєво при низьких температурах, коли через малу кінетичну енергію теплового руху молекули при зустрічах не можуть легко долати сили зчеплення. Указані чинники враховують шляхом уведення відповідних поправок у рівняння Клапейрона. Існують різні способи такого коригування і, відповідно, різні рівняння стану реального газу. Але через складні властивості сил міжмолекулярної взаємодії та їхрю залежність від хімічної природи газу всі такі поправки й рівняння є наближеними.

Рівняння Ван дер-Ваальса. Найпроостішим і зручним для якісного аналізу властивостей реальних газів є рівняння стану, що запропонував Ван-дер-Ваальс на основі досить простих міркувань. Об'єм V, який фігурує в рівняннях (2.6), (2.6а) є об'ємом посудини, в якій міститься газ. В дійсності через наявність власних розмірів для молекул доступний не весь об'єм посудини. Тому в рівнянні стану замість V має фігурувати величина $V - b$, де b - поправка на власний об'єм молекул, яка приблизно дорівнює добутку об'єму однієї молекули на їх кількість [3]. Відповідно, тиск 1 моля реального газу мав би виражатися, як

$$P = \frac{RT}{V_0 - b}, \qquad (2.11)$$



а) б)

Рис. 2.3

де V_0 - молярний об'єм газу (об'єм 1 моля).

Але на тиск газу впливають і сили притягання між молекулами. Для аналізу цього впливу вводять поняття радіуса молекулярної дії r_m, тобто максимальної відстані між молекулами, на якій сили притягання ще мають помітну величину (ця відстань є досить умовною і складає ~10^{-9} м, тобто декілька поперечників молекули). Якщо молекула знаходиться на відстані $r > r_m$ від стінки посудини, то вона оточена сусідніми молекулами з усіх боків у середньому рівномірно і їхня дія є компенсованою (рис.2.3а). Але біля стінки, коли $r < r_m$, кількість сусідніх молекул, які тягнуть дану молекулу донизу, переважає кількість тих, які тягнуть її вгору. Тому на кожну молекулу біля стінки з боку інших діє певна результуюча сила $\vec{f}$, спрямована всередину газу (рис. 2.3б). Такі молекулярні сили сповільнюють рух молекул на підльоті до стінки і, згідно з (2.2), зменшують тиск газу на стінку на деяку величину P_i. З урахуванням цього, рівняння (7.13) слід записувати у вигляді

$$P = \frac{RT}{V_0 - b} - P_i. \qquad (2.11а)$$

Величина P_i визначається сумарною силою, що діє на молекули в прилеглому до стінки посудини шарі газу. Ця сила є пропорційна як кількості молекул у пристіночному шарі, на які діють молекулярні сили, так і кількості молекул у сусідньому шарі, що створюють її. Обидва шари мають приблизно однакову товщину порядку r_m і, відповідно, однакову кількість молекул, прямо пропорційну концентрації газу n. Тому $P_i \sim n^2$. У свою чергу концентрація молекул є обернено пропорційна об'єму газу (для 1 моля $n = N_A/V_0$), отже можна записати $P_i = a/V_0^2$, де a - деяка стала. У підсумку, рівняння (2.11а) набуває вигляду

$$P = \frac{RT}{V_0 - b} - \frac{a}{V_0^2},$$

або

$$\left(P + \frac{a}{V_0^2}\right)(V_0 - b) = RT. \qquad (2.12)$$

Для ν молів газу $V = \nu V_0$, відповідно, рівняння (2.12) набуває вигляду:

$$\left(P + \frac{a\nu^2}{V^2}\right)(V - \nu b) = \nu RT. \qquad (2.12а)$$

Рівняння стану реального газу (2.12) або (2.12 а) називається ***рівнянням Ван-дер-Ваальса.***

Ізотерми Ван-дер-Ваальса. Критичний стан. Подивимось, які специфічні властивості реальних газів передрікає рівняння Ван-дер-Ваальса та як теоретичні прогнози узгоджуються з дослідними фактами. Для цього проаналізуємо ізотерми Ван-дер-Ваальса, тобто залежності тиску газу від об'єму при сталих температурах, які випливають із рівнянь (2.12), (2.12 а). На рис. 2.4 показаний характерний вигляд теоретичних ізотерм Ван-дер-Ваальса для різних температур газу. При високих температурах, як і слід очікувати, вони нагадують ізотерми ідеального газу - при розширенні тиск монотонно зменшується (крива 1). Але при низьких температурах на кривих з'являється характерна хвиляста ділянка (крива 3).

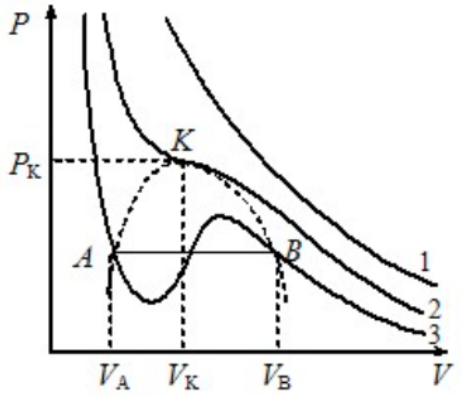


Рис. 2.4

Ці два різні типи залежностей розділяються так званою ***критичною ізотермою*** (крива 2), яка відповідає певній ***критичній температурі*** $T_к$ і має точку перегину *К*, яка називається критичною точкою. Цій точці, крім критичної температури $T_к$, відповідає ***критичнйм тиск*** $P_к$ і ***критичний об'єм*** $V_к$. Критичні параметри можна знайти, виходячи з того, що в точці К (точці перегину) перша та друга похідні тиску по об'єму дорівнюють нулю. Відповідні розрахунки на основі (2.12) для одного моля газу дають:

$$V_{0k} = 3b, \quad P_{0k} = \frac{a}{27b^2}, \quad T_k = \frac{8a}{27Rb} \qquad (2.13)$$

Як видно, критичні параметри визначаються виключно поправками Ван-дер-Ваальса a і b, тобто індивідуальними властивостями молекул. Але величина $K_к = RT_k/P_{0k}V_{0k}$, яка називається ***критичним коефіцієнтом***, є універсальною: згідно з (2.13) її значення, не≈залежно ні від чого, складає $K_к = 8/3 \approx 2{,}67$. Дослід свідчить, що це не зовсім так: критичний коефіцієнт є дещо більшим і залежним від газу[4]. Отже, рівняння Ван-дер-Ваальса не є точним. Про це свідчить і наявність підйомів на хвилястих ділянках ізотерм при $T < T_k$ (крива 3, рис. 2.4), які відповідають зростанню тиску при розширенні газу. Це є неможливо, бо означає збільшення тиску при зменшенні концентрації молекул при незмінній температурі. І дійсно, в експерименті указаний аномальний підйом на ізотермах реального газу не спостерігається. Але натомість виникає "поличка" (відрізок АВ на кривій 3, рис. 2.4): при зменшенні об'єму від $V_В$ до $V_А$ тиск газу залишається сталим, а потім стрімко зростає. Спостереження показують, що це пов'язано з конденсацією газу в рідину. У процесі стискання при об'ємі $V_В$ в посудині крім газу з'являється рідина, якої дедалі стає більше, а газу - менше. При об'ємі $V_А$ конденсація закінчується, і в посудині міститься лише рідина. Цим пояснюється й дуже стрімке зростання тиску при подальшому стисненні – адже рідина є практично нестисливою.

Таким чином, рівняння Ван-дер-Ваальса показує можливість переходу газу в рідину при ізотермічному стисканні. При цьому об'єм речовини під поршнем без зміни тиску зменшується від значення $V_А$, яка відповідає початковому об'єму газу, до $V_В$ – об'єму рідини, що утворилася з усього газу. Інтервал значень об'єму $V_В$ - $V_А$, при якому в посудині співіснують газоподібна та рідка фази речовини, що показаний штриховою кривою на рис. 2.4, із збільшенням температури зменшується і при $T = T_к$ зникає зовсім. При цьому, як показує дослід, при $V > V_k$ у посудині спостерігається лише газ, а при $V < V_k$ - лише рідина. Отже, при переході через критичну точку K увесь газ одномоментно перетворюється на рідину, проминаючи двофазну стадію. Самій точці $K(P_k, V_k, T_k)$ відповідає специфічний ***критичний стан*** речовини, при якому зникає будь-яка відміна між газом і рідиною. Зокрема, зникає меніск – межа поділу між рідиною та газом і спостерігається опалесценція - різке помутніння речовини по всьому об'єму безпосередньо перед переходом у рідкий стан. При температурах $T > T_к$ ізотерми Ван-дер-Ваальса не мають розглянутих особливостей, і можна припустити, що за таких умов газ весь час лишається газом, тобто його ніяким способом неможливо перевести в рідкий стан. Це дійсно так: практика отримання скраплених газів свідчить, що будь-який газ можна стисканням перетворити на рідину тільки при температурі, нижче критичної.

Підводячи підсумок, відмітимо, що рівняння Ван-дер-Ваальса якісно правильно відображає властивості реальних газів у широкому діапазоні параметрів стану. І це гідне подиву, оскільки в цьому рівнянні враховані не детальні властивості сил взаємодії між молекулами, а, за великим рахунком, лише сам факт існування таких сил.

Контрольні запитання

1. Що таке"ідеальний газ"? За яких умов природні гази можна розглядати як ідеальні?
2. Чому при високих тисках і низьких температурах властивості природних газів відрізняються від ідеальних?
3. Чим тиск газу на стінки посудини відрізняється від тиску рідини?
4. Якими мікроскопічними параметрами визначається тиск газу на стінки посудини?
5. Якими макроскопічними параметрами визначається тиск газу на стінки посудини?

6. Як визначається тиск суміші газів?

7. Що називається рівнянням стану та який вигляд воно має для ідеальних газів?

8. Чому рівняння Клапейрона можна використовувати для аналізу рівноважних процесів у газі і не можна – для нерівноважних?

9. Що таке ізопроцеси? Перелічіть ізопроцеси в газах.

10. Як виглядають ізотерми на діаграмі (*P, T*)? (*V,T)*?

11. Як виглядають ізобари на діаграмі (*P, V*)? (*P,T)*?

12. Як виглядають ізохори на діаграмі (*P, V*)? (*V,T)*?

13. Як впливає на тиск реального газу (порівняно з ідеальним) наявність розмірів у молекул? Як це пояснити, виходячи з механізму створення тиску?

14. Як впливає на тиск реального газу (порівняно з ідеальним) наявність сил притягання між молекулами? Як це пояснити, виходячи з механізму створення тиску?

15. Поясніть, які процеси в реальному газі відображають різні характерні ділянки ізотерм Ван-дер-Ваальса.

16. За яких умов газ при стисканні скраплюється (переходить у рідкий стан), а за яких – ні.

17. Що таке критичний стан речовини? Чи можна його спостерігати в експерименті?

---

[1] Але в деяких випадках доводиться враховувати рух і взаємодію атомів у молекулах.

[2] Атмосферний тиск вимірюють у позасистемних одиницях – *міліметрах ртутного стовпчика* (1 мм рт. ст. = 133,3 Па).

[3] Насправді вона дещо більша, оскільки навіть при щільній упаковці між молекулами лишаються певні не заповнені проміжки.

[4] для різних газів його величина лежить у межах від 3 до 5.

3 Розподіл Максвелла-Больцмана

У Вступі відмічалося, що в молекулярон-кінетичній теорії (МКТ) параметри макроскопічного стану (термодинамічні параметри) системи визначаються через середні значення її мікроскопічних параметрів, тобто, характеристик окремих молекул. Це наочно ілюструють рівняння (1.2), (2.2) або (2.2 а). Тому теоретичне визначення середніх значень мікропараметрів є принциповим завданням МКТ. Так само важливим є визначення кількості частинок у макросистемі, для яких значення того чи іншого мікропараметра лежить у заданих межах. Можливість розв'язання вказаних задач ґрунтується на тому, що в рівноважній термодинамічній системі існують певні ***статистичні розподіли*** частинок за різними мікропараметрами. Ці розподіли визначають стан системи й не змінюються з часом.

Далі розглядаються наступні питання:

1. Функція розподілу. Обчислення середніх значень
2. Розподіл Максвелла. Характеристичні швидкості молекул
3. Розподіл Больцмана

Контрольні запитання

1. Функція розподілу. Обчислення середніх значень

Статистичні розподіли аналітично задаються ***функцією розподілу***, яка визначає частку (відсоток) молекул у системі, що мають задану величину відповідного мікропараметра. Для різних систем і для різних параметрів функція розподілу може мати різний вигляд, але важливо те, що її можна встановити теоретично, виходячи з певних загальних уявлень про мікроскопічну будову системи та природу її об'єктів.

Поняття про функцію розподілу та про місце, яке вона посідає в МКТ, можна отримати на такому прикладі. Нехай у посудині є N молекул газу, який перебуває в стані термодинамічної рівноваги. Внаслідок випадкових зіткнень усі молекули газу в будь-який момент часу мають різні швидкості[1], а швидкість окремої молекули дуже часто й непередбачувано змінюється з часом. Але кількість молекул із заданими значеннями швидкості лишається незмінною. Зрозуміло, що кількість молекул $\mathrm{d}N_v$, які мають швидкість у вузькому інтервалі значень від заданого v до $v+\mathrm{d}v$[2], є прямо пропорційна загальній кількості молекул і ширині інтервалу: $\mathrm{d}N_v \approx N\mathrm{d}v$. Але вона залежить також і від величини v. До прикладу, молекул із дуже малими та великими швидкостями, є набагато менше, ніж із середніми. З огляду на це можна записати:

$$\mathrm{d}N_v = f(v)N\mathrm{d}v, \tag{3.1}$$

де залежна від швидкості величина

$$f(v) = \frac{\mathrm{d}N_v}{N\mathrm{d}v} \tag{3.2}$$

називається ***функцією розподілу*** молекул за модулем швидкості. Вона визначає

відносну кількість (частку) молекул, швидкості яких потрапляють в одиничний інтервал біля заданого значення v.

До цього питання можна підійти й з іншого боку. А саме, якби ми могли "вихоплювати" молекули із посудини по одній і вимірювати їхню швидкість, то з N таких випробувань у деякій кількості випадів $\mathrm{d}N_v$ отримали б числа в інтервалі від деякого значення v до $v+\mathrm{d}v$. При цьому величина

$$\frac{\mathrm{d}N_v}{N} = \mathrm{d}P(v) \tag{3.3}$$

визначає ймовірність (шанси) того, що одна навмання вихоплена молекула буде мати швидкість у вказаному елементарному інтервалі. Отже, вираз (3.2) можна записати як

$$f(v) = \frac{\mathrm{d}P}{\mathrm{d}v}. \tag{3.4}$$

Це дозволяє трактувати функцію розподілу як ***густину імовірності***, тобто імовірність того, що швидкість окремої молекули потрапить в *одиничний* інтервал біля заданого значення v. У такому разі ймовірність потрапляння швидкості окремої молекули в інтервал ширини $\mathrm{d}v$

$$\mathrm{d}P = f(v)\mathrm{d}v. \tag{3.4a}$$

Відповідно, ймовірність того, що окрема молекула буде мати швидкість у скінченному інтервалі $[v_1;\ v_2]$ визначається як

$$P(v_1;\ v_2) = \int\limits_{v_1}^{v_2} f(v)\mathrm{d}v. \tag{3.5}$$

Зокрема, якщо інтегрування провести по всіх можливих значеннях швидкості від 0 до нескінченності[3], то вийде одиниця, оскільки молекула достеменно має якесь значення швидкості:

$$\int\limits_0^\infty f(v)\mathrm{d}v = 1. \tag{3.6}$$

З цього приводу говорять, що функція розподілу є нормованою на одиницю, а інтеграл (3.6) називають ***інтегралом нормування***.

Функція розподілу $f(v)$ безпосередньо дозволяє визначити кількість молекул ΔN, які мають швидкість у заданому інтервалі значень $[v_1;\ v_2]$. Для цього треба проінтегрувати вираз (3.1) у відповідних границях:

$$\Delta N = N\int\limits_{v_1}^{v_2} f(v)\mathrm{d}v. \tag{3.7}$$

Для невеликих інтервалів швидкостей $\Delta v << v$ можна обійтися без інтегрування й використовувати наближений вираз

$$\Delta N \approx N f\left(v\right)\Delta v. \tag{3.7a}$$

За допомогою функції розподілу $f(v)$ можна визначати також середні значення будь-яких величин, залежних від v. Справді, нехай у макроскопічній системі з N частинок треба знайти середнє значення $\langle\Phi(v)\rangle$ певної величини $\Phi(v)$, яка є функцією швидкості. За означенням середнього

$$\langle\Phi(v)\rangle = \frac{\Phi_\Sigma}{N}, \tag{3.8}$$

де Φ_Σ - сума значень цієї величини для всіх N частинок. Через макроскопічну кількість частинок N безпосередньо обчислити Φ_Σ неможливо, але це можна зробити за допомогою функції розподілу. Будь-якому елементарному інтервалу $\mathrm{d}v$ відповідає

кількість частинок $\mathrm{d}N_v$, яка визначається виразом (3.1). Отже, сумарне значення величини $\Phi(v)$ на цьому інтервалі дорівнює

$$\mathrm{d}\Phi_\Sigma = \Phi(v)\mathrm{d}N_v = N\Phi(v)f(v)\mathrm{d}v.$$

Відповідно, для всіх частинок (і всіх значень v) маємо

$$\Phi_\Sigma = N\int\limits_v \Phi(v)f(v)\mathrm{d}v.$$

Отже, шукане середнє значення

$$\langle\Phi(v)\rangle = \int\limits_v \Phi(v)f(v)\mathrm{d}v. \qquad (3.9)$$

Цей вираз визначає загальний спосіб знаходження середнього значення будь-якої величини, яка є функцією певного мікропараметра системи, через функцію розподілу частинок за цим параметром.

2. Розподіл Максвелла. Характеристичні швидкості молекул

Розподіл Максвелла. Закон розподілу молекул газу по швидкостях був теоретично встановлений Максвеллом і носить його ім'я. Виведення цього закону потребує досить складних логічних міркувань і математичних викладок. Тому необхідні для аналізу його змісту вирази функцій розподілу приймемо без доведення.

Функція розподілу молекул за проєкцією швидкості на вісь координат, наприклад v_x, має вигляд:

$$f(v_x) = \left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}} e^{-\frac{mv_x^2}{2kT}}, \qquad (3.10)$$

де m - маса молекули, T - абсолютна температура газу, k - стала Больцмана. Ця функція, згідно з (3.1), визначає кількість молекул, які мають *x*-проєкцію швидкості в заданому інтервалі $\mathrm{d}v_x$ і будь-яку величину проєкції швидкості на кожну з двох інших осей координат. Оскільки всі напрямки руху молекул є рівноймовірними, функції розподілу по проєкціях v_y і v_z є такими самими:

$$f(v_y) = \left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}} e^{-\frac{mv_y^2}{2kT}}, \qquad (3.10а)$$

$$f(v_z) = \left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}} e^{-\frac{mv_z^2}{2kT}}. \qquad (3.10б)$$

Далі розглянемо ***функцію розподілу молекул по трьох компонентах швидкості*** $f(v_x, v_y, v_z)$, яка визначає молекули із заданим набором проєкцій, а отже, із заданою величиною та напрямком швидкості[4]. Згідно з (3.1), кількість молекул, *x*-проєкція швидкості котрих лежить в інтервалі від v_x до $v_x + \mathrm{d}v_x$, - це є $\mathrm{d}N_{v_x} = Nf(v_x)\mathrm{d}v_x$. З цих молекул *y*-проєкцію швидкості в інтервалі від v_y до $v_y + \mathrm{d}v_y$ мають $\mathrm{d}N_{v_x,v_y} = \mathrm{d}N_{v_x}f(v_y) = Nf(v_x)f(v_y)\mathrm{d}v_x\mathrm{d}v_y$ молекул. Аналогічно визначаються й ті з $\mathrm{d}N_{v_x,v_y}$ молекул, які мають ще й *z*-проєкцію швидкості в заданому інтервалі від v_z до $v_z + \mathrm{d}v_z$. Отже, кількість молекул, у яких усі три проєкції швидкості лежать у заданих інтервалах значень, складає

$$\mathrm{d}N_{v_x,v_y,v_z} = \mathrm{d}N_{v_x,v_y}f(v_x)f(V_y)f(v_z)\mathrm{d}v_x\mathrm{d}v_y\mathrm{d}v_x. \qquad (3.11)$$

Звідси, згідно з (3.2), визначаємо шукану ***функцію розподілу*** $f(v_x, v_y, v_z)$:

$$f(v_x, v_y, v_z) = \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} e^{-\frac{m(v_x^2+v_y^2+v_z^2)}{2kT}}. \qquad (3.12)$$

У показнику експоненти стоїть величина $v_x^2 + v_y^2 + v_z^2 = v^2$ - квадрат модуля швидкості. Це означає, що, хоча дана функція розподілу враховує як величину, так і напрям швидкості молекули, сама вона від напряму швидкості не залежить. У цьому немає нічого дивного, оскільки через хаотичність теплового руху молекул усі напрями швидкості є рівноймовірними.

Нарешті, розглянемо ***функцію розподілу молекул газу за модулем швидкості*** $f(v)$, яка визначає кількість молекул із швидкостями з величиною в інтервалі від v до $v + \mathrm{d}v$ і будь-яким напрямком. Відповідний розрахунок дає такий вираз $f(v)$:

$$f(v) = 4\pi\left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} v^2 e^{-\frac{mv^2}{2kT}}. \qquad (3.13)$$

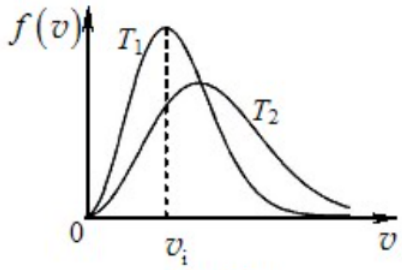


Рис. 3.1

Вигляд графіка цієї функції для двох температур ($T_2 > T_1$) показано на рис. 3.1.

Зауважимо, що всі наведені функції розподілу нормовані і, згідно з (3.6), площі під кривими на рис. 3.1 дорівнюють одиниці. Тому при підвищенні температури і зростанні середньої швидкості теплового руху молекул максимум функції розподілу переміщується в область більших швидкостей, а сама крива $f(v)$ стає ширшою та нижчою, тобто відбувається певне вирівнювання швидкостей молекул.

Характеристичні швидкості молекул. У різних задачах для характеристики теплового руху молекул використовують три різні так звані ***характеристичні*** швидкості: найбільш імовірну, середню арифметичну та середню квадратичну.

Найбільш імовірна швидкість молекул v_i відповідає максимуму функції розподілу $f(v)$ (рис. 3.1). Її можна визначити, взявши похідну від функції (3.13) і прирівнявши її до нуля. Результат такий:

$$v_i = \sqrt{\frac{2kT}{m}}. \qquad (3.14)$$

Домноживши чисельник і знаменник під коренем на сталу Авогадро N_A, отримаємо інший варіант формули:

$$v_i = \sqrt{\frac{2RT}{M}}, \qquad (3.14a)$$

де R - універсальна газова стала, M - молярна маса газу.

Середня арифметична швидкість молекул $\langle v \rangle$, згідно із загальним правилом (3.9), визначається як

$$\langle v\rangle = 4\pi\Big(\frac{m}{2\pi kT}\Big)^{\frac{3}{2}}\int\limits_0^\infty v^3 e^{-\frac{mv^2}{2cT}}\,\mathrm{d}v.$$

Інтеграл у цьому виразі є табличним[5] і дає

$$\langle v\rangle = \sqrt{\frac{8kT}{\pi m}} \qquad (3.15)$$

або

$$\langle v\rangle = \sqrt{\frac{8RT}{\pi M}}. \qquad (3.15a)$$

Середньою квадратичною швидкістю називається величина $v_{\text{кв}} = \sqrt{\langle v^2\rangle}$. Величина під коренем знаходиться аналогічно до $\langle v\rangle$:

$$\langle v^2\rangle = 4\pi\Big(\frac{m}{2\pi kT}\Big)^{\frac{3}{2}}\int\limits_0^\infty v^4 e^{-\frac{mv^2}{2kT}}\,\mathrm{d}v = \frac{3kT}{m}.$$

Отже, середня квадратична швидкість визначається формулами:

$$v_{\text{кв}} = \sqrt{\frac{3kT}{m}}. \qquad (3.16)$$

або

$$v_{\text{кв}} = \sqrt{\frac{3RT}{M}}. \qquad (3.16a)$$

Зауважимо, що всі характеристичні швидкості прямо пропорційні до $\sqrt{T/m}$ і мало відрізняються за величиною, а саме, $v_1 : \langle v\rangle : v_{\text{кв}} = 1 : 1{,}13 : 1{,}22$. Самі ж величини характеристичних швидкостей для газів при кімнатній температурі (~ 300 К) мають $10^2 \div 10^3$ м/с. До прикладу, для кисню O_2 $\langle v\rangle \approx 450$ м/с, а для водню H_2 — 900 м/с.

Зведений розподіл молекул за швидкостями. У деяких випадках розгляд задач, пов'язаних із розподілом молекул по швидкостях, спрощується, якщо користуватися безрозмірною ***зведеною швидкістю*** молекул $u = v/v_i$. Тоді кількість молекул $\mathrm{d}N$ із швидкостями в заданому інтервалі значень, яка визначається формулою (3.1), через функцію розподілу $f(u)$ виражається, як $Nf(u)\mathrm{d}u$. Отже, $f(v)\mathrm{d}v = f(u)\mathrm{d}u$, звідки

$$f(u) = f(v)\cdot\frac{\mathrm{d}v}{\mathrm{d}u} = f(v)v_i. \qquad (3.17)$$

Зробивши у виразі $f(v)$ (3.13) заміну $v = uv_i$ і врахувавши (3.14), із співвідношення (3.17) отримаємо зведену функцію розподілу $f(u)$:

$$f(u) = \frac{4}{\sqrt{\pi}}\cdot u^2 e^{-u^2}. \qquad (3.18)$$

Розподіл молекул за кінетичними енергіями. Кількість молекул із кінетичними енергіями в заданому інтервалі значень $\mathrm{d}E$ дорівнює кількості молекул із швидкостями у відповідному інтервалі $\mathrm{d}v$, отже $f(E)\mathrm{d}E = f(v)\mathrm{d}v$, де $E = mv^2/2$.

Тому, врахувавши (3.13), і зробивши заміни $v = \sqrt{2E/m}$ та $\mathrm{d}E = mv\mathrm{d}v$, знайдемо функцію розподілу молекул за кінетичними енергіями ***поступального*** **руху**:

$$f(E) = \frac{2\pi}{(\pi kT)^{3/2}} \cdot \sqrt{E} e^{-\frac{E}{kT}}. \qquad (3.19)$$

За допомогою цієї функції розподілу та виразу (3.9) можна визначити середню енергію поступального руху молекули:

$$\langle E \rangle = \frac{2\pi}{(\pi kT)^{3/2}} \int_0^\infty E\sqrt{E} e^{-\frac{E}{kT}} \mathrm{d}E. \qquad (3.20)$$

Цей інтеграл зводиться до табличного і дає результат:

$$\langle E \rangle = \frac{3}{2} kT, \qquad (3.21)$$

що співпадає з (1.2). Цей результат можна також отримати, виходячи з того, що

$$\langle E \rangle = \frac{m\langle v^2 \rangle}{2} = \frac{mv_{\text{кв}}^2}{2},$$

і врахувавши вираз (3.16).

За допомогою розподілу Максвелла також можна строго вивести вирази (2.2) і (2.2а) для тиску ідеального газу, і, з урахуванням (3.21), теоретично вивести рівняння стану ідеального газу (2.3) і (2.6), а також і всі наслідки, що випливають з них.

Дослідна перевірка розподілу Максвелла. Всі теоретичні прогнози, про які щойно говорилося, справджуються на дослідіі, що доводить правильність розподілу Максвелла. Але існують і прямі способи його дослідної перевірки. Розглянемо той, що був здійснений Ламмертом.

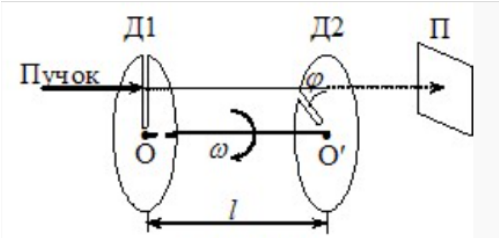


Рис. 3.2

Схема досліду Ламмерта така.У камеру з малим отвором уміщують досліджувану речовину (метал) і підіймають температуру вище точки плавлення. Як наслідок, у камері створюється достатня кількість пари даного металу, котру можна вважати ідеальним газом. З атомів, що вилітають через отвір, за допомогою діафрагм [6] формують вузький і практично нерозбіжний пучок, який спрямовують на вимірювальну установку, схематично показану на рис. 3.2. Вона складається з двох дисків Д1 і Д2, жорстко закріплених на спільній, паралельній до пучка осі OO′, та скляної пластинки П, на якій осаджуються атоми, що пройшли крізь установку. В дисках зроблені вузькі радіальні прорізи, що зміщені на невеликий кут φ ~ 1° один відносно одного. Відстань l між дисками заздалегідь вимірюють. Поки диски нерухомі, всі атоми, що проходять крізь щілину в диску Д1, повністю затримуються диском Д2. Але, коли диски почати обертати з необхідною кутовою швидкістю ω у показаному на рис. 3.2 напрямку, певна частина атомів пройде крізь прорізи в обох

дисках і осяде на пластинці П. Зрозуміло, що пройдуть тільки ті атоми, для яких час прольоту τ від першого диска до другого співпаде з часом повороту дисків на кут φ. Це дозволяє знайти швидкість v таких атомів. Оскільки $l = v\tau$ і $\varphi = \omega\tau$, то

$$v = \frac{l}{\varphi} \cdot \omega.$$

Таким чином, кожному значенню ω відповідає цілком визначена швидкість атомів, які проходять крізь обидва диски. Ці молекули осідають на пластинці П і створюють на ній тонкий напівпрозорий шар, товщина і прозорість якого прямо пов'язані з кількістю атомів, що осіли. Тому, обертаючи диски з різними кутовими швидкостями протягом однакового проміжку часу і, вимірюючи ступінь почорніння пластинок П, отримують залежність відносної кількості атомів у пучку від швидкості їх руху, тобто - функцію розподілу. Результати таких вимірювань, як й інші подібні досліди, повністю узгоджуються з Максвеллівським законом розподілу молекул за швидкостями.

3. Розподіл Больцмана

За відсутності зовнішніх сил частинки системи, наприклад молекули газу, завдяки хаотичності теплового руху та великій частоті зіткнень, розподіляються в просторі однорідно, і всі можливі значення координат для молекули є рівноймовірними. Однак у силовому полі однорідність розподілу частинок у просторі порушиться. Для з'ясуваня впливу силового поля на розподіл молекул по координатах, розглянемо, як сили тяжіння впливають на тиск газу.

Барометрична формула. Уявімо дуже високий вертикальний циліндр із газом, на молекули якого діє однорідна сила тяжіння, і температура в якому скрізь однакова.

Виділимо на висоті h від основи циліндра шар газу гранично малої товщини dh (рис. 3.3). Цей шар, як і весь газ, перебуває в механічній рівновазі. Отже, сила тиску газу на нижню основу цього шару $\vec{F}$ зрівноважує силу $\vec{F}'$, яка створюється силою тиску газу на верхню основу шару та його вагою:

$$F = F' \quad \Rightarrow \quad PS = (P + \mathrm{d}P)S + \mathrm{d}m \cdot g \quad \Rightarrow \quad S\mathrm{d}P = -g\mathrm{d}m, \tag{3.22}$$

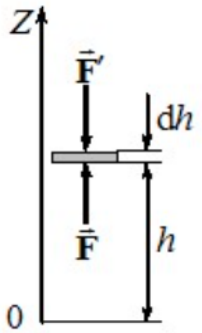


Рис. 3.3

де S - площа основи циліндра, dm - маса газу в шарі dh, P і P + dP - тиск газу на рівні нижньої та верхньої основи шару, відповідно. (Величина dP < 0, оскільки з висотою тиск зменшується). Маса газу в шарі визначається його густиною ρ і об'ємом dV = Sdh: $\mathrm{d}m = \rho S\mathrm{d}h$. Відтак вираз (3.22) набуває вигляду

$$\mathrm{d}P = -\rho g\mathrm{d}h. \tag{3.23}$$

Згідно з рівнянням Клапейрона (2.6), густина газу

$$\rho = \frac{\mathrm{d}m}{\mathrm{d}V} = \frac{PM}{RT},$$

де M - молярна маса газу, T - його температура, R - універсальна газова стала. Підставляючи цей вираз у (3.23), після інтегрування отримуємо:

$$\frac{\mathrm{d}P}{P} = -\frac{Mg}{RT}\mathrm{d}h \qquad \Rightarrow \qquad \ln\frac{P}{P_0} = -\frac{Mg}{RT}\cdot h. \tag{3.24}$$

Звідси

$$P = P_0 e^{-\frac{Mg}{RT}h}, \tag{3.25}$$

де P_0 - тиск біля основи циліндра (при h = 0).

Отримана залежність тиску газу від висоти стовпа (3.25) називається ***барометричною формулою***. Із застереженнями її можна застосовувати й до земної атмосфери, розглядаючи P_0 як тиск на рівні моря. Ці застереження такі. Повітря є сумішшю різних газів, тому молярну масу M слід розглядати як усереднену величину, що враховує відсотковий вміст у повітрі різних газів, а тиск P – як суму парціальних тисків компонент P_i. Але для компонент повітря з різними молярними масами M_i тиски P_i і концентрації молекул $n_i = P_i/kT$, згідно з (3.25), мають по-різному змінюватися з висотою - важчі швидше, ніж легші. Тому вміст компонентів і молярна маса повітря M мали би залежати від висоти h. Насправді цього не спостерігається, позаяк земна атмосфера не є ізотермічною. Температура повітря зменшується з висотою й неоднорідно розподіляється по поверхні Землі. Через це в атмосфері постійно циркулюють потужні повітряні потоки, які й вирівнюють співвідношення концентрацій компонентів повітря на різних висотах. Проте зміна температури повітря з висотою є суттєвою – на висоті 10 - 20 км вона на десятки градусів нижча, ніж на рівні моря. Тому барометрична формулу (3.25) для земної атмосфери є принципово неточною. Але, оскільки на рівні моря температура складає ~ 300 К, *відносна* зміна температури з висотою є порівняно невеликою, як і спричинена нею неточність формули (3.25). Тому за допомогою барометра, що проградуйований не в одиницях тиску, а в одиницях відстані, можна досить точно визначати висоту. Такі прилади називаються альтиметрами й використовуються на практиці, зокрема, в авіації.

Розподіли Больцмана та Максвелла-Больцмана. У випадку однорідного газу $M = mN_A$ (m - маса молекули), а, згідно з (2.3), $P = nkT$ (n - концентрація молекул). Окрім того $R = kN_A$. Зробивши такі підстановки в (3.25), отримаємо залежність концентрації молекул від висоти:

$$n = n_0 e^{-\frac{mgh}{kT}} \qquad \Rightarrow \qquad n = n_0 e^{-\frac{E_n}{kT}}, \tag{3.26}$$

де $E_n = mgh$ - потенціальна енергія молекули в полі сил тяжіння, n_0 - концентрація молекул на нульовому рівні енергії h = 0. Корисно відмітити, що така залежність концентрації молекул від висоти зумовлена протиборством між двома протилежними тенденціями – тепловим рухом, який намагається “розкидати” молекули рівномірно по всьому об’єму циліндра, та силою тяжіння, що намагається “опустити” всі молекули на дно циліндра.

Л. Больцман довів, що формула (3.26) є чинною не лише для поля сил тяжіння, а й для будь-якого іншого потенціального поля. Тому (3.26) є функцією розподілу частинок за потенціальними енергіями. Вона називається ***розподілом Больцмана***. Оскільки потенціальна енергія частинки в заданому полі визначається тільки її

координатами: $E_п = E_п(x, y, z)$, розподіл Больцмана по суті є ***розподілом частинок по координатах у силовому полі***. Відповідно до цього, кількість частинок $dN = ndV$ в об'ємі $dV = dxdydz$ в околі точки з координатами (x, y, z) визначається, як

$$dN_{x,y,z} = n_0 e^{-\frac{En(x,y,z)}{kT}} dxdydz. \qquad (3.27)$$

Слід зауважити, що через дуже малу масу молекул (~10^{-26} кг) помітне зменшення концентрації молекул повітря в полі сил тяжіння спостерігається лише на великих висотах[7]. За таких умов атмосферу не можна вважати ізотермічною, тому безпосередньо експериментально перевірити формули (3.25) і (3.26) неможливо. Але це можна зробити, досліджуючи розподіл за висотою броунівських частинок - дуже малих (~10^{-7} м) твердих частинок, які зважені в рідині. Броунівські частинки внаслідок ударів з боку молекул рідини здійснюють неупорядкований рух, який відображає тепловий рух молекул рідини. Тому броунівські частинки можна розглядати як гігантські молекули і вважати, що для них є чинними розподіли Максвелла та Больцмана. Такі досліди, що їх виконав Перрен, підтвердили правильність теоретичної формули (3.27) і дозволили експериментально визначити сталу Авогадро.

Розподіл Больцмана (3.27) та розподіл Максвелла (3.12) можна об'єднати в більш загальний ***розподіл Максвелла-Больцмана***, який дозволяє визначати кількості молекул, які знаходяться в заданій елементарній області простору (точці) і мають задану величину та напрям швидкості теплового руху. Вираз (3.27) визначає у кожному місці кількість частинок $dN_{x,y,z}$ із будь-якими швидкостями. З цієї кількості, відповідно до (3.11), задану величину й напрям швидкості мають $dN_{v_x,v_y,v_z,x,y,z} = dN_{x,y,z} f(v_x, v_y, v_z) dv_x dv_y dv_z$ частинок. Підставивши сюди вирази (3.27) і (3.12), отримаємо

$$dN_{v_x,v_y,v_x,x,y,z} = n_0 \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} e^{-\frac{m(v^2+v_y^2+v_z^2)/2+E_n(x,y,z)}{kT}} dv_x dv_y dv_z dxdydz \qquad (3.28)$$

або

$$dN_{v_x,v_y,v_z,x,y,z} = n_0 \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} e^{-\frac{E_k+E_n}{kT}} dv_x dv_y dv_z dxdydz, \qquad (3.28a)$$

де E_k - кінетична енергія частинки.

Вираз (3.28), чи (3.28а) називається ***розподілом Максвелла-Больцмана***. Він визначає кількість частинок маси m із концентрацією n_0, швидкості яких за величиною й напрямом нескінченно мало відрізняються від заданого значення і які знаходяться в заданій елементарній області простору (точці).

Контрольні запитання

1. Дайте означення поняття "функція розподілу молекул за швидкостями". Чи можна сказати, що функція розподілу визначає кількість молекул, які мають задану величину швидкості?
2. Покажіть на одному кресленні вигляд функції розподілу за швидкостями $f(v)$ для водню H_2 та гелію He при однаковій температурі. Як співвідносяться положення максимумів $v_1 : v_2$ та "висоти" піків на графіках $f(v)$?
3. Виходячи з розподілу молекул за модулем швидкості (3.13), отримати функцію розподілу молекул за енергіями (3.19).
4. Виходячи з розподілу молекул за енергіями (3.19), отримати функцію розподілу молекул за модулем швидкості (3.13).

5. Як за допомогою функції розподілу молекул за швидкостями (3.13) можна знайти середню кінетичну енергію поступального теплового руху молекул $\langle E \rangle$? Як це можна зробити за допомогою функції розподілу молекул за енергіями (3.19)?

6. Знайдіть за допомогою функції розподілу молекул за енергіями (3.19) найбільш імовірну енергію поступального теплового руху молекули.

7. Чи можна при визначепнні атмосферного тиску на деякій висоті h розглядати величину P_0 у формулі (3.25) як тиск не на рівні моря, а на довільній висоті H_0?

8. При кімнатній температурі молекули газу не осідають на дно, а практично рівномірно розподіляються по всьому об'єму посудини, бо можуть долати сили тяжіння за рахунок кінетичної енергії теплового руху. А як буде при дуже низькій температурі? Перевірте для гелію ($M = 4 \cdot 10^{-3}$ кг/моль) при температурі $T = 5$ К і нормальному атмосферному тиску.

9. Чому на легких планетах немає атмосфери?

10. Чому в атмосфері Землі кількість водню складає дуже малу частку від кількості кисню, попри те, що загальні кількості Гідрогену та Оксигену на планеті є сумірними?

11. Як би змінювався із часом відсотковий склад атмосфери Землі, якби вона не поповнювалася за рахунок природних і техногенних процесів?

---

[1] Замість швидкості можна говорити й про будь-який інший мікропараметр системи, наприклад, про імпульс або енергію окремої молекули.

[2] Більш точна, на перший погляд, постановка задачі про кількість молекул, які мають *задане* значення швидкості насправді є некоректною. Кількість молекул у системі хоча й велика, проте скінченна, тому цілком може статися, що якогось заданого точного значення швидкості не має жодна з них.

[3] Те, що замість граничної швидкості c використовують ∞, є лише зручним формальним прийомом, який спрощує обчислення. Це зумовлено тим, що підінтегральна функція $f(v)$ досить швидко спадає при великих швидкостях, і її вклад в інтеграл стає несуттєвим задовго до досягнення значень $v \sim c$.

[4] Точніше - молекули, для яких модулі та напрямки вектора швидкості гранично мало відрізняються від заданих значень.

[5] Див., напр, Г.Б.Двайт "Таблицы интегралов", М., "Наука", 1977.

[6] Пластинки з малим отвором, які розміщується напроти отвору печі.

[7] До прикладу, на висоті 100 м тиск і концентрація молекул повітря всього на 1 % менші, ніж на рівні моря.

ІІ. ЕЛЕМЕНТИ ТЕРМОДИНАМІКИ

Термодинаміка побудована на чотирьох істинах, які прийнято називати ***законами*** або ж ***началами***. Ці закони (начала) встановлені на основі аналізу всієї сукупності відомих експериментальних фактів і не можуть бути виведені з якихось інших начал.

Нульове начало термодинаміки пов'язують з фактом існування рівноважного стану системи.

Перше начало термодинаміки було встановлене у працях Роберта Майєра, Джеймса Джоуля та Германа Гельмгольца.

Друге начало термодинаміки вперше було висловлене Саді Карно у формі, яка відповідає уявленням того часу про теплоту. Пізніше Рудольф Клаузіус удосконалив ідеї Карно, трактуючи теплоту як форму енергії. Саді Карно, Рудольф Клаузіус та Уільям Томсон (лорд Кельвін) уважаються засновниками термодинаміки. Подальший розвиток термодинаміки повязують з такими вченими як Герман Гельмгольц, Джозайя Гіббс, які ввели поняття про відну енергію, термодинамічні потенціали і розвинули вчення про необоротні процеси.

Третє начало термодинаміки про принципову недосяжність температури абсолютного нуля (теорема Нернста) було сформульовано Вальтером Нернстом.

У цьому розділі будуть розглянуті такі питання

4. Перше начало термодинаміки

5. Теплоємність. Політропічні процеси

6. Явища переносу

7. Друге начало термодинаміки

8. Ентропія

4. Перше начало термодинаміки

Викладені вище відомості про властивості газів добре ілюструють сутність і можливості молекулярно-кінетичного (статистичного) методу (МКТ). Але, як вже говорилося, в молекулярній фізиці широко використовується й термодинамічний метод дослідження, що ґрунтується на декількох основних законах – началах термодинаміки. Ці закони, встановлені дослідним шляхом, відображають такі зв'язки між макроскопічними характеристиками системи, які не пов'язані з її мікроскопічною будовою. Тому закони термодинаміки та всі наслідки, що з них випливають, є універсальними, тобто, чинними для будь-яких систем. Слід також зауважити, що закони термодинаміки мають не лише теоретичне, а й велике практичне значення, оскільки лежать в основі отримання та використання людством різних видів енергії[1].

У термодинаміці, крім параметрів стану, використовують й інші термодинамічні величини. Найперше це стосується внутрішньої енергії, теплоти та роботи.

Нижче розглянуті наступні питання:

1. Внутрішня енергія
2. Макроскопічна робота та теплота
3. Перше начало термодинаміки

Контрольні запитання

1. Внутрішня енергія

У широкому розумінні до внутрішньої енергії довільної фізичної системи входять усі види енергії, крім кінетичної енергії руху системи як цілого та її потенціальної енергії у зовнішніх полях. Стосовно речовини внутрішня енергія включає кінетичну енергію руху молекул відносно центра мас системи (інакше – теплового руху), потенціальну енергію взаємодії між молекулами, а також енергію руху та взаємодії атомів усередині молекули.

Внутрішня енергія одноатомного ідеального газу. В найпростішій термодинамічній системі - одноатомному ідеальному газі – внутрішня енергія U складається тільки з кінетичної енергії теплового руху атомів і дорівнює добутку кількості частинок N на середню кінетичну енергію однієї частинки $\langle E \rangle$. Згідно з (1.1), (1.1а), (1.2), (2.5),

$$U = \frac{3}{2}\nu RT = \frac{3}{2} \cdot \frac{m}{M} RT. \qquad (4.1)$$

Для одного моля

$$U = \frac{3}{2} RT. \qquad (4.1a)$$

Урахувавши рівняння Клапейрона (2.6), маємо ще один вираз U:

$$U = \frac{3}{2} PV. \qquad (4.2)$$

Записані вирази відображають дуже важливу властивість внутрішньої енергії не лише ідеального газу, а й будь-якої іншої термодинамічної системи. Ця властивість полягає в тому, що

внутрішня енергія однозначно визначається параметрами стану системи.

Відповідно, зміна внутрішньої енергії системи в будь-якому процесі визначається тільки зміною параметрів стану. До прикладу, при переведенні одноатомного газу зі стану 1 (P_1, V_1, T_1,) у стан 2 (P_2, V_2, T_2) зміна внутрішньої енергії складає

$$\Delta U = U_1 - U_1 = \frac{3}{2} \cdot \frac{m}{M} R(T_2 - T_1) = \frac{3}{2} \cdot \frac{m}{M} R\Delta T \qquad (4.3)$$

або

$$\Delta U = \frac{3}{2}(P_2V_2 - P_1V_1). \qquad (4.3a)$$

Отже, зміна внутрішньої енергії не залежить від процесу, тобто, способу переведення системи з одного стану в інший. Тому говорять, що

внутрішня енергія є функцією стану системи.

Ступені свободи. Молекули одноатомного ідеального газу не мають якоїсь конфігурації й можуть рухатися тільки поступально. У випадку складніших молекул при визначенні внутрішньої енергії необхідно враховувати й інші види їхніх рухів як цілого, а також рухи та взаємодію атомів усередині молекули. Для цього використовують поняття кількості ступенів свободи.

> ***Кількістю ступенів свободи i системи (молекули) називається кількість незалежних координат, яка є необхідною для однозначного визначення положення молекули в просторі.***

Зміна координат означає рух, отже, можна сказати, що кількість ступенів свободи молекули визначає кількість різних рухів, які вона може здійснювати. До прикладу, поступальний рух тіла визначається зміною з часом трьох координат центра мас. Відповідно, кожна молекула, в тому числі й одноатомна, має $i_п = 3$ поступальні ступені свободи. Для визначення положення жорсткої лінійної молекули[2] (рис. 4.1), окрім положення центра мас, необхідно вказати напрям її осі. Для цього треба задати ще дві координати, наприклад, кути α і β, що їх утворює вісь молекули з двома осями декартової системи (третій кут γ однозначно визначається через два інших умовою: $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$). Зміна цих кутів визначає обертання молекули навколо двох осей, перпендикулярних до осі молекули[3]. Тому жорстка лінійна молекула має $i_{об} = 2$ обертальні ступені свободи й повну кількість ступенів свободи $i = i_п + i_{об} = 5$.

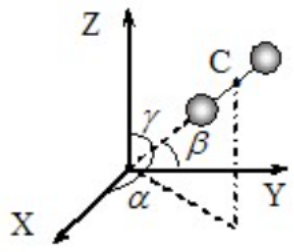


Рис. 4.1

У випадку жорсткої молекули, що складається з трьох і більше атомів, які не лежать на одній осі, є ще один можливий рух - обертання навколо осі, що проходить через будь-які два атоми. Тому багатоатомна жорстка молекула має $i_{об} = 3$ обертальні ступені свободи й повну кількість ступенів свободи $i = 6$. Якщо в системі з N частинок між ними немає жорсткого зв'язку, то положення кожної частинки треба задавати окремо, через що повна кількість ступенів свободи системи $i = 3N$. Ця кількість, окрім поступальних і обертальних ступенів свободи, що визначають рух системи як цілого, включає також ступені свободи, які пов'язані з рухом частинок системи відносно центра мас. Як відзначалося на початку цього розділу (п. 1.3), атоми в молекулах

зв'язані пружними силами й здійснюють коливальний рух, тому вказані додаткові ступені свободи молекул називаються *коливальними*. Кількість коливальних ступенів свободи дорівнює різниці між повною кількістю ступенів свободи молекули та сумарною кількістю поступальних і обертальних ступенів свободи: $i_{кол} = i - (i_п + i_{об})$. Коливальні ступені свободи молекули визначаються кількістю та конфігурацією атомів, із яких вона складається. При цьому для лінійної молекули з N атомів $i_{кол} = 3N - 5$, а для всіх інших конфігурацій $i_{кол} = 3N - 6$. Наприклад, для двохатомної пружної молекули кількість коливальних ступенів свободи $i_{кол} = 3 \cdot 2 - 5 = 1$, для трьохатомної лінійної $i_{кол} = 3 \cdot 3 - 5 = 4$, а для трьохатомної "трикутної" - $i_{кол} = 3 \cdot 3 - 6 = 3$.

Розподіл енергії за ступенями свободи. Згідно з (1.2), середня кінетична енергія поступального руху одноатомної молекули (матеріальної точки) дорівнює

$$\langle E_n \rangle = \frac{i_n}{2} kT,$$

де $i_n = 3$ - кількість поступальних ступенів свободи. Отже, на кожен поступальний ступінь свободи припадає середня кінетична енергія, що дорівнює $kT/2$. Л. Больцман довів, що це вірно й для кінетичної енергії обертального та коливального руху частинок. У цьому полягає класичний ***закон рівнорозподілу енергії по ступенях свободи:***

> ***у стані термодинамічної рівноваги на кожен ступінь свободи системи припадає однакова середня кінетична енергія***
>
> $$\langle E_1 \rangle = \frac{1}{2} kT.$$

Це зводить проблему визначення внутрішньої енергії довільної системи до підрахунку кількості її ступенів свободи. Правда, розглядаючи внесок в енергію рухів атомів усередині молекули, треба враховувати не лише кінетичну, а й потенціальну енергію їхньої взаємодії. Але це не складно, оскільки коливання атомів у молекулах близькі до гармонічних, а за таких умов середні значення кінетичної та потенціальної енергії коливань виявляються однаковими. Тож середня енергія, що припадає на один коливальний ступінь свободи, дорівнює двом половинкам kT. Через це, при підрахунках, кількість коливальних ступенів свободи подвоюють і обчислюють *середню енергію молекули* за загальною формулою

$$\langle E \rangle = \frac{i}{2} kT, \qquad (4.4)$$

де величина i визначається виразом:

$$i = i_n + i_{об} + 2i_{кол}. \qquad (4.5)$$

Відповідно до сказаного, для визначення внутрішньої енергії *довільного ідеального газу* в (4.1), (4.1а) і (4.2) трійку треба замінити на i:

$$U = \frac{i}{2} \nu RT = \frac{i}{2} \cdot \frac{m}{M} RT, \qquad (4.6)$$

$$U_m = \frac{i}{2} RT, \qquad (4.6a)$$

$$U = \frac{i}{2} PV. \qquad (4.7)$$

Для інших термодинамічних систем, зокрема для реальних газів, ці формули не підходять, оскільки вони не враховують потенціальну енергію взаємодії між молекулами.

2. Макроскопічна робота та теплота

Фізичні процеси завжди супроводжуються обміном енергією між тілами та її перетвореннями з одних видів на інші. Зокрема, в термодинамічних процесах, за винятком ізотермічних, змінюється внутрішня енергія системи. Внутрішню енергію можна змінювати різними конкретними способами, але є тільки дві різні причини такої зміни. Це – виконання макроскопічної роботи та теплообмін.

Макроскопічна робота. Добре відомо, що при швидкому стисканні газу в циліндрі під поршнем його температура, отже й внутрішня енергія, значно зростає[4]. Це пояснюється тим, що при русі поршня молекули газу після зіткнення з ним отримують додатковий імпульс і відлітають із більшими швидкостями, ніж підлітають. З іншого боку, для стискання газу до поршня має бути прикладена зовнішня сила, яка при переміщенні поршня виконує відповідну *макроскопічну роботу*[5].

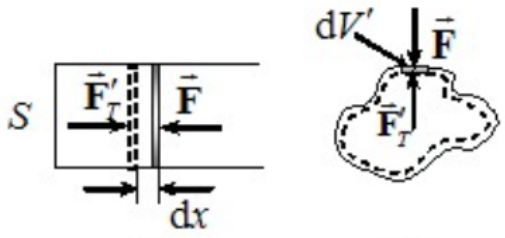


Рис. 4.2

Розглянемо стискання газу в циліндрі з поршнем площею S під дією зовнішньої сили $\vec{F}$ (рис 4.2 а), вважаючи, що стискання відбувається квазістатично (рівноважно). При переміщенні поршня на нескінченно малу відстань dx сила F виконує елементарну роботу $\delta A' = F\mathrm{d}x$. Оскільки процес є рівноважним, зовнішня сила за модулем дорівнює силі тиску газу на поршень $F_т = PS$ (P - тиск). Отож, $F = PS$, і робота $\delta A' = PS\mathrm{d}x$, або

$$\delta A = -P\mathrm{d}V, \qquad (4.8)$$

dV - зміна об'єму газу, знак "-" стоїть тому, що $\delta A' > 0$, коли $\mathrm{d}V < 0$, і навпаки. Відповідно, газ виконує таку саму по модулю, але протилежну за знаком роботу

$$\delta A = -\delta A' = P\mathrm{d}V. \qquad (4.8a)$$

Отримані вирази є чинними й у загальному випадку. Це легко довести. Нехай газ міститься в еластичній оболонці довільної форми (рис 4.2 б). При нескінченно малому переміщенні dl якоїсь однієї елементарної ділянки оболонки площею dS газ виконує роботу $P\mathrm{d}V'$, де $\mathrm{d}V' = \mathrm{d}S\cdot\mathrm{d}l$ - часткова зміна об'єму газу, зумовлена переміщенням цієї ділянки. Оскільки тиск газу в усіх точках однаковий, робота, зумовлена елементарним переміщенням усіх ділянок $\delta A = P\mathrm{d}V$, де dV – сума всіх dV', тобто елементарний приріст об'єму всього газу. Отже, в термодинаміці робота пов'язана не з переміщенням системи як цілого, а із зміною її об'єму, тобто – із переміщенням стінок, які відокремлюють систему від оточення.

У процесі стискання або розширення газу тиск, на загал, не лишається сталим. Тому при скінченній зміні об'єму газу робота зовнішніх сил A' та робота газу A визначаються інтегруванням виразів (4.8) і (4.8 а):

$$A' = -\int P\mathrm{d}V. \qquad (4.9)$$

$$A = \int P\mathrm{d}V. \qquad (4.9a)$$

Обчислення роботи газу в різних процесах.

В ***ізохорному*** (V = const) процесі $\mathrm{d}V = 0$, і робота не виконується: $\delta A = 0$.

В ***ізобарному*** процесі P = const і при зміні об'єму від V_1 до V_2 газ, згідно з (4.9), виконує роботу

$$A = \int_{V_1}^{V_2} P\mathrm{d}V = P\int_{V_1}^{V_2} \mathrm{d}V \quad \Rightarrow \quad A = P(V_2 - V_1) = P\Delta V. \qquad (4.10)$$

Робота зовнішніх сил над газом визначається, як

$$A' = P(V_1 - V_2) = -P\Delta V. \qquad (4.10a)$$

За допомогою рівняння Клапейрона (2.6) роботу газу в ізобарному процесі можна виразити й через зміну температури:

$$A = \nu R(T_2 - T_1) = \nu R\Delta T. \qquad (4.10б)$$

В ***ізотермічному*** процесі T = const. Урахувавши це та, підставивши в (4.9) вираз тиску з рівняння Клапейрона (2.6), отримаємо для роботи газу при ізотермічній зміні об'єму від V_1 до V_2:

$$A = \nu RT\int_{V_1}^{V_2} \frac{\mathrm{d}V}{V} = \nu RT\ln\frac{V_2}{V_1} = \frac{m}{M}RT\ln\frac{V_2}{V_1}. \qquad (4.11)$$

Оскільки в ізотермічному процесі PV = const, можна також записати:

$$A = \nu RT\ln\frac{P_1}{P_2} = \frac{m}{M}RT\ln\frac{P_1}{P_2}.$$

Робота зовнішніх тіл над газом при незмінній температурі, відповідно до (4.9), виражається як

$$A' = \nu RT\ln\frac{V_1}{V_2}. \qquad (4.11a)$$

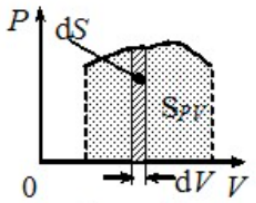


Рис. 4.3

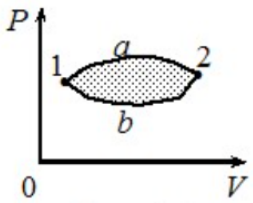


Рис. 4.4

Термодинамічна робота має простий геометричний зміст. На рис. 4.3 показано графік якогось процесу на діаграмі P-V ("у координатах" (P, V)). Видно, що при нескінченно малій зміні об'єму робота газу $\delta A = P\mathrm{d}V$ чисельно дорівнює площі $\mathrm{d}S$ смужки ширини $\mathrm{d}V$ під графіком процесу. Так само, робота при скінченній зміні об'єму від V_1 до V_2 чисельно дорівнює всій площі S під відповідною ділянкою графіка процесу на діаграмі (P, V):

$$A = S_{\mathrm{PV}}.$$

Такий геометричний підхід, звичайно, не знімає необхідності обчислення інтегралів (4.9) при визначенні роботи (крім випадку лінійної залежності $P(V)$), але він є дуже зручним для якісного аналізу термодинамічних процесів. Наприклад, він дозволяє легко зрозуміти важливу обставину: на відміну від внутрішньої енергії, макроскопічна робота не є функцією стану системи, оскільки залежить від способу переведення системи з початкового стану в кінцевий. Це добре видно з рис. 4.4: роботи, виконані при переході системи із стану 1 в стан 2 у процесах  та  не однакові й відрізняються на величину, що чисельно дорівнює площі заштрихованої області на графіку. Тому говорять, що

макроскопічна робота є функцією процесу.

Теплота. Внутрішня енергія може змінюватись і без виконання роботи, наприклад, при безпосередньому контакті тіл із різною температурою.

Процес переходу внутрішньої енергії від одних тіл до інших без виконання макроскопічної роботи називається теплообміном.

У процесі теплообміну частина внутрішньої енергії гарячого тіла переходить до холодного за рахунок зіткнень між більш швидкими молекулами гарячого тіла з повільнішими молекулами холодного. Отже, можна сказати, що при теплообміні внутрішня енергія змінюється за рахунок *мікроскопічної роботи*, що виконується на рівні окремих молекул. Слід, одначе, зауважити, що внутрішня енергія без виконання макроскопічної роботи може змінюватися й за рахунок випромінювання, та при зміні агрегатного стану речовини (плавлення - тверднення, пароутворення - конденсація), або в хімічних реакціях (як от при горінні).

Величина внутрішньої енергії, яку отримує або втрачає система без виконання макроскопічної роботи називається кількістю теплоти Q.

Дослід свідчить, що для зміни внутрішньої енергії системи на задану величину в різних процесах потрібна не однакова кількість теплоти. Тому, як і робота,

кількість теплоти є функцією процесу.

3. Перше начало термодинаміки

Між внутрішньою енергією, роботою та теплотою існує універсальний зв'язок, який встановлюється першим началом термодинаміки.

Формулювання першого начала термодинаміки. Як видно з попереднього, коли не виконується макроскопічна робота ($A' = 0$), то зміна внутрішньої енергії системи дорівнює отриманій нею кількості теплоти: $\Delta U = Q$. При виконанні макроскопічної роботи внутрішня енергія також змінюється. Дослідним шляхом установлено, що за відсутності теплообміну зміна внутрішньої енергії системи точно дорівнює виконаній над нею макроскопічній роботі: $\Delta U = A'$ при $Q = 0$. Отже, у загальному випадку

$$\Delta U = Q + A'. \qquad (4.12)$$

Це рівняння виражає ***перше начало термодинаміки*:**

зміна внутрішньої енергії довільної термодинамічної системи дорівнює сумі отриманої системою теплоти та виконаної над нею макроскопічної роботи.

Перше начало термодинаміки є ***законом збереження енергії,*** узагальненим на теплові процеси.

Для елементарного процесу (тобто, гранично малої зміни стану системи) перше начало записується так:

$$\mathrm{d}U = \delta Q + \delta A', \qquad (4.12а)$$

де $\mathrm{d}U$ - приріст внутрішньої енергії системи, δQ і $\delta A'$ - елементарна кількість теплоти, отримана системою, та елементарна робота, виконана над нею зовнішніми тілами. Слід зауважити, що всі величини в рівняннях (4.12) і (4.12а) є алгебраїчними. Наприклад, якщо система при теплообміні передає енергію іншим тілам, то $\delta Q < 0$. Так само, якщо об'єм системи збільшується, то $\delta A < 0$.

Рівняння (4.12) і (4.12а) можна записати інакше, якщо в (4.12) і (4.12а) зовнішню роботу замінити на роботу самої системи, згідно з (4.8а). Тоді

$$\delta Q = \mathrm{d}U + \delta A, \qquad (4.13)$$

$$Q = \Delta U + A. \qquad (4.13а)$$

Відповідно до цього, можна дати інше формулювання першого начала термодинаміки:

кількість теплоти, передана системі, йде на зміну її внутрішньої енергії та на виконання системою роботи над зовнішніми тілами.

У такій редакції перше начало термодинаміки прямо вказує на можливість виконання системою макроскопічної, тобто механічної, роботи за рахунок теплової енергії. Інакше говорячи, перше начало вказує на принципову можливість перетворення енергії невпорядкованого теплового руху молекул на механічну енергію, тобто енергію руху макроскопічних тіл. На цьому ґрунтується робота так званих теплових двигунів, які складають переважну більшість двигунів, що використовуються людиною.

Історично у становленні термодинаміки як науки вирішальну роль відіграли саме дослідження, спрямовані на підвищення ефективності теплових двигунів. У цьому зв'язку принциповим є висновок, який прямо випливає з (4.13) і показує неможливість так званого ***вічного двигуна першого роду*** - пристрою, котрий міг би необмежено довго виконувати роботу, не отримуючи енергію ззовні. Справді, для будь-якої системи, що не отримує енергії від інших тіл або навколишнього середовища, перше начало термодинаміки (4.13а) набуває вигляду

$$\Delta U + A = 0 \quad \Rightarrow \quad A = -\Delta U = U_0 - U_к,$$

де U_0 і $U_к$ – початкова та кінцева внутрішня енергія системи, відповідно. Звідси очевидно, що максимальна робота, яку міг би виконати будь-який двигун, навіть теоретично, не може перевищити його початкову внутрішню енергію: $A_{max} = U_0$. Отже, жоден двигун не може бути "вічним", він за будь-яких умов не зможе працювати довше, ніж доки вистачить запасу внутрішньої енергії[6]. Цей висновок є настільки принциповим, що його розглядають як третє формулювання першого начала термодинаміки:

вічний двигун першого роду є неможливим.

Термодинаміка ізопроцесів. За допомогою першого начала термодинаміки можна з'ясувати істотні деталі стосовно умов здійснення та перебігу різних процесів у термодинамічних системах, зокрема, в газах. Для цього запишемо рівняння (4.13) розлогіше, врахувавши вираз для роботи системи (4.8а), і застосуємо його до різних ізопроцесів.

$$\delta Q = \mathrm{d}U + P\mathrm{d}V. \qquad (4.14)$$

1. *Ізотермічний процес* (T = const). У цьому процесі, згідно з (4.6) і (4.7), внутрішня енергія не змінюється. Тому в рівнянні (4.12 а) dU = 0, і

$$\delta Q_{\mathrm{T}} = P\mathrm{d}V = \delta A \quad \text{і} \quad Q_{\mathrm{T}} = A. \tag{4.15}$$

Отже, в ізотермічному процесі вся надана системі теплота йде на виконання системою макроскопічної роботи, тобто, перетворюється на механічну енергію. При цьому виникає резонне запитання: яким чином система, отримуючи певну кількість внутрішньої енергії у формі теплоти δQ, зберігає свою температуру незмінною? При довільній зміні об'єму, дійсно, так не може бути. Але, якщо газ помістити в циліндр з ідеально теплопровідними стінками, і проводити процес квазістатично (гранично повільно), то температура скрізь буде лишатись однаковою й рівною температурі навколишнього середовища, тобто такий ідеалізований процес буде строго ізотермічним. Стискання або розширення газу за реальних умов завжди відбувається зі скінченною швидкістю, а стінки циліндра мають скінченну теплопровідність. Тому практично можна здійснювати лише наближено ізотермічні процеси.

2. *Ізохорний процес* (V = const). У цьому процесі робота не виконується (dV = 0), і перше начало термодинаміки (4.12а) набуває вигляду

$$\delta Q_V = \mathrm{d}U, \tag{4.16}$$

або, для скінченних величин,

$$Q_V = \Delta U = U_2 - U_1. \tag{4.16а}$$

Таким чином, позаяк у газі ізохорний процес здійснюється в циліндрі із закріпленим поршнем, товнутрішня енергія змінюється лише за рахунок теплообміну при контакті системи із більш гарячими чи більш холодними тілами, або за рахунок випромінювання.

3. *Ізобарний процес* (P = const). Такий процес у газі завичай здійснюється шляхом повільного нагрівання або охолодження газу в циліндрі, закритому рухомим поршнем. За таких умов газ або розширюється, або стискається так, що його тиск весь час лишається сталим і рівним сумі зовнішнього тиску на поршень і тиску, що створюється його вагою. Тому в ізобарному процесі не лише змінюється внутрішня енергія системи, а й виконується робота. Процес іде відповідно до рівняння (4.10), і, згідно з (4.12 а),

$$Q_P = \Delta U + P\mathrm{d}V, \quad \text{або} \quad Q_P = U_2 - U_1 + P(V_2 - V_1). \tag{4.17}$$

Контрольні запитання

1. Чи зміниться внутрішня енергія газу в закритій посудині, якщо вона почне швидко рухатися?
2. Говорять, що внутрішня енергія є функцією стану системи, а робота та теплота – функціями процесу. Що означають ці вислови?
3. Скільки обертальних і скільки коливальних ступенів свободи має лінійна молекула вуглекислого газу O-C-O та молекула води, в якій атоми утворюють трикутник?
4. Який відсоток від повної енергії складає енергія поступального теплового руху в жорсткій та в пружній двохатомній молекулі?
5. Які існують способи зміни внутрішньої енергії системи? Що таке "кількість теплоти"?
6. В якому процесі вся отримана системою теплота йде на зміну внутрішньої енергії?

7. В якому процесі вся отримана системою теплота перетворюється на механічну роботу?
8. Коли для підвищення температури заданої кількості газу на задану величину ΔT доведеться витратити більше тепла, – нагріваючи газ при сталому тиску, чи при сталому об'ємі?
9. Чи може система виконувати механічну роботу за рахунок своєї внутрішньої енергії?
10. Чи можна перетворити всю внутрішню енергію системи на механічну роботу?

---

[1] У цьому контексті слід сказати, що основи термодинаміки заклав французький інженер С. Карно в дослідженнях, спрямованих на вдосконалення теплових двигунів.

[2] Тобто, такої, в якій атоми нерухомі відносно центра мас і розташовані на одній прямій.

[3] Говорити про обертання двохатомної молекули навколо третьої (власної) осі беззмістовно, оскільки атоми розглядаються як матеріальні точки.

[4] За рахунок цього, зокрема, відбувається самозагоряння паливної суміші в дизельних двигунах.

[5] Термін "макроскопічна робота" у термодинаміці вживають, аби відрізнити роботу, що виконується при переміщенні макроскопічних тіл (тут - поршня) від "мікроскопічної роботи", що виконується силами взаємодії між окремими молекулами. Але коли це не викликає непорозумінь, слово "макроскопічна" опускають.

[6] Як показує друге начало термодинаміки (Лекція 7, п. 7.3), в дійсності навіть такої максимальної роботи не можна отримати.

5. Теплоємність. Політропні (політропічні) процеси

Нижче мова піде про таке:

1. Теплоємності ідеального газу

2. Політропні процеси

3. Процес Джоуля-Томсона

Контрольні запитання

1. Теплоємності ідеального газу

Теплоємність. При теплообміні змінюються внутрішня енергія та температура системи (тіла)[1]. При цьому одна й та сама передана кількість теплоти нагріває одні тіла сильніше, а інші слабше. Тож різні тіла мають не однакову здатність до нагрівання, котра кількісно визначається ***теплоємністю тіла*** C_0 **(Дж/К)**:

$$C_0 = \frac{\delta Q}{\mathrm{d}T}\left(\frac{\text{Дж}}{\text{К}}\right),$$

де δQ – отримана кількість теплоти і $\mathrm{d}T$ – зміни температури тіла.

Величина C_0 характеризує задане тіло. Але для теорії важливо знати теплові властивості не конкретного тіла, а речовини, з якої воно складається. Для цього уводяться ***молярна*** та ***питома*** теплоємності речовини.

Молярною теплоємністю C називається теплоємність одного моля даної речовини:

$$C_0 = \frac{\delta Q}{\mathrm{d}T}\left(\frac{\text{Дж}}{\text{К}}\right), \qquad (5.1)$$

де $\delta Q_m = \delta Q/\nu$ – кількість теплоти, що припадає на 1 моль, ν – кількість молів речовини, що міститься в тілі. Отже,

$$C = \frac{C_0}{\nu} \quad \text{або} \quad C = \frac{MC_0}{m},$$

де M - молярна маса речовини, m - маса тіла. Одиницею молярної теплоємності є 1 Дж/(моль·К).

Питома теплоємність c - це теплоємність одиниці маси речовини:

$$c = \frac{C_0}{m} = \frac{\delta Q}{m\mathrm{d}T}. \qquad (5.1a)$$

Вона вимірюється у Дж/(кг·К) і пов'язана з молярною теплоємністю очевидним співвідношенням:

$$c = \frac{C}{M}. \qquad (5.2)$$

У попередній лекції було з'ясовано, що в різних процесах заданій зміні внутрішньої енергії та температури тіла відповідає не однакова кількість теплоти. Тому, як і кількість теплоти,

теплоємність є функцією процесу.

Це означає, що теплоємність залежить не лише від властивостей речовини, а й від способу передачі тілу теплової енергії термодинамічній системі ("тілу"). Це випливає з першого начала термодинаміки й добре легко прослідковується в ізобарному та

ізохорному процесах. Згідно з виразами (4.14) і (4.16), при наданні системі однакової кількості теплоти зміна внутрішньої енергії, тож і температури, в ізохорному процесі буде більшою, ніж в ізобарному, оскільки при незмінному об'ємі на нагрівання йде все надане тепло, а при сталому тиску – лише його частина. Інша ж частина витрачається на роботу при розширенні (зміні об'єму) тіла. Тому для збільшення температури тіла на 1 К в ізобарному процесі треба витратити більше тепла, ніж в ізохорному. Отже, для будь-якої системи теплоємності при сталому тиску C_P та при сталому об'ємі C_V, є не однакові, причому $C_P > C_V$. Проте, на практиці ця різниця істотна лише для газів, оскільки зміна об'єму й робота розширення при нагріванні рідин і твердих тіл є несуттєвою. Тому для рідин і твердих тіл C_P і C_V не розрізняють, і в довідкових таблицях їхні теплоємності наводять без указівки на процес.

Теплоємності ідеального газу C_V і C_P. Внутрішня енергія ідеального газу залежить тільки від температури (формули (4.6), (4.6а)). Тому його молярні теплоємності C_V і C_P, які посідають важливе місце в теорії, легко визначаються з першого начала термодинаміки. Згідно з означенням (5.1) та виразами (4.16) і (4.14), маємо:

$$C_V = \frac{\delta Q_V}{\mathrm{d}T} = \frac{\mathrm{d}U}{\mathrm{d}T} \qquad (5.3)$$

і

$$C_P = \frac{\delta Q_P}{\mathrm{d}T} = \frac{\mathrm{d}U}{\mathrm{d}T} + \frac{P\mathrm{d}V}{\mathrm{d}T}. \qquad (5.3a)$$

Зміна внутрішньої енергії, що відповідає заданій зміні температури тіла, не залежить від процесу, то ж вираз C_P можна переписати так:

$$C_P = C_V + P\frac{\mathrm{d}V}{\mathrm{d}T}, \qquad (5.4)$$

де другий доданок визначає роботу, яку виконує 1 моль ідеального газу при ізобарному нагріванні на 1 К. Цю роботу легко визначити з рівняння Клапейрона (2.6 а) з урахуванням умови P = const:

$$\frac{\mathrm{d}V}{\mathrm{d}T} = \frac{R}{P} \quad \Rightarrow \quad P\frac{\mathrm{d}V}{\mathrm{d}T} = R.$$

Отже, робота одного моля ідеального газу при його ізобарному нагріванні на 1 К чисельно дорівнює універсальній газовій сталій[2]. Таким чином, між молярними теплоємностями ідеального газу є універсальне співвідношення, яке називається рівнянням Майєра:

$$C_P = C_V + R \quad \text{або} \quad C_P - C_V = R. \qquad (5.5)$$

Що ж до самих теплоємностей, то вони залежать лише від кількості ступенів свободи молекул газу. Справді, із (5.3) і (4.6а) і (5.5) маємо:

$$C_V = \frac{i}{2}R \quad \text{і} \quad C_P = \frac{i+2}{2}R, \qquad (5.6)$$

де i – кількість ступенів свободи молекул, яка визначається виразом (4.5).

Корисно відмітити, що, згідно з отриманим виразом C_V і формулами (4.6) і (4.6 а), внутрішня енергія ідеального газу виражається, як

$$U = \nu C_V T \quad \text{і} \quad U_m = C_V T. \qquad (5.7)$$

У теорії часто використовується величина

$$\gamma = \frac{C_P}{C_V}, \tag{5.8}$$

яка називається ***показником адіабати***. З (5.6) випливає, що для ідеального газу

$$\gamma = \frac{i+2}{i}. \tag{5.8a}$$

Зважаючи на останній вираз, у відповідних формулах замість кількості ступенів свободи i можна оперувати показником адіабати g. Наприклад, формули теплоємностей (5.6) та внутрішньої енергії ідеального газу (4.6) і (4.7) можна подати у вигляді:

$$C_V = \frac{R}{\gamma - 1}, \quad C_P = \frac{\gamma R}{\gamma - 1}; \tag{5.9}$$

$$U = \frac{\nu RT}{\gamma - 1} = \frac{m}{M} \cdot \frac{RT}{\gamma - 2}; \tag{5.10}$$

$$U = \frac{PV}{\gamma - 1}. \tag{5.11}$$

Неповнота класичної теорії теплоємностей. З формул (5.6) випливає. що теплоємності ідеальних газів визначаються тільки кількістю ступенів свободи, тобто, просторовою структурою молекул. Зокрема, для молекул виходить $C_V = 3R/2$ у випадку одноатомних молекул (i = 3), $C_V = 5R/2$ для двохатомних жорстких молекул (i = 5) і $C_V = 3R$ для багатоатомних жорстких молекул (i = 6). Досліди свідчать, що такий прогноз, більшою чи меншою мірою, справджується тільки при не дуже низьких і не дуже високих температурах. Це ілюструє рис. 5.1, де схематично показано температурну залежність теплоємності C_V водню H_2.

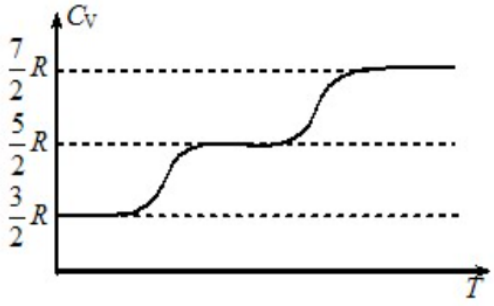


Рис. 5.1

При проміжних температурах маємо очікуване для жорстких двохатомних молекул значення $C_V = 5R/2$. Але при охолодженні газу теплоємність починає зменшуватись і при дуже низьких температурах набуває значення $3R/2$, яке притаманне одноатомному газу. Аналогічно, при значному нагріванні теплоємність зростає, наближаючись до величини $7R/2$, яка є характерною для двохатомних пружних молекул через наявність і коливального ступеня свободи. Пояснити ці розбіжності класична фізика не здатна, оскільки вони пов'язані зі специфічними квантовими властивостями молекул та інших мікроскопічних частинок. Як показано у квантовій механіці, енергія обертального та коливального руху мікрочастинок може змінюватися тільки дискретними порціями-квантами, причому енергія квантів обертальної енергії $\varepsilon_{об}$ набагато менша, ніж коливальної $\varepsilon_{кол}$. З іншого боку, при теплообміні молекули передають одна одній енергію порядку kT. Тому при дуже низьких температурах, коли $kT < \varepsilon_{об} << \varepsilon_{кол}$, обертальний рух молекул та коливальний рух атомів у них не впливають на процес теплообміну й, відповідно, на теплоємність газу. При цьому молекули поводяться, як одноатомні (i = 3, $C_V = 3R/2$), тому говорять, що

при низьких температурах обертальні та коливальні ступені свободи є "замороженими". При підвищенні температури обертальні ступені свободи поступово починають "розморожуватись", тобто, все більша кількість молекул H_2 починає давати вклад у внутрішню енергію за рахунок не тільки поступального, а й обертального руху. Внаслідок цього теплоємність поступово підвищується до величини $C_V = 5R/2$, що відповідає жорстким двохатомним молекулам (i = 5). Так само при дуже високих температурах "розморожується" й коливальний ступінь свободи, відтак, згідно з (4.5), $i = 5 + 2i_{кол} = 7$. Через це теплоємність зростає до значення $C_V = 7R/2$.

2. Політропні процеси

Окрім розглянутих раніше ізопрцесів у газах, важливу роль у термодинаміці відіграють адіабатний процес і політропні процеси, в яких змінюються всі три параметри стану: P, V, T.

Адіабатний процес. Так називається процес, який відбувається без теплообміну системи з навколишнім середовищем. Для такого процесу, згідно з (4.14), $\mathrm{d}U + P\mathrm{d}V = 0$. Тож, скориставшись виразом (5.11), отримаємо:

$$\frac{\mathrm{d}(PV)}{\gamma - 1} + p\mathrm{d}V = 0 \quad \Rightarrow \quad \mathrm{d}(PV) + (\gamma - 1)P\mathrm{d}V = 0 \quad \Rightarrow$$

$$V\mathrm{d}P + P\mathrm{d}V + (\gamma - 1)P\mathrm{d}V = 0 \quad \Rightarrow \quad V\mathrm{d}P + \gamma P\mathrm{d}V = 0.$$

Поділимо останній вираз на PV і врахуємо відоме з диференціального числення співвідношення $\mathrm{d}x/x = \mathrm{d}(\ln x)$. Тоді

$$\frac{\mathrm{d}P}{P} + \gamma\frac{\mathrm{d}V}{V} = 0 \quad \Rightarrow \quad \mathrm{d}\left(\ln P\right) + \gamma\mathrm{d}(\ln V) = 0 \quad \Rightarrow \quad \mathrm{d}(\ln P) + \mathrm{d}(\ln V^{\gamma}) = 0 \quad \Rightarrow$$

$$\Rightarrow \quad \mathrm{d}(\ln P + \ln V^{\gamma}) = 0 \quad \Rightarrow \quad \mathrm{d}(\ln PV^{\gamma}) = 0.$$

З цього виразу випливає рівняння адіабатного процесу, яке називається рівнянням Пуассона:

$$PV^{\gamma} = \mathrm{const} \tag{5.12}$$

За допомогою (2.6) це рівняння можна подати і в інших змінних:

$$TV^{\gamma-1} = \mathrm{const}; \tag{5.12а}$$

$$TP^{\frac{1-\gamma}{\gamma}} = \mathrm{const}. \tag{5.12б}$$

Зауважимо, що показником степеня в рівнянні Пуассона є введена в попередньому пункті величина $\gamma = C_P/C_V$, чим і пояснюється її назва "показник адіабати".

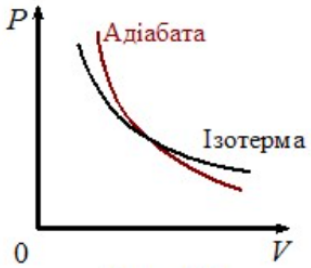


Рис. 5.2

Адіабатний процес нагадує ізотермічний тим, що тиск і об'єм знаходяться в оберненому зв'язку. Але в адіабатному процесі зміна тиску відбувається швидше, як це показано на рис. 5.2. Крім того, в адіабатному процесі $\delta Q = 0$, і, згідно з (4.13), робота газу виконується не за рахунок теплообміну, як в ізотермічному процесі, а за рахунок зміни внутрішньої енергії системи:

$$0 = \mathrm{d}U + \delta A \quad \Rightarrow \quad \delta A = -\mathrm{d}U \quad \Rightarrow \quad \mathrm{d}U = -P\mathrm{d}V.$$

Звідси випливає, що при адіабатному розширенні газ охолоджується, а при стисканні – нагрівається. Завдяки останньому, як уже відмічалося, відбувається самозапалювання паливної суміші в двигунах дизельного типу. Пониження температури газу при адіабатному розширенні використовується, зокрема, в кріогенній техніці та камері Вільсона – приладі, що дозволяє спостерігати траєкторії руху швидких елементарних частинок. При цьому слід зауважити, що в реальності процес у газі на може бути стого адіабатично через наявність теплообміну з довкіллям. Але при великій швидкості протікання, як у наведених прикладах, він є достатньо близьким до адіабатичного.

Але виявляється, що не будь-яке швидке розширення газу супроводжується його охолодженням. У цьому зв'язку повчальними є результати дослідів по розширенню газу в пустоту, які були здійснені Гей-Люссаком і Джоулем. У цих дослідах використовувалися два балони, що були з'єднані трубкою із закритим краном. Один балон був заповнений повітрям, а інший – відкачаний. Дослід полягав у вимірюванні температури газу до та після відкривання крана. На перший погляд результат досліду легко передбачити: після відкривання крана повітря розширюється дуже швидко, отже практично без теплообміну з оточенням (адіабатно), і тому має охолодитися. Але, навіть при високому початковому тиску, що в 20 разів перевищував атмосферний, ніякої зміни внутрішньої енергії та температури повітря при розширенні не спостерігалось. Цей "парадокс" не важко зрозуміти, взявши до уваги, що при вказаних вище значеннях тиску повітря ще поводиться, як ідеальний газ. Розширюючись *у пустоту*, повітря не виконує макроскопічної роботи (4.8 а), оскільки не рухає ніяких зовнішніх тіл типу поршня. Не виконується й мікроскопічна робота силами взаємодії між молекулами, оскільки їх в ідеальному газі немає. За таких умов внутрішня енергія, отже й температура повітря, при розширенні не змінюються, що й показали досліди Джоуля. Таким чином, при адіабатному розширенні ідеального газу в пустоту тиск і об'єм значно змінюються при незмінній температурі, що суперечить рівнянням адіабати (5.12 а) і (5.12 б). Але в цьому немає ніякого протиріччя, оскільки всі рівняння термодинаміки стосуються тільки рівноважних процесів, у той час як розширення газу в пустоту відбувається дуже швидко і не є рівноважним.

Політропні процеси. Теплоємність є функцією процесу і, загалом, може змінюватися в ході процесу. Але існують процеси, в яких теплоємність системи лишається сталою: $C = const$. Такі процеси називаються ***політропними*** (говорять також "політропічними"). Установимо рівняння політропного процесу. Для цього запишемо перше начало термодинаміки (4.14) для одного моля газу, виразивши δQ і $\mathrm{d}U$ з (5.1) та (5.3):

$$C\mathrm{d}T = C_V\mathrm{d}T + P\mathrm{d}V \quad \Rightarrow \quad (C - C_V)\mathrm{d}T - P\mathrm{d}V = 0.$$

З рівняння Клапейрона (2.6 а) випливає, що $\mathrm{d}T = \mathrm{d}(PV)/R$. Зробивши цю підстановку в попередній вираз і, ввівши позначення $\alpha = (C - C_V)/R$, отримаємо: $\alpha\mathrm{d}(PV) - P\mathrm{d}V = 0$. Далі, вчиняючи так, як раніше при виведенні рівняння Пуассона, дістанемо:

$$PV^{1-\frac{1}{\alpha}} = \mathrm{const}.$$

Якщо, підставити в показнику вираз α і врахувати, що $R = C_P - C_V$, то отримаємо ***рівняння політропи***:

$$PV^n = \mathrm{const}, \qquad (5.13)$$

де число

$$n = \frac{C - C_P}{C - C_V} \tag{5.14}$$

називається ***показником політропи***.

Рівняння (5.13) аналогічне до (5.12), тому за аналогією з (5.12 а) і (5.12 б) можна записати:

$$TV^{n-1} = \text{const} \quad \text{і} \quad TP^{\frac{1-n}{n}} = \text{const}. \tag{5.14a}$$

До політропних належать усі розглянуті раніше процеси. Справді, ізобарний процес $P = const$ задовольняє рівнянню (5.13) при $n = 0$, для ізотермічного процесу $PV = const$ і $n = 1$, для адіабатного– $n = \gamma$. Нарешті, якщо (5.13) переписати у вигляді $P^{1/n}V = const$, стане зрозуміло, що ізохорний процес $V = const$ теж формально можна розглядати як політропний з $n \to \infty$. Формула (5.14) дозволяє виразити теплоємність системи в будь-якому політропному процесі через показник політропи:

$$C = \frac{nC_V - C_P}{n - 1}. \tag{5.15}$$

З цієї формули випливає, що в ізотермічному процесі (n =1) теплоємність $C \to \infty$, що природно, оскільки для цього процесу у виразі (5.1) $\mathrm{d}T = 0$. Для адіабатного процесу $n = \gamma$ і з (5.15) виходить $C = 0$. Це теж зрозуміло, адже в цьому процесі $\delta Q = 0$. Але для всіх політропних процесів, проміжних між ізотермічним та адіабатним, із формули (5.15) випливає, що $C < 0$. Щоб переконатися в цьому досить у (5.15) зробити заміну $C_P = \gamma C_V$ і врахувати, що для таких процесів n < γ. Цей висновок може здатися несподіваним, але в ньому немає нічого парадоксального. Він означає тільки те, що в подібних процесах газ весь час виконує більшу роботу, ніж отримувана ним кількість теплоти. З цієї причини внутрішня енергія зменшується і газ охолоджується, тож $\mathrm{d}T < 0$ і $C < 0$.

Робота газу в політропних процесах. Найзагальнішим способом обчислення роботи в термодинаміці є визначення інтегралів (4.9). Але для політропного процесу із заданим показником політропи це можна зробити легше за допомогою першого начала термодинаміки (4.13а), згідно з яким $A = Q - \Delta U$. Виразивши Q і ΔU через теплоємності за формулами (5.1) і (5.3), отримаємо для одного моля $A = (C - C_V)\Delta T$ і для ν молів

$$A = \nu(C - C_V)\Delta T. \tag{5.16}$$

Різницю теплоємностей легко виразити через показник політропи з (5.14):

$$n = \frac{C - C_P}{C - C_V} = \frac{(C - C_V) + C_V - C_P}{C - C_V} = 1 - \frac{R}{C - C_V} \quad \Rightarrow \quad C - C_V = -\frac{R}{n - 1}.$$

Підставивши цей вираз у (5.16), отримаємо

$$A = \frac{\nu R(T_1 - T_2)}{n - 1}, \tag{5.17}$$

або, з урахуванням (2.6),

$$A = \frac{P_1V_1 - P_2V_2}{n - 1}. \tag{5.17a}$$

Слід зауважити, що ці вирази не можна використовувати для ізотермічного процесу (n = 1), оскільки і чисельники, і знаменники дорівнюють нулю.

3. Процес Джоуля-Томсона

Усі конкретні питання, пов'язані з першим началом термодинаміки, розглядалися вище на прикладі ідеальних газів. Термодинаміка процесів у реальних газах є значно складнішою. Це зумовлено тим, що у внутрішню енергію реального газу, крім кінетичної, входить ще й потенціальна енергія взаємодії між молекулами $U_{вз}$. Остання залежить від відстані між молекулами, тобто, від об'єму газу. Тому при адіабатній (без обміну енергією із зовнішніми тілами) зміні об'єму реального газу має змінюватись не тільки потенціальна, а й кінетична енергія молекул, отже – температура газу. Причому, на відміну від ідеального газу, в реальних газах температура має змінюватись і при розширенні в пустоту, коли зміна об'єму газу не супроводжується виконанням роботи.

Ефект зміни температури при адіабатному розширенні газу в пустоту виявили Джоуль і Томсон, які здійснили дослід за схемою рис. 5.3. Теплоізольована труба розділяється на дві половини "дроселем" Др (пористим корком), по обидва боки від якого розміщені рухомі нетеплопровідні поршні П1 і П2. У початковому стані поршень П2 прилягає до дроселя, а об'єм між поршнем П1 і дроселем заповнений газом. Зовнішні тиски на поршні однакові і дорівнюють тиску газу між поршнем П1 і дроселем. Дослід полягає в повільному продавлюванні газу крізь дросель за рахунок невеликої додаткової сили, прикладеної до поршня П1. При цьому обидва поршні переміщуються з положень 1, 2 в положення 1′, 2′. Такий процес називається "*дроселюванням*", або *процесом Джоуля-Томсона*.

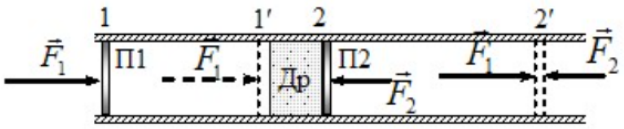


Рис. 5.3

При повільному переміщенні поршнів дроселювання можна вважати рівноважним (квазістатичним) процесом, так що тиск і температура газу в кожній половині труби (P_1, T_1) і (P_2, T_2) лишаються сталими, причому $P_1 > P_2$. Застосуємо до цього процесу перше начало термодинаміки (4.12), урахувавши відсутність теплообміну:

$$\Delta U = A',$$

де робота A' складається з робіт A'_1 і A'_2, що їх виконують над газом сили зовнішнього тиску на поршні:

$$A' = A'_1 + A'_2.$$

Позначивши початковий об'єм зліва V_1 і кінцевий об'єм праворуч V_2, на основі (4.10а) маємо: $A'_1 = P_1 V_1$ і $A'_2 = -P_2 V_2$. Відтак зміна внутрішньої енергії виражається, як

$$\Delta U = A' = P_1 V_1 - P_2 V_2 \quad \Rightarrow \quad \Delta U = -\Delta(PV). \qquad (5.18)$$

Без урахування енергії взаємодії між молекулами $U_{вз}$, тобто, за умови ідеальності, внутрішня енергія одного моля газу, згідно з (5.7), $U = C_V T$. Отже, для моля реального газу

$$U = C_V T + U_{вз}.$$

Підставивши цей вираз в (5.18), отримаємо:

$$\Delta(C_V T + U_{\text{вз}}) = -\Delta(PV) \quad \Rightarrow$$
$$C_V \Delta T + \Delta U_{\text{вз}} = -\Delta(PV) \quad \Rightarrow \quad \Delta T = -\frac{\Delta U_{\text{вз}} + \Delta(PV)}{C_V}. \qquad (5.19)$$

Таким чином, при описаному розширенні в пустоту реального газу температура має змінюватись. Це явище, яке дійсно спостерігається, називається *ефектом Джоуля-Томсона*, а співвідношення (5.19) є його загальним аналітичним виразом. Для отримання конкретного виразу ΔT треба мати конкретну модель газу, тобто, явні вирази енергії взаємодії між молекулами та рівняння стану газу. Але деякі особливості ефекту Джоуля-Томсона можна з'ясувати й на основі загальних міркувань. По-перше, в процесі Джоуля-Томсона поршні виконують зовнішню роботу A_1' і A_2'. Тому може здатися, що саме ця робота є причиною зміни температури, і ніякого "ефекту", власне, немає. Аби відкинути такий сумнів, застосуємо вираз (5.19) до ідеального газу, для якого $\Delta U_{\text{вз}} = 0$ і, згідно з (2.6а), $\Delta(PV) = R\Delta T$. Тоді вийде:

$$\Delta T = -\frac{R\Delta T}{C_V} \quad \Rightarrow \quad (C_V + R)\Delta T = 0 \quad \Rightarrow \quad \Delta T = 0.$$

Отже, зміна температури в ефекті Джоуля-Томсона зумовлена саме "неідеальністю" газу. По-друге, для різних газів і при різних тисках і температурах кожен доданок і весь чисельник (5.19) можуть мати різні знаки. Тому температура газу при дроселюванні може як зменшуватися, так і збільшуватися, а в окремих випадках і лишатися незмінною. Практика підтверджує такий прогноз: для кожного газу існує своя *температура інверсії*, при переході через яку ефект змінює знак.

Наостанок, укажемо, що можливість зменшення температури при дроселюванні використовують у промисловості для скраплення газів і отримання дуже низьких ("кріогенних") температур.

Контрольні запитання

1. Що означають терміни "теплоємність", "молярна теплоємність" та "питома теплоємність"? Як пов'язані між собою ці величини?
2. Чи можна твердити, що теплоємність визначається тільки властивостями речовини тіла?
3. Чому в довідниках фізичних характеристик матеріалів є таблиці теплоємностей для рідин і твердих тіл, але немає – для газів?
4. Чи може теплоємність системи бути від'ємною?
5. Що таке ізохорна теплоємність C_V та ізобарна теплоємність системи C_P? Чому ізобарна теплоємність більша за ізохорну? На скільки?
6. Що таке показник адіабати? Отримайте формули (5.9) – (5.11).
7. Який процес називається адіабатним? Наведіть різні варіанти запису рівняння адіабати.
8. Як змінюється температура газу в адіабатному процесі? Чи може зменшуватися температура при стисканні газу?
9. Відомо, що при швидкому (адіабатному) розширенні ідеального газу в пустоту його температура не змінюється. Чому?
10. Які процеси називаються політропними? Як пов'язані між собою тиск та об'єм газу в політропному процесі?

11. Отримайте рівняння політропи (5.13) і (5.14а).

12. У чому полягає ефект Джоуля-Томсона? Чому він не спостерігається в ідеальних газах?

---

[1] Виняток становить тільки ізотермічний процес, у якому внутрішня енергія та температура лишаються незмінними.

[2] Це можна розглядати як термодинамічне означення універсальної газової сталої.

6. Явища переносу

Розглянуті раніше відомості стосувалися рівноважних станів і процесів, але за відповідних умов їх можна використовувати й для нерівноважних систем. В різних частинах нерівноважної замкненої системи певні параметри стану (температура, тиск, тощо) та інші характеристики системи (приміром, концентрація молекул даного сорту в суміші) мають не однакові значення. У процесі встановлення термодинамічної рівноваги відбувається вирівнювання подібних неоднорідностей, яке супроводжується виникненням потоків відповідних фізичних характеристик системи. Скажімо, якщо в одній частині системи температура вища, ніж в іншій, то від більш нагрітої частини до менш нагрітої буде переноситися теплова енергія, при вирівнюванні концентрації частинок від області з більшою їхньою концентрацією до області з меншою буде переноситися маса, тощо. Тому подібні явища називаються ***явищами переносу***.

Явища переносу спостерігаються і в незамкнених системах. При цьому, якщо неоднорідність розподілу характеристик системи штучно підтримувати на незмінному рівні, перенесення відповідних величин буде стаціонарним, тобто, їхні значення не залежитимуть від часу, так само, як у рівноважній системі. Тому для стаціонарних явищ переносу є чинними закони молекулярно-кінетичної теорії та термодинаміки.

Нижче розглянуті наступні питання:

1. Емпіричні закони явищ переносу

2. Середня довжина вільного пробігу молекул

3. Молекулярно-кінетична теорія явищ переносу в газах

Контрольні запитання

1. Емпіричні закони явищ переносу

Потік і густина потоку. Кількісно інтенсивність перенесення деякої величини A характеризують ***потоком* Ф,** який визначається як кількість величини А, що переноситься через задану поверхню за одиницю часу:

$$\Phi = \frac{\Delta A}{\Delta t}, \qquad (6.1)$$

де ΔA – кількість величини A, що переноситься крізь цю поверхню за час Δt.

У загальному випадку потік Ф може бути неоднорідним, коли в різних точках інтенсивність перенесення є не однаковою. Тому для опису перенесення величини A в кожній точці простору вводять ***густину потоку***

$$j = \frac{\Delta \Phi}{\Delta S_{\perp}}, \qquad (6.1а)$$

де $\Delta\Phi$ – потік крізь елементарну площинку $\Delta S_{\perp}$, перпендикулярну до напрямку переносу.

Перенесення величини A вздовж певного напряму завжди є зумовлене неоднорідним розподілом у цьому напрямі якоїсь іншої величини B. Наприклад, перенесення теплоти Q від батареї опалення до якоїсь віддаленої точки в кімнаті зумовлене наявністю перепаду температури T між батареєю та цією точкою. Мірою неоднорідності величини B у заданому напрямку ОХ є її ***градієнт***, тобто швидкість зміни $\mathbf{d}B/\mathbf{d}x$ у цьому напрямі.

Розглянуті поняття й величини дозволяють аналітично виразити емпіричні (встановлені на досліді) закони стаціонарних явищ переносу, що розглядаються далі.

Теплопровідність. Уявімо дві паралельні плоскі пластини, розміщені в однорідному теплопровідному середовищі, котрі підтримуються при сталих температурах T_1 і T_2 ($T_1 > T_2$). За таких умов між пластинами у показаному на рис. 6.1 напрямі OX встановлюється стаціонарне перенесення тепла[1].

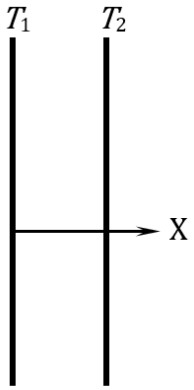


Рис. 6.1

Дослід показує, що густина теплового потоку в кожній точці є прямо пропорційною градієнтові температури в цій точці, що відображається ***рівнянням теплопровідності,*** або ***рівнянням Фур'є***:

$$j_Q = -\varkappa \frac{\mathrm{d}T}{\mathrm{d}x}, \qquad (6.2)$$

де коефіцієнт пропорційності κ ***коефіцієнтом теплопровідності***

Відповідно, кількість теплоти, що переноситься через ділянку перерізу потоку площею ΔS за проміжок часу Δt дорівнює :

$$\Delta Q = -\varkappa \frac{\mathrm{d}T}{\mathrm{d}x} \Delta S \Delta t. \qquad (6.3)$$

(Примітка. Знак "–" у виразах (6.2) і (6.3) є зумовлений тим, що перенесення будь-якої величини (тут Q) завжди відбувається у напрямку менших значень тієї величини (тут T), градієнт якої є причиною переносу. Тож у напрямку перенесення тепла $Q > 0$, тоді як

Дифузія. Так називається взаємопроникнення різних речовин у сумішах, якщо концентрації частинок у них нерівномірно розподілені по об'єму. До прикладу, проявом дифузії є поширення запаху парфумів у повітрі, або розширення області забарвлення при вміщенні у воду чи на фільтрувальний папір краплі рідкого барвника. Дифузія спостерігається й у хімічно однорідній речовині (газі), якщо концентрація частинок різна в різних місцях. У такому випадку говорять про самодифузію.

Імовірність теплового руху окремої частинки за всіма напрямками однакова, тому будь-якої миті кількість частинок, які рухаються в бік меншої концентрації, більша, ніж кількість частинок, які рухаються в зустрічному напрямку. Через це виникає потік частинок у напрямку зони меншої концентрації. Якщо в газі підтримується сталий градієнт концентрації частинок у напрямку певної осі OX, то, як свідчить дослід, сумарна кількість частинок, які проходять за час Δt крізь площинку ΔS, перпендикулярну до напрямку потоку, дорівнює

$$N = -D \frac{\mathrm{d}n}{\mathrm{d}x} \Delta S \Delta t, \qquad (6.4)$$

де $\mathrm{d}n/\mathrm{d}x$ – градієнт концентрації, і D – величина, що залежить від індивідуальних властивостей частинок речовини і називається ***коефіцієнтом дифузії***. Відповідно, густина потоку частинок $j_n = N/(\Delta S \Delta t)$ визначається рівнянням

$$j_n = -D\frac{\mathrm{d}n}{\mathrm{d}x}, \tag{6.5}$$

яке називається ***рівнянням дифузії,*** або ***рівнянням Фіка***.

Якщо обидві частини рівняння (6.5) помножити на масу однієї молекули m_0 і врахувати, що $nm_0 = \rho$ – густина речовини, то в правій частині отримаємо градієнт густини $\mathrm{d}\rho/\mathrm{d}x$, а в лівій – густину потоку маси $j_m = m/(\Delta S \cdot \Delta t)$, де $m = m_0 N$ – маса речовини, що переноситься. Відтак рівняння (6.5) трансформується в рівняння густини масового потоку:

$$j_m = -D\frac{\mathrm{d}\rho}{\mathrm{d}x}. \tag{6.5a}$$

Внутрішнє тертя (в'язкість). До класичних явищ переносу відноситься також внутрішнє тертя, яке спостерігається в рухомій рідині або газі, коли сусідні шари речовини мають різні швидкості течії u_1 і u_2, як показано на рис. 6.2.

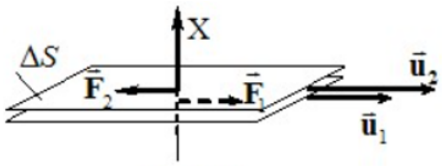


Рис. 6.2

В цьому випадку частинки речовини за рахунок теплового хаотичного руху переходять з одного шару в інший, переносячи із собою відповідний імпульс упорядкованого руху. Унаслідок такого обміну повільний шар буде поповнюватися молекулами, що рухаються у напрямку течії з більшою швидкістю u_2, а швидкий – молекулами, що мають меншу швидкість u_1. Тому повільний шар буде прискорюватись, а швидкий – сповільнюватись. Це означає, що між шарами є певні сили взаємодії $\vec{F}_1$ і $\vec{F}_2$, які називаються силами внутрішнього тертя. На досліді встановлено, що величина сили внутрішнього тертя залежить від індивідуальних властивостей речовини, площі взаємодіючих шарів ΔS та градієнту їхньої швикості й визначається, як

$$F = \eta\left|\frac{\mathrm{d}u}{\mathrm{d}x}\right|\Delta S. \tag{6.6}$$

Поділивши цей вираз на ΔS, отримаємо ***рівняння Ньютона***, яке визначає силу внутрішнього тертя між шарами одиничної площі:

$$f = \eta\left|\frac{\mathrm{d}u}{\mathrm{d}x}\right|. \tag{6.6a}$$

Похідна під знаком модуля є градієнтом швидкості, який показує, як стрімко змінюється швидкість руху шарів рідини чи газу при переході від одного до іншого шару. Величина η називається ***коефіцієнтом внутрішнього тертя*** або ***динамічною в'язкістю*** речовини. Крім динамічної, розглядають і так звану кінематичну в'язкість

$$\nu = \frac{\eta}{\rho}, \tag{6.7}$$

де ρ - густина речовини.

Сила внутрішнього тертя $\vec{f}$ виникає внаслідок перенесення імпульсу від швидких шарів рідини чи газу до повільних. За другим законом Ньютона $f = |\mathrm{d}\vec{p}/\mathrm{d}t|$, причому в даному випадку права частина – то є імпульс, який переноситься за одиницю часу через одиничну площинку, тобто, густина потоку імпульсу j_p. Тому, з урахуванням напрямку переносу, рівняння (6.6) можна, подібно до (6.2) – (6.5а), подати як рівняння переносу імпульсу:

$$j_p = -\eta \frac{\mathrm{d}u}{\mathrm{d}x}. \qquad (6.8)$$

2. Середня довжина вільного пробігу молекул

Вирівнювання характеристик ізольованої системи по всьому її об'єму зумовлюється тим, що молекули неперервно стикаються одна з одною й переходять з одних місць в інші. При цьому кожна молекула між двома послідовними зіткненнями проходить певну відстань, яка називається довжиною вільного пробігу. Через хаотичність і велику швидкість теплового руху ця відстань для кожної молекули дуже часто й непрогнозовано змінюється від зіткнення до зіткнення. Тому мірою вільного руху молекул у системі є ***середня довжина вільного пробігу λ.*** Оцінимо, чим вона визначається в ідеальному газі. Для цього будемо розглядати молекули як пружні кульки, й уявимо спочатку, що рухається тільки одна молекула, а всі інші перебувають у спокої. За проміжок часу Δt ця молекула проходить середню відстань $\Delta l = \langle v \rangle \Delta t$, де $\langle v \rangle$ – середня арифметична швидкість молекули. При цьому дана молекула стикається з усіма тими молекулами, центри котрих потрапляють у циліндр висоти Δl і радіуса d, рівного діаметру молекули, як показано на рис. 6.3.

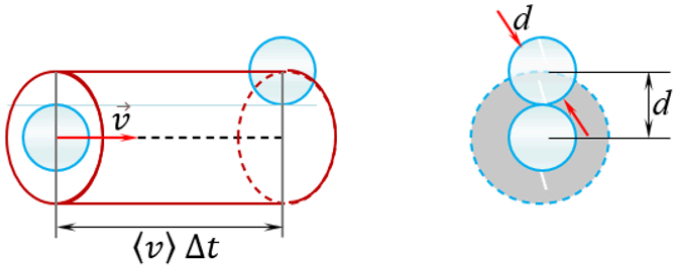


Рис. 6.3

За таких умов середня кількість зіткнень $< z' >$ на довжині Δl (тобто, за час Δt) дорівнює кількості молекул у вказаному циліндрі: $< z' > = n\pi d^2 \Delta l$, де n – концентрація молекул, $\sigma = \pi d^2$ – так званий ефективний переріз молекули[2]. В дійсності рухається не одна, а всі молекули, тому ймовірність зіткнення даної молекули з якоюсь іншою визначається їхньою середньою *відносною* швидкістю, яка є в $\sqrt{2}$ разів більша за середню арифметичну швидкість $\langle v \rangle$.[3] Тому коректнішою оцінкою кількості зіткнень є величина

$$z = \sqrt{2} n \pi d^2 \Delta l = \sqrt{2} n \sigma \Delta l. \qquad (6.9)$$

Якщо на відстані Δl відбувається z зіткнень даної молекули з іншим, то середня довжина вільного пробігу $\lambda = \Delta l / z$ і, згідно з (6.9),

$$\lambda = \frac{1}{\sqrt{2}\pi\sigma n} = \frac{1}{\sqrt{2}\pi d^2 n}. \qquad (6.10)$$

Концентрація молекул пов'язана з тиском газу співвідношенням $P = nkT$. Тому середня довжина вільного пробігу молекул обернено пропорційна тиску газу, тож

$$\lambda_1 P_1 = \lambda_2 P_2. \qquad (6.11)$$

Середню довжину вільного пробігу молекул у газах можна визначати експериментально на основі явищ переносу. Тому з виразу (6.10) можна знайти ефективний переріз і ефективний діаметр молекули. Щодо точності таких вимірів слід зауважити наступне. По-перше, при підрахунку кількості зіткнень z (6.9) вважалося, що рухома молекула стикається з іншими молекулами в об'ємі *прямого* циліндра довжини Δl, тобто, що вона при зіткненнях не змінює напрямку руху. Насправді це не так. Тому указаний циліндр є ламаним і має дещо більший об'єм за рахунок "колін". По-друге, молекули розглядались як кулі, хоча в дійсності довільна молекула складається з декількох атомів і не має симетрії кулі. Тому вираз (6.10) слід розглядати як оціночний. Але, попри це, уявлення про ефективний переріз та ефективний діаметр молекули виявляється досить плідним і загально вживаним.

Наостанок зауважимо, що із середньою довжиною вільного пробігу молекул пов'язане поняття вакуумного стану газу або технічного вакууму. При відкачуванні закритої посудини з газом середня довжина вільного пробігу в міру зниження тиску зростає й при певному тиску P_0 стає рівною чи більшою за лінійні розміри посудини. Такий стан газу називається ультрарозрідженим або вакуумним. У вакуумному стані молекули газу практично не стикаються між собою, а тільки зі стінками посудини. Тому властивості ультрарозрідженого газу істотно відрізняються від його властивостей за звичайних умов. Наприклад, при високому вакуумі в газі практично відсутнє внутрішнє тертя та теплопровідність. Останнє використовують у посудинах Дьюара – посудинах із подвійними стінками, простір між якими відкачано до високого вакууму. Такі посудини забезпечують добру теплову ізоляцію від зовнішнього середовища. Прикладом використання посудин Дьюара в побуті є добре відомий кожному термос. В науці та техніці високоякісні посудини Дюара використовують, зокрема, для зберігання скраплених неорганічних газів (азот, кисень, водень, гелій), які мають дуже низькі (кріогенні) температури.

3. Молекулярно-кінетична теорія явищ переносу в газах

Розглянуті явища переносу спостерігаються в будь-яких речовинах мають спільні механізми і, як наслідок, виражаються схожими співвідношеннями (6.2) – (6.8). Але побудувати їхню кількісну теорію на основі молекулярно-кінетичних уявлень можна лише для газів, у яких практично відсутні сили зчеплення між молекулами.

Дифузія. Нехай у газі підтримується незалежний від часу нерівномірний розподіл концентрації молекул у напрямку осі OX, як показано на рис. 6.4.

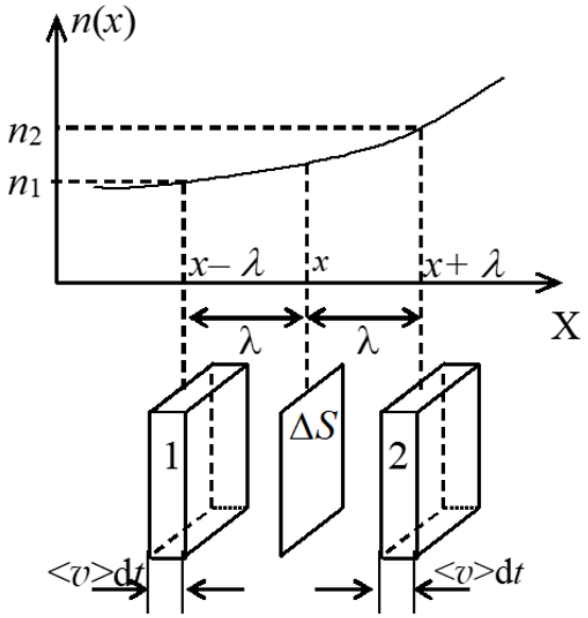


Рис. 6.4

За таких умов у напрямі осі OX створюється стаціонарний потік молекул. Визначимо густину цього потоку в довільній точці x. Для цього виділимо в газі два тонкі шари 1 і 2 на відстані вільного пробігу молекул λ від точки x, для кожного з яких площа складає ΔS і товщина $h << \lambda$. В такому разі можна вважати, що крізь площинку ΔS проходять усі молекули кожного шару, котрі вилітають в її напрямку паралельно до осі OX. Також, ураховуючи рівноймовірність усіх напрямків теплового руху молекул, будемо вважати, що в кожному з двох можливих напрямів кожної з трьох координатних осей рухається (1/6) частина всіх молекул газу. Для виділених шарів це складає, відповідно, $N_1 = (1/6)n_1\Delta Sh$ і $N_2 = (1/6)n_2\Delta Sh$. Проліт усіх цих молекул крізь площинку ΔS триває протягом часу $\Delta t = h/\langle v\rangle$, який потрібен, аби найвіддаленіші з них подолали відстань h між стінками шару при середній швидкості руху $\langle v\rangle$. Відтак для густини потоку молекул отримуємо:

$$j = \frac{1}{6}n\langle v\rangle \qquad (6.12)$$

Відтак, згідно з означеннями (6.1а) і (6.1а) , густина зустрічних потоків молекул крізь площинку ΔS виражається, відповідно, як $j_1 = (1/6)n_1\langle v\rangle$ і $j_2 = (1/6)n_2\langle v\rangle$, а результуюча густина потоку дифузії

$$j_n = \frac{1}{6}(n_1 - n_2)\langle v\rangle = -\frac{1}{6}\Delta n\,\langle v\rangle,$$

де $\Delta n = n_2 - n_1$ — приріст концентрації молекул уздовж осі OX на відстані $\Delta x = 2\lambda$.

Позаяк величина λ є дуже малою, то за канонами математичного аналізу можна записати $\Delta n = (\mathrm{d}n/\mathrm{d}x)\Delta x = 2\lambda(\mathrm{d}n/\mathrm{d}x)$[4].

Відтак для густини потоку частинок при самодифузії остаточно отримаємо:

$$j_n = -\frac{1}{3}\lambda\langle v\rangle\frac{\mathrm{d}n}{\mathrm{d}x}, \qquad (6.13)$$

де $(\mathrm{d}n/\mathrm{d}x)$ — градієнт концентрації молекул, а множник

$$D = \frac{1}{3}\lambda\langle v\rangle \qquad (6.13а)$$

— коефіцієнт дифузії.

Таким чином, рівняння (6.13) повністю узгоджується з емпіричним рівнянням (6.5) та теоретично визначає коефіцієнта дифузії D ідеального газу.

Теплопровідність. Як говорилося в п.1, в середовищі між двома паралельними площинами, що підтримуються при сталих температурах T_1 і T_2, встановлюється певний не залежний від часу розподіл температур уздовж перпендикулярної до площин осі OX (рис. 6.5) та стаціонарний тепловий потік, зумовлений теплопровідністю, тобто не пов'язаний з рухом речовини (конвекцією).

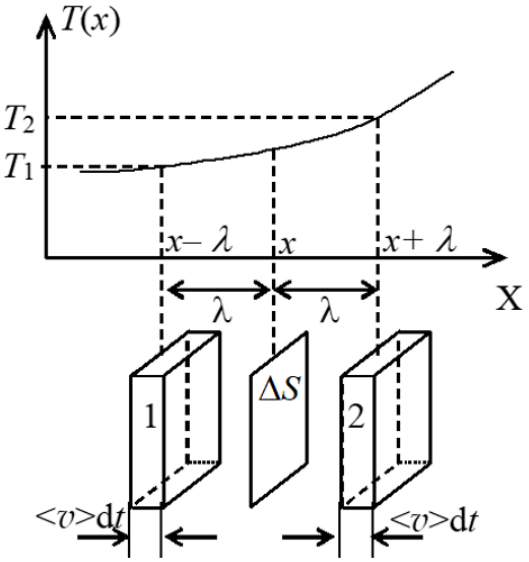


Рис. 6.5

Теоретично закономірності цього процесу в ідеальному газі легко встановити за тією самою формальною схемою, що й у випадку дифузії, врахувавши, що при теплопровідності перенесення тепла спричинюється градієнтом температури, а не концентрації молекул. Тож знову виділимо в газі обабіч площинки ΔS на відстанях вільного пробігу молекул λ два шари 1 і 2 з такою самою площею і товщиною $h << \lambda$. Тоді можна вважати, що крізь площинку ΔS проходять усі молекули з кожного шару, котрі рухаються в її напрямі паралельно до осі OX. При цьому молекули з шару 1 є більш "гарячими", ніж із шару 2, тож переносять із собою більшу кінетичну енергію E тепловового руху. Тому зустрічний рух обох пучків молекул супроводжується перенесенням теплові енергії від гарячих до холодних зон, яку легко розрахувати.

Очевидно, що густина потоку енергії крізь площинку ΔS, що створюється кожним пучком молекул, дорівнює $j = j_м E$, де $j_м$ – густина потоку молекул, і E – середня енергія теплового руху *однієї молекули* в пучку. Тож густина теплового потоку $j_q = j_1 - j_2$, що створюється крізь площинку ΔS при зустрічному русі обох пучків, дорівнює

$$j_q = j_м(E_1 - E_2)$$

Енергія полекули визначається її кількістю ступенів свободи $\boldsymbol{i}$ та температурою газу за формулою (4.4): $E = (i/2)kT$. Отже,

$$E_1 - E_2 = \frac{ik}{2}(T_1 - T_2) = -\frac{ik}{2}\Delta T$$

Відтак, урахувавши, що при малій відстані між шарами $\Delta x = 2\lambda$ зміна температури $\Delta T = (\mathrm{d}T/\mathrm{d}x)\Delta x = 2\lambda(\mathrm{d}T/\mathrm{d}x)$[4], а густина потоку молекул виражається формулою (6.12), отримуємо наступний вираз густини теплового потоку при теплопровідності:

$$j_q = -\frac{1}{3}\lambda\langle v\rangle\frac{ikn}{2}\cdot\frac{\mathrm{d}T}{\mathrm{d}x}.$$

Множник $ikn/2$ можна подати зручніше, домноживши чисельник і знаменник на сталу Авогадро N_A та масу молекули m_0. Тоді, враховуючи, що $m_0 n = \rho$ – густина, $m_0 N_A = M$– молярна маса газу, а також формули (2.5), (5.2) і (5.6), отримаємо:

$$j_q = -\frac{1}{3}\lambda\langle v\rangle \rho c_V \cdot \frac{\mathrm{d}T}{\mathrm{d}x}. \qquad (6.14)$$

Це і є теоретичне рівняння теплопровідності, причому коефіцієнт теплопровідності визначається виразом

$$\kappa = \frac{1}{3}\lambda\langle v\rangle \rho c_V, \qquad (6.15)$$

де ρ - густина газу, c_V - питома теплоємність при сталому об'ємі.

***Внутрішнє тертя* (*в'язкість*).** Як відмічалося раніше, внутрішнє тертя між шарами газу є зумовлене переносом імпульсу впорядкованого руху від шару до шару в напрямку зменшення швидкості (від'ємний напрям осі X на рис. 6.2, що безпосередньо відображено в рівнянні (6.8). Молекулярний механізм в'язкості є повністю аналогічним до щойно розглянутого механізму теплопровідності. Тому для отримання теоретичного рівняння внутрішнього тертя в газі, не повторюючи всіх попередніх міркувань, достатньо заміни у виразах густини потоків енергію E на імпульс упорядкованого руху молекули $p = m_0 u$. Тоді густина потоку імпульсу крізь площинку ΔS дорівнює:

$$j_p = \frac{1}{6}n\langle v\rangle m_0(u_1 - u_2) = -\frac{1}{6}n\langle v\rangle m_0 \Delta u \Rightarrow j_p = -\frac{1}{3}\lambda\langle v\rangle n m_0 \frac{\mathrm{d}u}{\mathrm{d}x}.$$

Урахувавши, що $m_0 n = \rho$ – густина газу, отримаємо теоретичне рівняння густини потоку імпульсу:

$$j_p = -\frac{1}{3}\lambda\langle v\rangle \rho \frac{\mathrm{d}u}{\mathrm{d}x}, \qquad (6.16)$$

а відтак і рівняння Ньютона (6.6 а), з таким виразом коефіцієнта внутрішнього тертя:

$$\eta = \frac{1}{3}\lambda\langle v\rangle \rho. \qquad (6.17)$$

Із виразів (6.13). (6.15) і (6.17) випливає, що, визначаючи коефіцієнти переносу дослідним шляхом, можна знайти середню довжину пробігу молекул, і, згідно з (6.10), ефективний переріз та діаметр молекули – параметри, які відіграють важливу роль у багатьох задачах молекулярної фізики. Тому цікаво проаналізувати, наскільки розглянута елементарна теорія узгоджується з експериментом.

Порівняння теоретичних та емпіричних рівнянь переносу показує, що теорія правильно відображає лінійний характер залежності густини потоку величини, які переноситься, від градієнту відповідного термодинамічного параметра газу. Розглянута теорія також адекватно показує, від яких величин залежать коефіцієнти переносу D, η і κ. Зокрема те, що всі коефіцієнти переносу повинні зростати при підвищенні температури, оскільки $\langle v\rangle \sim \sqrt{T}$. Інший теоретичний прогноз полягає в тому, що коефіцієнти теплопровідності (6.15) та внутрішнього тертя (6.17) газів не залежать від тиску, оскільки густина $\rho = nm_0$ є прямо пропорційною, а довжина вільного пробігу молекул λ– обернено пропорційною до концентрації молекул, а отже, й тиску газу. Дослід це підтверджує, що може здатися дивним, – адже зменшуючи концентрацію молекул, ми зменшуємо кількість "переносників" енергії у випадку теплопровідності та імпульсу у випадку внутрішнього тертя. Це справді так, але при тому, такою ж мірою збільшується довжини вільного пробігу молекул, тож і різниця значень величини, що переноситься кожною парою зустрічних молекул, а отже й

густина потоку. Водночас, дійсна величина числових множників у коефіцієнтах переносу D, η і κ дещо відрізняється від 1/3 і є не однаковою для різних явищ переносу. Це пояснюється тим, що при виведенні розглянутих рівнянь відповідні процеси в газі розглядалися доволі спрощено.

Контрольні запитання

1. Які процеси називаються явищами переносу? За яких умов вони спостерігаються?
2. За яких умов до явищ переносу можна застосовувати положення молекулярної фізики рівноважних процесів?
3. Що таке потік даної фізичної величина та густина потоку? Який зв'язок існує між ними?
4. Чи супроводжується тепловий потік переміщенням шарів газу від більш нагрітої області до менш нагрітої?
5. При прокачуванні рідини по трубі швидкість її руху є найбільшою на осі труби й найменша біля стінки. Який напрям має градієнт швидкості зумовлений цим потік імпульсу?
6. Що таке вакуумний стан газу? Якими мають бути розміри посудини, щоб газ у ній перебував у вакуумному стані?
7. Який "принцип дії" термоса? Оцініть, при якому максимальному тиску термос із відстанню між стінками 5 мм ще буде виконувати свою функцію – забезпечувати теплову ізоляцію вмісту від довкілля.

---

[1] Самі пластини не показані й розташовані перпендикулярно до площини рисунка.

[2] Як видно, ефективний переріз молекули дорівнює площі круга, в який має потрапити центр іншої молекули, щоби вони зіткнулися, тобто, достатньо сильно змінили свої швидкості внаслідок взаємодії. Діаметр та ефективний переріз молекули є певною мірою умовними величинами, зокрема, вони залежать від температури.

[3] Див. Розділ III, п. 2.

[4] Це випливає з того, що при малій зміні аргумента похідну можна наближено розглядати як відношення приросту функції до приросту аргумента: $(\mathrm{d}n/\mathrm{d}x) \approx (\Delta n/\Delta x)$

7. Друге начало термодинаміки

Серед різноманітних трансформацій енергії особливе місце посідають взаємні перетворення механічної та теплової енергії. Перетворення енергії руху тіл на внутрішню (теплову) енергію відбувається в будь-якій реальній механічній системі через наявність сил тертя та опору середовища. Кожному відоме відчуття тепла при потиранні рук або розтиранні тіла грубим рушником. За допомогою інтенсивного тертя можна навіть видобувати вогонь, як це робили прадавні люди. У свою чергу, як це випливає з першого начала термодинаміки, теплова енергія може перетворюватися на механічну. Теоретична можливість цього доводиться першим началом термодинаміки, а практична – існуванням теплових двигунів. Але в указаних взаємоперетвореннях є принципова асиметрія – на тепло можна перетворити будь-яку частину, чи навіть усю механічну енергію системи[1], але ніяк не можна перетворити на механічну роботу всю теплову (внутрішню) енергію системи. Це є одним із проявів фундаментального закону природи, який називається другим началом термодинаміки.

Нижче будуть розглянуті такі питання:

1. Колові процеси
2. Теплові двигуни
3. Друге начало термодинаміки
4. Теорема Карно

Контрольні запитання

1. Колові процеси

Усі теплові двигуни – пристрої для виконання механічної роботи за рахунок теплової енергії – працюють циклічно. Для отримання механічної енергії в них використовуються так звані колові процеси, що можуть багаторазово повторюватись. При цьому перетворення теплової енергії на механічну здійснюється завдяки зміні об'єму тіл при теплообміні. Тому як робочу речовину в теплових двигунах використовують гази, бо вони найсильніше змінюють свій об'єм при зміні температури. Тож наступні загальні положення термодинаміки будемо розглядати на прикладі ідеального газу.

Колові процеси (цикли). Робота циклу. Коловим процесом або ***циклом*** називається такий процес, у якому система, пройшовши через низку[2] не співпадаючих станів, повертається у вихідний стан. Графік квазістатичного (рівноважного) колового процесу зображується замкненою кривою, як показано на діаграмі *P-V* (рис. 7.1).

Коловий процес характеризується і розширенням, і стисканням газу й має певний напрям, тобто, послідовність змін параметрів стану. Якщо цикл іде "за годинниковою стрілкою", як показано стрілками на рис.7.2а, то цикл називається прямим. При зворотному напрямі зміни параметрів стану (рис.7.2б) цикл називається оберненим.

Квазістатичні цикли є оборотними, тобто, при проведенні такого циклу спочатку в одному, а потім у зворотному напрямі, в зовнішніх тілах не лишається будь-яких змін.

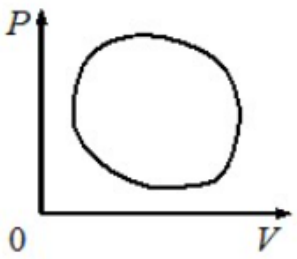
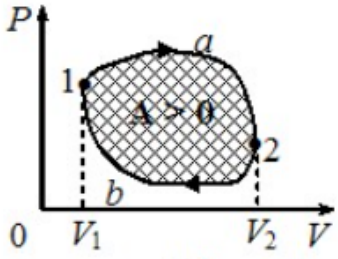
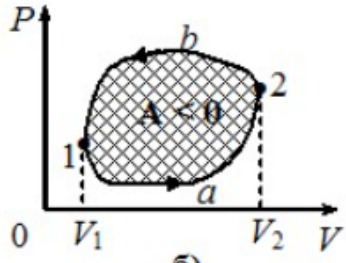


Рис. 7.1



а)



б)

Рис. 7.2

У коловому процесі газ виконує роботу, яка на ділянках розширення $1 \xrightarrow{a} 2$ є додатньою, а на ділянках стискання $2 \xrightarrow{b} 1$ – від'ємною. Тому графічно робота циклу визначається різницею площ під указаними ділянками й чисельно дорівнює взятій з відповідним знаком площі, обмеженій графіком циклу на діаграмі (P, V) (штриховка на рис. 7.2)[3]. У прямому циклі (рис. 7.2а) розширення (ділянка $1 \xrightarrow{a} 2$) відбувається при більшому тиску, ніж стискання (ділянка $2 \xrightarrow{b} 1$), тому повна робота циклу має знак "+". Відповідно, в оберненому циклі знак роботи "–".

Розглянемо детальніше термодинаміку прямого циклу, записавши перше начало термодинаміки (4.13а) для ділянок розширення й стискання газу та для циклу в цілому:

$$(1 \xrightarrow{a} 2) \quad Q_{12} = U_2 - U_1 + A_{12},$$

$$(2 \xrightarrow{b} 1) \quad Q_{21} = U_1 - U_2 + A_{21},$$

$$\text{увесь цикл} \quad Q_{12} + Q_{21} = A_{12} + A_{21}.$$

На ділянці розширення об'єм і температура газу зростають, отже $A_{12} > 0$ і $U_2 > U_1$. Тому $Q_{12} > 0$, тобто, газ дійсно отримує тепло від зовнішніх тіл. Але на ділянці стискання $A_{12} < 0$, $U_1 < U_2$ і $Q_{21} < 0$, отже, газ віддає тепло іншим тілам, а роботу реально виконує не сам газ, а зовнішні тіла. Відповідно до цього позначимо отриману газом кількість тепла як $Q_1 = Q_{12}$, віддану кількість тепла як $Q_2' = -Q_{21}$, і сумарну роботу циклу як $A = A_{12} + A_{21} = A_{12} - A_{21}'$, де A_{21}' - додатня робота зовнішніх тіл по стисканню газу. Тоді робота газу в циклі виразиться, як

$$A = Q_1 - Q_2'. \qquad (7.1)$$

Отже, робота в прямому коловому процесі виконується виключно за рахунок теплової енергії, реально отриманої газом від зовнішніх тіл.

2. Теплові двигуни

Принцип дії та склад теплового двигуна. В прямому циклі після виконання певної механічної роботи $A > 0$ газ повертається у вихідний стан. Тому процес можна повторювати знов і знов. На цьому ґрунтується робота ***теплових двигунів*** - циклічно діючих пристроїв для отримання механічної енергії за рахунок тепла. Через це прямий коловий процес називають іще циклом теплового двигуна.

З наведеного аналізу дії прямого циклу випливає, що, незалежно від конструкції, тепловий двигун повинен мати три обов'язкових структурних елементи: 1) ***робоче тіло*** - речовину (газ), у якій здійснюється робочий цикл двигуна, тобто, прямий коловий процес, в якому відбувається перетворення тепла на механічну енергію; 2) ***нагрівник*** - термодинамічну систему з високою температурою та великим запасом теплової енергії, яка може передаватися робочому тілу і 3) ***холодильник*** - систему з низькою температурою й дуже великою теплоємністю, котра може приймати від робочого тіла не використане тепло практично без зміни свого стану. Отже, в

кожному циклі двигуна робоче тіло: 1) отримує від нагрівника тепло Q_1, 2) виконує за рахунок цього корисну роботу A і 3) віддає холодильнику певну кількість тепла Q_2', повертаючись у вихідний стан. Описана термодинамічна схема теплового двигуна показана на рис 7.3а. Слід зауважити, що вона відображає лише загальний принцип дії теплового двигуна. Що ж до конструктивного втілення, то всі три елементи у вигляді окремих пристроїв існували тільки в класичних парових машинах – перших і довгий час єдиних теплових двигунах, які використовувалися в техніці. В сучасних теплових двигунах, до прикладу, автомобільних, джерелом теплової енергії є паливна суміш у циліндрах, яка після запалення одночасно стає і нагрівником, і робочим тілом – штовхає поршні й виконує корисну роботу. *А* в якості холодильника використовується навколишнє повітря, куди викидається відпрацьоване робоче тіло (вихлопні гази) разом із не використаною тепловою енергією.

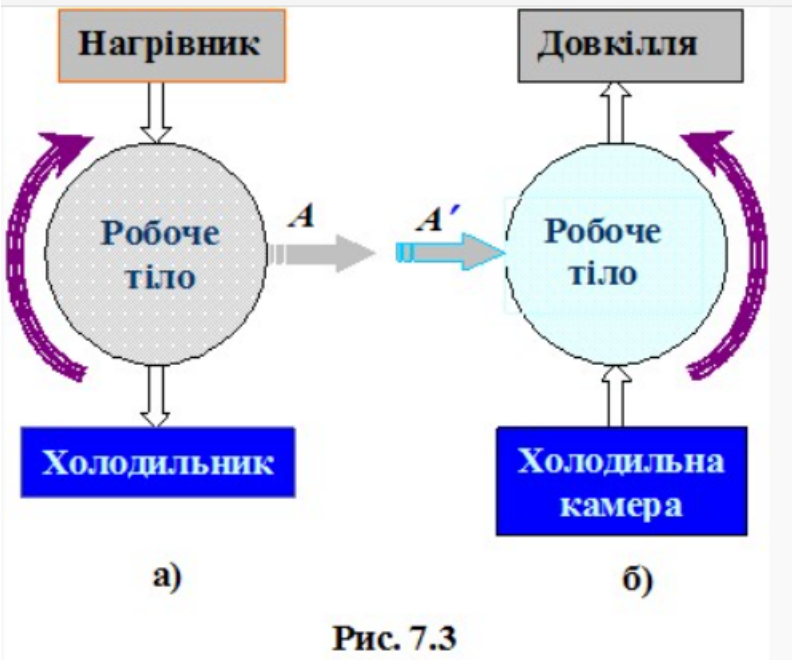


Рис. 7.3

Холодильна машина. В робочому тілі (газі) можна здійснити й обернений цикл, умовно показаний на рис. 7.2б. Це можна, приміром, зробити так. Спочатку стиснений газ привести у добрий тепловий контакт із тілом з меншою температурою (холодним тілом), дати йому можливість розширитися при температурі цього тіла (ділянка $1 \xrightarrow{a} 2$), а потім привести газ у контакт із більш нагрітим (гарячим) тілом і примусовим стисканням повернути його у вихідний стан (ділянка $2 \xrightarrow{b} 1$). У такому разі газ отримує певну кількість тепла Q_1 від холодного тіла й передає певну кількість тепла Q_2' гарячому тілу. Як наслідок, холодне тіло ще більше охолоджується, а гаряче нагрівається. Саме це відбувається в кімнатному холодильнику, в якому тепло від холодильної камери, що має низьку температуру, завдяки роботі компресора передається повітрю в кімнаті, яке має вищу температуру. Тому обернений цикл називають циклом ***холодильної машини.*** Слід одразу наголосити на тому, що, на відміну від теплового двигуна, цикл холодильної машини можна здійснити тільки примусово за рахунок виконання більшої зовнішньої роботи по стисканню газу на ділянці ($2 \xrightarrow{b} 1$), ніж виконує сам газ на ділянці розширення ($1 \xrightarrow{a} 2$). Тому холодильна машина може працювати тільки в комбінації з якимось двигуном (зазвичай – електричним), який дає механічну енергію, необхідну для стискання робочого тіла. Принципова схема холодильної машини показана на рис.7.3б.

3. Друге начало термодинаміки

Ефективність перетворення теплової енергії на механічну в робочому циклі теплового двигуна визначається його коефіцієнтом корисної дії (ККД) η, тобто, відношенням роботи циклу A (формула (7.1)) до кількості тепла Q_1, отриманого робочим тілом від нагрівника[4]:

$$\eta = \frac{A}{Q_1} = \frac{Q_1 - Q_2'}{Q_1}. \qquad (7.2)$$

Отже, ККД теплового двигуна показує, яка частка отриманої теплової енергії перетворюється на механічну, і є тим більшим, чим менша кількість тепла Q_2' передається холодильнику. Тому виникає природне питання, чи не можна обійтися без холодильника й побудувати тепловий двигун, у якому $Q_2' = 0$ і $\eta = 1$, адже це це не суперечить першому началу термодинаміки. Використовуючи в такому двигуні як нагрівник, скажімо, світовий океан із його практично безмежним запасом внутрішньої енергії, можна було би необмежено довго отримувати механічну енергію за рахунок внутрішньої енергії Землі. Оскільки такий гіпотетичний пристрій не суперечить закону збереження енергії, він називається ***вічним двигуном другого роду***. Але, як доводить вся історія розвитку фізики та техніки,

вічний двигун другого роду є неможливим.

У цьому полягає один із фундаментальних законів природи – ***друге начало термодинаміки.*** Більш розлого воно виражається ***формулюванням Кельвіна-Планка***, зміст якого полягає в тому, що

> ***принципово неможливо збудувати циклічно працюючий механізм, єдиним результатом дії якого було би виконання роботи за рахунок внутрішньої енергії нагрівника.***

Отже, друге начало термодинаміки стверджує неможливість повного перетворення отриманого робочим тілом тепла на механічну енергію в тепловому двигуні (або коловому процесі). Але воно не забороняє цього взагалі. До прикладу, при ізотермічному розширенні газу все отримане тепло перетворюється на роботу (див. Розділ IV, (4.15)), але такий процес не є циклічним і не може бути повторений без того, щоб якась кількість теплоти перейшла від газу до інших тіл.

Існують й інші формулювання другого начала, зокрема, ***формулювання Клаузіуса***:

> ***теплота не може спонтанно (саме по собі) переходити від менш нагрітого тіла до більш нагрітого.***

Це твердження теж зовсім не означає, що перехід тепла від холодного тіла до гарячого взагалі неможливий. До прикладу, такий перехід відбувається під час роботи кімнатного холодильника або кондиціонера. Ідеться про те, що ніяким способом неможливо відібрати тепло від тіла з меншою температурою й повністю передати його тілу з більшою температурою так, ***аби в навколишньому середовищі не відбулося жодних змін.*** А якщо перехід тепла від холодного тіла до гарячого все ж таки відбувається, то він обов'язково супроводжується якимось іншими процесами. Приміром, перехід тепла в холодильнику від морозильної камери до повітря в кімнаті супроводжується роботою компресора.

Розглянуті формулювання другого начала термодинаміки Кельвіна-Планка та Клаузіуса словесно є істотно різними, і видається дивним, як вони можуть виражати один і той самий закон. Але можна довести, що обидва формулювання є еквівалентними. Припустимо, що якийсь тепловий двигун отримав від нагрівника кількість тепла Q_1 і віддав холодильнику кількість тепла Q_2', виконавши додатню роботу $A = Q_1 - Q_2'$. А тепер уявімо, що, всупереч формулюванню Клаузіуса, тепло

Q_2' можна повернути від холодильника до нагрівника так, аби в довкіллі не відбулося ніяких змін. Тоді єдиним наслідком такого комбінованого процесу було б виконання роботи A за рахунок теплоти $Q = Q_1 - Q_2'$ і повернення двигуна у вихідний стан. Але це неможливо, згідно з формулюванням Кельвіна-Планка. Отже, із неможливості повного перетворення теплоти в роботу в коловому процесі з необхідністю випливає неможливість спонтанного переходу тепла від холодного тіла (холодильника) до гарячого (нагрівника). Тобто, із формулювання Клаузіуса випливає формулювання Кельвіна-Планка. Подібним чином можна довести й те, що з формулювання Кельвіна-Планка випливає формулювання Клаузіуса.

Друге начало термодинаміки дозволяє краще зрозуміти, чому при встановленні термодинамічної рівноваги в ізольованій системі всі тіла набувають однакової температури: процеси в замкненій системі є спонтанними й тому супроводжуються переходом тепла виключно від більш нагрітих тіл до менш нагрітих, аж доки ступінь нагрітості всіх тіл, тобто їхні температури, не зрівняються. Отже, можна сказати, що друге начало термодинаміки визначає напрям спонтанних термодинамічних процесів у природі.

4. Теорема Карно

Цикл Карно. Позаяк навіть теоретична можливість існування теплового двигуна з ККД $\eta = 1$ забороняється другим началом термодинаміки, постає принципове питання, яким може бути максимальний коефіцієнт корисної дії теплових двигунів. Відповідь отримав французький інженер і фізик Саді Карно, котрий винайшов коловий процес ("цикл Карно"), що має найбільший теоретично можливий ККД.

Цикл Карно є ***оборотним*** (квазістатичним) коловим процесом, котрий складається з двох ізотерм (ділянки $1 \to 2$ і $3 \to 4$ на рис.7.4) та двох адіабат (ділянки $2 \to 3$ і $4 \to 1$) Висока ефективність такого циклу пояснюється тим, що робоче тіло виконує роботу в ізотермічному та адіабатному процесах (ділянки $1 \to 2 \to 3$), тобто без будь-яких непродуктивних втрат теплової енергії.

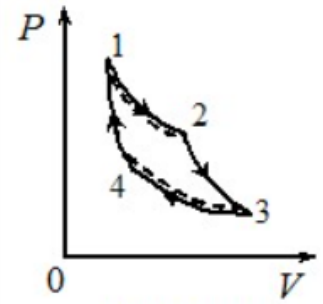


Рис. 7.4

Визначимо ККД циклу Карно, в котрому робочим тілом є ідеальний газ. Згідно з (7.2), можна записати:

$$\eta = \frac{A}{Q_1} = 1 - \frac{Q_2'}{Q_1}, \qquad (7.2a)$$

де Q_1 - кількість тепла, отримана газом від нагрівника, Q_2' - тепло, віддане холодильнику. Оскільки цикл Карно є оборотним, всі процеси в ньому є рівноважними. Тому газ отримує тепло, маючи температуру нагрівника T_1, а віддає не використане тепло, маючи температуру холодильника T_2. В ізотермічному процесі, згідно з (4.15), $Q = A$, тому в (7.2а) $Q_1/Q_2' = A_{12}/A_{34}'$, де A_{12} - робота газу на ділянці

ізотермічного розширення $1 \to 2$, і A'_{34} - робота зовнішніх тіл на ділянці ізотермічного стискання газу $3 \to 4$. Урахувавши це та вирази (4.11) і (4.11а), для ККД отримаємо:

$$\eta_c = 1 - \frac{T_2 \ln(V_3/V_4)}{T_1 \ln(V_2/V_1)}. \qquad (7.3)$$

Далі візьмемо до уваги, що пари точок (2, 3) і (4, 1) належать кожна своїй адіабаті. Тому, згідно з (4.29а), можна записати:

$$\begin{aligned} T_1 V_2^{\gamma-1} &= T_2 V_3^{\gamma-1} \\ T_1 V_1^{\gamma-1} &= T_2 V_2 V_4^{\gamma-1} \end{aligned} \quad \Rightarrow \quad \left(\frac{V_2}{V_1}\right)^{\gamma-1} = \left(\frac{V_3}{V_4}\right)^{\gamma-1} \quad \Rightarrow \quad \ln(V_2/V_1) = \ln(V_3/V_4).$$

Отже, із (7.3) маємо остаточно:

$$\eta_c = 1 - \frac{T_2}{T_1}, \quad \text{або} \quad \eta_c = \frac{T_1 - T_2}{T_1}. \qquad (7.4)$$

Теорема Карно. Вираз (7.4) ми отримали для циклу Карно в ідеальному газі. Але можна довести, що він є універсальним і визначає ККД будь-якого теплового двигуна, що працює на циклі Карно, незалежно ані від речовини робочого тіла, ані від конструкції. Карно також довів, що будь-який інший мислимий цикл має ККД менший, ніж (7.4). Це можна проілюструвати на прикладі циклу, що, як і цикл Карно, складається з двох ізотерм і двох адіабат, але є необоротним (нестатичним). У такому циклі теплообмін відбувається зі скінченною швидкістю, тому температура робочого тіла на ізотермічних ділянках не встигає вирівнюватись і при розширенні є нижчою за температуру нагрівника, а при стисканні – вищою за температуру холодильника. Тому ізотерми проходять, як показано штриховими лініями на рис. 7.4. Через це площа, обмежена графіком необоротного циклу, тож і робота, є меншою, ніж робота (площа) оборотного циклу Карно. Відповідно, меншим є й ККД. Отже, констатуємо, що ККД будь-якого теплового двигуна задовольняє фундаментальній нерівності:

$$\eta \le \frac{T_1 - T_2}{T_1} \quad \text{або} \quad \eta \le 1 - \frac{T_2}{T_1}, \qquad (7.5)$$

де T_1 – максимальна температура, при якій тепло переходить від нагрівника до робочого тіла, а T_2 – мінімальна температура, при якій робоче тіло передає тепло холодильнику.

Підводячи підсумок, можна сказати, що

ніякий тепловий двигун не може мати ККД вищий, ніж двигун, який працює на оборотному циклі Карно.

Наведені твердження й підсумковий вираз (7.5) складають зміст ***теореми Карно***, яка посідає провідне місце в термодинаміці.

З практичної точки зору вираз (7.5) показує, що для збільшення ККД теплових двигунів треба підвищувати температуру нагрівника й знижувати температуру холодильника. Але реальні можливості підвищення температури нагрівника обмежені відносно невисокою жаростійкістю матеріалів і тепловим розширенням деталей, що призводить до виходу двигуна з ладу при перегріві. Що ж до температури холодильника, то її мінімальне обґрунтоване значення обмежене температурою довкілля. Причина полягає в тому, що для охолодження холодильника до більш низької, ніж навколишня, температури треба виконувати роботу, що, як випливає з

другого начала термодинаміки, є більшою, ніж приріст корисної роботи двигуна за рахунок підвищення ККД. З указаних причин ККД навіть найкращих теплових двигунів є істотно меншим за одиницю.

Теорема Карно має не лише практичне, а й велике теоретичне значення. З попереднього зрозуміло, що теплота, передана холодильнику, вже не може бути ефективно перетворена на механічну енергію. Відбувається своєрідне знецінення теплової енергії, коли все більша її частина стає непридатною для отримання механічної енергії. Причому, це стосується не лише теплових двигунів, а взагалі будь-яких процесів перетворення теплової енергії на механічну.

Контрольні запитання

1. Що таке “коловий процес” (цикл)? Які бувають цикли? Який цикл називається оборотним, а який – оберненим?
2. Чому дорівнює робота системи в прямому циклі? Як вона відображається на графіку циклу?
3. Чи можливі процеси, в яких усе отримане системою тепло йде на виконання нею роботи?
4. Чи можливо перетворити на роботу все отримане системою тепло в коловому процесі?
5. Що таке “тепловий двигун”? Наведіть термодинамічну схему теплового двигуна. Чи можна віднести електромотор до теплових двигунів?
6. Що таке “вічний двигун другого роду”? У чому полягає друге начало термодинаміки?
7. Дайте формулювання другого начала термодинаміки Кельвіна-Планка та Клаузіуса і доведіть їхню еквівалентність.
8. Який коловий процес називається циклом Карно? Чому дорівнює ККД циклу Карно?
9. Про що трактує теорема Карно?

---

[1] Більш за те, це відбувається навіть усупереч нашому бажанню, так що в замкненій системі з часом уся механічна енергія може перетворитися на теплову.

[2] Нагадаємо, що квазістатичним називається достатньо повільний процес, коли можна вважати, що система проходить через низку рівноважних станів.

[3] Див. Ррзділ IV, п. 2.

[4] Це так званий термодинамічний ККД. Реальний ККД, тобто, відношення механічної енергії на валу двигуна до теплової енергії, що виділяється при згорянні палива, ясна річ, менший через недосконалість конструкції та втрати на тертя у вузлах двигуна.

8. Ентропия

Теорема Карно не тільки визначає обмеження, накладені природою на можливість перетворення теплової енергії на механічну. З неї випливає існування фізичної величини, що дозволє аналітично виразити друге начало термодинаміки й висвітлює його фундаментальний зміст. Ця величина, котра називається ентропією, розглядається в наступних питаннях:

1. Ентропія

2. Обчислення ентропії. Теорема Нернста

3. Нерівність Клаузіуса. Закон зростання ентропії

4. Ентропія й імовірність. Формула Больцмана

Контрольні запитання

1. Ентропія

Рівність Клаузіуса. Розглянемо цикл Карно не прагматично, як найкращий робочий цикл теплового двигуна, а як конкретний оборотний коловий процес у довільній термодинамічній системі. В такому разі доцільно не розрізняти отриману та віддану системою теплоту, а скрізь розглядати її як алгебраїчну величину. Відтак, зробивши у виразі (7.2 а) заміну $Q_2' = -Q_2$ і прирівнявши його до (7.4), отримаємо:

$$1 + \frac{Q_2}{Q_1} = 1 - \frac{T_2}{T_1} \Rightarrow \frac{Q_2}{Q_1} = -\frac{T_2}{T_1} \Rightarrow \frac{Q_1}{T_1} = \frac{Q_2}{T_2},$$

або

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0. \qquad (8.1)$$

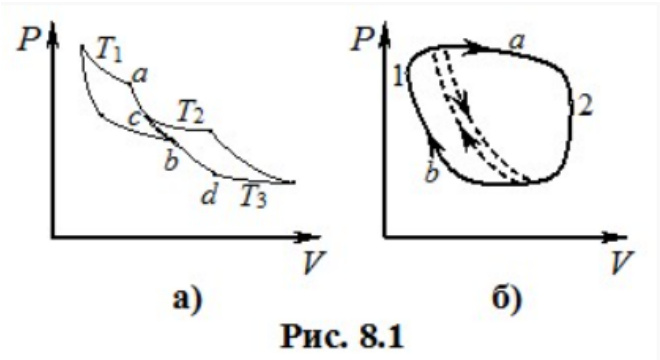


Рис. 8.1

Клаузіус назвав відношення кількості теплоти, отриманої системою, до її температури ***зведеною кількістю теплоти***. Тож сума зведених кількостей теплоти, отриманих системою в циклі Карно, дорівнює нулю. Відтак розглянемо оборотний коловий процес, який складається з чотирьох ізотерм і чотирьох адіабат (рис. 8.1*а*), ділянки *a*-*b* і *c*-*d* яких лежать на одній адіабаті (точки *b* і *c* не з'єднані). Якщо подумки продовжити адіабату *ac* від точки *c* до точки *b*, а адіабату *db* – від точки *b* до точки *c,* то вказаний комбінований процес розпадається на два окремі цикли Карно, котрі здійснюються між температурами T_1 і T_4 та T_2 і T_3. При цьому, позаяк вставлені віртуальні процеси між точками *b* і *c* є адіабатичними ($Q = 0$), то вони ніяк не впливають на зведені теплоти. Відтак для всього "ступінчастого" циклу маємо:

$$\frac{Q_1}{T_1}+\frac{Q_2}{T_2}+\frac{Q_3}{T_3}+\frac{Q_4}{T_4}=0 \quad \Rightarrow \quad \odot\sum_i \frac{Q_i}{T_i}=0, \tag{8.2}$$

де кружечок біля знаку суми означає, що додавання виконується по всьому коловому процесу (по замкненому "шляху"). Довільний коловий процес (рис. 8.1б) також можна розчленувати віртуальними адіабатами на безліч елементарних циклів Карно (на рис.8.1б штрихами показано один із таких елементарних циклів). При цьому на кожній елементарній ділянці циклу система отримує нескінченно малу зведену кількість теплоти $\delta Q/T$, тому дискретна сума (8.2) вироджується в інтеграл по всьому коловому процесу:

$$\oint \frac{\delta Q}{T}=0. \tag{8.3}$$

Вирази (8.1), (8.2) і (8.3) називаються ***рівностями Клаузіуса***. Вони означають, що

> ***в будь-якому оборотному коловому процесі сумарна зведена кількість теплоти дорівнює нулю***.

Зауважимо, що найбільш загальним є вираз (8.3), а (8.1) і (8.2) є його окремими випадками.

Ентропія. Отриманий результат можна подати інакше. Поділимо весь цикл (рис. 8.1 б) на дві ділянки: $1 \xrightarrow{a} 2$ та $2 \xrightarrow{b} 1$ і, відповідно, інтеграл (8.3), на дві частини. Тоді

$$\oint \frac{\delta Q}{T}=\int_{1a2}\frac{\delta Q}{T}+\int_{2b1}\frac{\delta Q}{T}=0 \quad \Rightarrow \quad \int_{1a2}\frac{\delta Q}{T}=-\int_{2b1}\frac{\delta Q}{T}. \tag{8.3а}$$

Оскільки цикл оборотний, то при проведенні процесу $2 \xrightarrow{b} 1$ у зворотному напрямку по шляху $1 \xrightarrow{b} 2$ в кожній точці зміниться тільки напрям теплообміну й знак величини $\delta Q/T$, тому

$$-\int_{2b1}\frac{\delta Q}{T}=\int_{1b2}\frac{\delta Q}{T}$$

Отже, (8.3*а*) можна переписати, як

$$\int_{1a2}\frac{\delta Q}{T}=\int_{1b2}\frac{\delta Q}{T}. \tag{8.4}$$

Це дуже важливий результат, який означає, що при оборотному переході системи із якогось стану 1 у якийсь стан 2 сумарна зведена кількість теплоти, отримана системою, не залежить від способу переходу (тобто, виду процесу) й визначається тільки параметрами початкового та кінцевого стану системи. Схожа ситуація спостерігається в механіці[1] стосовно роботи консервативних сил, причому незалежність вказаної роботи від траєкторії руху частинки дозволяє ввести поняття потенціальної енергії. Подібним чином у термодинаміці вводиться поняття ***ентропії S***.

> ***Ентропією системи називається величина, зміна котрої при переході системи з одного стану в інший будь-яким квазістатичним (оборотним) способом дорівнює сумарній***

***зведеній кількості теплоти, що її отримує система при цьому переході*:**

$$\Delta S = S_2 - S_1 = \left(\int_{1\to 2} \frac{\delta Q}{T} \right)_{\text{об}}. \tag{8.5}$$

Відповідно, при елементарній зміні стану приріст ентропії

$$\mathrm{d}S = \frac{\delta Q}{T}. \tag{8.5а}$$

Одразу звернімо увагу на такі загальні властивості ентропії. По-перше, хоча кількість теплоти є функцією процесу, величина $\mathrm{d}S = \delta Q/T$ і, відповідно, ΔS не залежить від способу зміни стану, отже,

***ентропія є функцією стану системи*.**

По-друге, кількість теплоти, отримана системою в будь-якому квазістатичному процесі, є прямо пропорційною масі (кількості речовини) системи. Тому

***ентропія є адитивною величиною*,**

тобто, ентропія системи дорівнює сумі ентропій її окремих частин. Це дозволяє поширити поняття ентропії на нерівноважні стани системи, якщо її можна поділити на окремі частини (підсистеми), що перебувають у відмінних, але квазірівноважних станах. Нарешті, з (8.3) і (8.5) випливає що в оборотному коловому процесі

$$\Delta S = 0, \tag{8.6}$$

тобто,

***в оборотних колових процесах ентропія не змінюється*.**

2. Обчислення ентропії. Теорема Нернста

Те, що ентропія є функцією стану, дозволяє при визначенні зміни ентропії системи обирати в (8.5) будь-який зручний для обчислень шлях інтегрування. Розглянемо деякі приклади.

***Ентропія ідеального газу*.** Нехай ідеальний газ у кількості ν молів оборотно переходить із стану 1 (P_1, V_1, T_1) у стан 2 (P_2, V_2, T_2). Визначимо зміну ентропії газу. Згідно з першим началом термодинаміки (4.14), при елементарній зміні стану системи $\delta Q = \mathrm{d}U + P\mathrm{d}V$. Урахувавши (5.7) і (2.6), будемо мати: $\delta Q = \nu C_V \mathrm{d}T + \nu RT\mathrm{d}V/V$. Звідси для елементарної зведеної кількості теплоти дістанемо:

$$\frac{\delta Q}{T} = \nu C_V \cdot \frac{\mathrm{d}T}{T} + \nu R \cdot \frac{\mathrm{d}V}{V} \tag{8.7}$$

Інтегруючи цей вираз відповідно до (8.5), отримаємо

$$\Delta S = \nu C_V \int_{T_1}^{T_2} \frac{\mathrm{d}T}{T} + \nu R \int_{V_1}^{V_2} \frac{\mathrm{d}V}{V} \quad \Rightarrow$$

$$\Rightarrow \quad \Delta S = S_2 - S_1 = \nu \left(C_V \ln \frac{T_2}{T_1} + R \ln \frac{V_2}{V_1} \right). \tag{8.8}$$

Величину ΔS можна подати й через інші параметри стану, зробивши у (8.8) відповідні заміни на основі рівняння Клапейрона (2.6). Зокрема, можна отримати такий симетричний вираз:

$$\Delta S = \nu \left(C_V \ln \frac{P_2}{P_1} + C_P \ln \frac{V_2}{V_1} \right). \qquad (8.8a)$$

З (8.8) і (8.8*a*) випливають прості формули для обчислення зміни ентропії ідеального газу в ізопроцесах. До прикладу, в ізотермічному процесі $\Delta S = \nu R \ln(V_2/V_1)$, або $\Delta S = \nu R \ln(P_1/P_2)$, в ізобарному процесі $\Delta S = \nu C_P \ln(V_2/V_1)$, або $\Delta S = \nu C_P \ln(T_1/T_1)$ в ізохорному процесі $\Delta S = \nu C_V \ln(P_2/P_1)$, або $\Delta S = \nu C_V \ln(T_2/T_1)$. Принагідно відмітимо, що в адіабатному процесі $\delta Q = 0$, тому й $\Delta S = 0$. Отже, адіабатний процес – то є ***ізоентропний (ізоентропійний)*** процес. Відповідно, на діаграмі (T, S) графік циклу Карно зображується прямокутником (рис. 8.2).

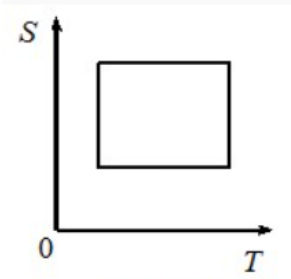


Рис. 8.2

Це, зокрема, дозволяє дуже легко отримати вираз (7.4) для ККД циклу Карно. Справді, оскільки в цьому циклі теплообмін відбувається тільки на ізотермах при сталих температурах T_1 і T_2, то, відповідно до (8.5), $Q_1 = T_1(S_2 - S_1)$ і $Q_2 = T_2(S_2 - S_1)$. Підставивши ці вирази в (7.2) з урахуванням знаків, одразу отримуємо формулу (7.4).

Зміна ентропії при необоротній зміні стану газу. Нехай в одній із двох теплоізольованих посудин, які з'єднані тонкою трубкою з перекритим краном, знаходиться ν молів ідеального газу, а в другій – вакуум. Об'єми посудин V_1 і V_2. Кран відкривають і газ, розширюючись в пустоту, заповнює обидві посудини й переходить у новий рівноважний стан. Такий процес є необоротним: для повернення газу у вихідний стан його треба стиснути, а для цього треба виконати роботу та забезпечити теплообмін, що суттєво змінить стан зовнішніх тіл. Тому годі навіть і думати про обчислення зміни ентропії за допомогою виразу (8.5) безпосередньо для цього процесу, адже для необоротного процесу саме поняття зведеної кількості теплоти втрачає зміст через невизначеність температури системи. Але розширення ідеального газу в пустоту відбувається без зміни температури[2], тому замість реального процесу можна розглянути квазістатичне ізотермічне розширення газу від об'єму V_1 до об'єму $V_1 + V_2$. Тоді, з урахуванням (4.15) і (2.8), із (8.5) отримаємо:

$$\Delta S = \int_{V_1}^{V_1+V_2} \frac{P\mathrm{d}V}{T} = \nu R \int_{V_1}^{V_1+V_2} \frac{\mathrm{d}V}{V} \quad \Rightarrow \quad \Delta S = \nu R \ln \frac{V_1 + V_2}{V_1} \qquad (8.9)$$

Отже, в розглянутому процесі ентропія газу зростає.

Теорема Нернста. Означення (8.5) визначає не саму ентропію, а тільки її зміну при заданій зміні стану системи. Тому для визначення самої ентропії необхідно в будь-якому обраному стані покласти ентропію рівною нулю, подібно до того, як це робиться в механіці стосовно потенціальної енергії[2]. Але у випадку ентропії вибір нульового рівня, по суті, є однозначним, завдяки так званій *тепловій теоремі Нернста*. Нернст на основі емпіричних даних установив, що при наближенні температури до 0 К ентропія будь-якої системи прямує до певної границі, що не залежить від жодних параметрів стану системи. Тому цю границю покладають рівною нулю, і приймають таке формулювання:

> ***при наближенні температури до абсолютного нуля ентропія будь-якої системи прямує до нуля:***

$$S \to 0, \text{ якщо } T \to 0. \tag{8.10}$$

Слід зауважити, що теорема Нернста логічно не випливає з інших законів термодинаміки. Тому її ще називають ***третім началом термодинаміки***.

Теорема Нернста дозволяє дати означення не лише для зміни, а й для самої ентропії системи:

$$S = \int_0^T \frac{\delta Q}{T}. \tag{8.11}$$

---

3. Нерівність Клаузіуса. Закон зростання ентропії

Усі реальні термодинамічні процеси є необоротними, тому ККД довільного циклу визначається не формулою (7.4), а нерівністю (7.5). Через це в загальному випадку замість співвідношень (8.1), (8.2) і (8.3) виконуються ***нерівності Клаузіуса***:

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} \leq 0, \tag{8.12}$$

$$\odot \sum_i \frac{Q_i}{T_i} \leq 0, \tag{8.13}$$

$$\oint \frac{\delta Q}{T} \leq 0, \tag{8.14}$$

в яких знак “<” відноситься до необоротних, а знак “=”– до оборотних колових процесів.

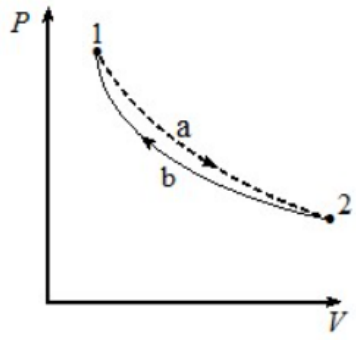


Рис. 8.3

Нерівність (8.14) дозволяє з'ясувати загальну поведінку ентропії в необоротних процесах. Нехай система перейшла зі стану 1 у стан 2 в якомусь необоротному процесі, показаному умовно[3] штриховою лінією $1 \xrightarrow{a} 2$ на рис. 8.3. Повернемо подумки систему у вихідний стан, здійснивши певний оборотний процес, який показано лінією $2 \xrightarrow{b} 1$. Відтак отримуємо необоротний цикл, для якого, згідно з (8.14) можна записати:

$$\int_{1a2} \frac{\delta Q}{T} + \int_{2b1} \frac{\delta Q}{T} < 0.$$

У першому інтегралі під T слід розуміти температуру навколишнього середовища, при якій відбувається теплообмін, у другому інтегралі ця температура є й температурою системи. Процес $2 \xrightarrow{b} 1$ є оборотним, тому, згідно з (8.5), другий інтеграл дорівнює зміні ентропії $S_1 - S_2$ при переході системи із стану 2 у стан 1. Отже,

$$\int_{1a2} \frac{\delta Q}{T} + (S_1 - S_2) < 0 \quad \rightarrow \quad S_2 - S_1 > \int_{1a2} \frac{\delta Q}{T}. \qquad (8.15)$$

Зведена кількість теплоти (інтеграл у правій частині кінцевої нерівності), яку отримує система, може мати будь-який знак, залежно від напрямку та інтенсивності теплообміну на окремих стадіях процесу. Тому в довільному необоротному процесі ентропія системи може як зростати, так і зменшуватись. Але, якщо система є теплоізольованою, тобто, не обмінюється теплом із навколишніми тілами, то інтеграл у (8.15) дорівнює нулю, відтак $\Delta S = S_2 - S_1 > 0$. Отже, при необоротних процесах у теплоізольваній системі ентропія обов'язково зростає. Об'єднуючи цей результат із (8.6), отримуємо універсальне співвідношення

$$\Delta S \geq 0, \qquad (8.16)$$

яке означає, що

ентропія теплоізольованої системи не може зменшуватися – вона або зростає, або лишається незмінною.

Цей результат називають ***законом зростання ентропії.*** Він безпосередньо випливає з теореми Карно, котра, у свою чергу, є прямим наслідком другого начала термодинаміки і, по суті, виражає його. Тому закон зростання ентропії (8.16) розглядають як ще одне формулювання другого начала термодинаміки. З нього, зокрема, випливає висновок про те, що в замкнених системах можливі лише такі процеси, при яких ентропія не зменшується. А практично,позаяк строго оборотних

процесів у природі не існує, в усіх процесах у замкнених системах ентропія зростає[4]. Тим самим, друге начало термодинаміки визначає напрям процесів, які можуть відбуватися в природі.

4.Ентропія й імовірність. Формула Больцмана

Раніше вже говорилося про те, що будь-яка ізольована нерівноважна система з часом обов'язково переходить у стан термодинамічної рівноваги. З іншого боку, друге начало термодинаміки свідчить, що в такій системі можливі лише процеси, в яких ентропія зростає. Звідси випливає логічний висновок, що в стані термодинамічної рівноваги ентропія системи має максимальне можливе значення, тобто,

умовою термодинамічної рівноваги системи є максимум її ентропії.

В цьому сенсі можна говорити про те, що ентропія показує ступінь наближеності системи до стану рівноваги: чим ближча ентропія до свого максимуму, тим ближчим є стан системи до рівноважного. Така думка підтримується й тим, що перехід замкненої системи до рівноваги є необоротним, а саме необоротність термодинамічних процесів є формальною причиною зростання ентропії у замкнених системах. Тому постає питання про причину цієї необоротності, адже термодинамічні процеси визначаються рухом і взаємодією між частинками макроскопічної системи, а в законах руху (законах механіки) ніякої необоротності немає. Наприклад, якби Земля рухалася навколо Сонця в протилежному до існуючого напрямку, ні на Землі, ні у Всесвіті взагалі, нічого б не змінилось.

Виявляється, що причина необоротності теплових процесів полягає в макроскопічності термодинамічних систем, тобто, в тому, що вони містять гігантську кількість частинок. У цьому можна переконатися, позірно[5] розглянцвши явно необоротний процес розширення газу в пустоту на основі статистичного (молекулярно-кінетичного) методу. Отже, нехай маємо посудину, розділену перегородкою на дві однакові частини 1 і 2, причому частина 1 містить N молекул газу, а 2 – порожня. Якщо перегородку прибрати, виникне потік молекул із частини 1 у частину 2. При цьому, внаслідок хаотичності теплового руху, одразу виникне й зустрічний потік молекул з частини 2 в 1. Очевидно, що в міру зменшення кількості молекул в частині 1 і збільшення їхньої кількості в частині 2, перший потік буде з часом послаблюватись, а другий – підсилюватись. Отож через певний час, коли кількості молекул в обох посудинах стануть однаковими, ці потоки зрівняються, перетікання газу з частини 1 у частину 2 припиниться, й у системі встановиться рівноважний стан. Проте ця рівновага є динамічною – перехід молекул з однієї половини посудини в іншу не припиняється, просто кількості молекул, які перетинають у зустрічних напрямках умовну поверхню, що відділяє одну половину посудини від іншої, стають *у середньому* однаковими. Але в окремі моменти, знову таки через хаотичність теплового руху, вказана рівність порушується й виникає короткочасне перетікання газу з однієї частини посудини в іншу, причому байдуже, в якому напрямку. Отже, спонтанний перехід газу з *другої частини в першу*, тобто,

обернений до розширення процес, є принципово можливим. Але йдеться тільки про дуже незначні частки молекул і про гранично малі проміжки часу. Ніхто й ніколи не спостерігав переходу з другої частини посудини в першу значної частки молекул, тим паче – всього газу. Проаналізуємо, чому? Нехай спочатку $N = 1$, тобто в частині 1 у початковий момент є лишень одна молекула. Після прибирання перегородки молекула буде рухатися по всій посудині, з однаковою ймовірністю перебуваючи або в другій, або в першій частині посудини. Але щоразу, коли молекула опиняється в першій половині посудини, наш "газ" спонтанно повертається в вихідний стан, причому ймовірність такого переходу дорівнює $P = 1/2$, тобто, є дуже великою. Нехай тепер кількість молекул $N = 2$. Якщо ми будемо спостерігати за цими молекулами певний час, за який зробимо n спостережень, то виявимо, що перша молекула знаходиться в першій частині посудини в $n/2$ випадків. У свою чергу друга молекула в половині з цих $n/2$ випадків буде теж знаходиться в першій частині посудини. Тому обидві молекули будуть разом знаходитись у першій половині посудини в $(1/2)(n/2) = n/4$ випадків. Отже, ймовірність спонтанного повернення у вихідний стан для "газу" з двох молекул уже менша й становить $P = (1/2)\cdot(1/2) = 1/2^2$. Міркуючи так само, для $N = 3$ отримаємо $P = (1/2)\cdot(1/2)\cdot(1/2) = 1/2^3$, і в загальному випадку $P = 1/2^N$, де N - кількість молекул у посудині. Як бачимо, спонтанний перехід газу у вихідний стан після встановлення рівноваги є принципово можливим, але ймовірність такого процесу стрімко зменшується при збільшенні кількості молекул у посудині. Так, при $N = 10$ імовірність $P = 1/2^{10} = 1/1024 \approx 0{,}001$. Це мала, проте помітна величина: якщо "фотографувати" розподіл молекул у посудині протягом великого часу, то в середньому на кожні 1000 кадрів буде траплятись один, на якому всі молекули знаходяться в першій частині посудини. Це повністю підтверджується комп'ютерним моделюванням. Але вже при $N = 100$ ймовірність спонтанного повернення у вихідний стан стає $P \approx 10^{-30}$, тобто, виявити таку подію немає жодного шансу. Реальна кількість молекул в 1 $см^3$ газу складає ~10^{20} і $P \approx 10^{-310^{19}}$. Наскільки мале це число, неможливо уявити. Тому спонтанний перехід усіх молекул газу знов у першу частину посудини абстракно є можливий, але в дісності ніколи не відбувається, і розширення газу в пустоту є необоротним. Таким чином, молекулярно-кінетична теорія не тільки пояснює причину необоротності термодинамічних процесів, а й уточнює саме поняття необоротності. А саме, необоротні процеси – це не такі процеси, що принципово не можуть самі по собі відбуватись у зворотньому напрямі, а такі, для яких сронтанний зворотній процес є гранично мало ймовірними. Відповідно, й кінцевий стан системи не є математично однозначно заданим, і

рівноважний стан є найбільш імовірним макроскопічним станом системи.

У світлі сказаного важливо з'ясувати, чим визначається ймовірність того чи іншого макроскопічного стану системи. Відповідь не важко знайти, якщо пригадати, що макроскопічний стан системи реалізується через її мікроскопічні стани, тобто, через сукупність станів окремих частинок системи (див. Вступ). Але зв'язок між мікро- та макростанами системи не є взаємно однозначним. Тоді як кожному мікроскопічному станові відповідає визначений макроскопічний стан системи, заданому макроскопічному станові відповідає безліч різних мікроскопічних станів. Наприклад,

в розглянутому позірному процесі розширення газу всіх випадках (мікростанах), коли в даній частині посудини знаходиться задана кількість молекул, термодинамічні параметри (концентрація частинок, тиск, внутрішня енергія, тощо) є однаковими, тобто газ перебуває в одному й тому самому макроскопічному стані. Зважаючи на хаотичність руху молекул і величезну частоту зіткнень між ними, в МКТ прийнятий постулат про рівноймовірність усіх мікростанів, які відповідають заданому макроскопічному станові системи, тобто, – що всі мікростани трапляються однаково часто. Відтак стає зрозумілим, що заданий макроскопічний стан буде спостерігатися тим частіше, отже, тим імовірніше, чим більша кількість мікроскопічних станів йому відповідає. Тому в МКТ за міру ймовірності макроскопічного стану системи приймають його ***термодинамічну ймовірність*** або ***статистичну вагу*** **Ω** – кількість способів (мікростанів), якою реалізується даний макроскопічний стан системи. Наведена нижче таблиця ілюструє сказане для "газу" з чотирьох молекул 1, 2, 3, 4. У ній показані всі можливі розподіли молекул по частинах посудини (мікростани), відповідні макростани та їхні термодинамічна Ω й математична *P* ймовірності. Як видно, рівноважний стан, у якому молекули розподілені по посудині рівномірно, є найбільш імовірним (Ω = 6), а стани, в яких усі молекули самодовільно скупчуються тільки в частині 1, або тільки в частині 2, є найменш імовірними (Ω = 1).

Способи реалізації макростанів		Ω	*P*
A	B		
0	1, 2, 3, 4	1	1/16
1	2, 3, 4	4	4/16
2	1, 3, 4		
3	1, 2, 4		
4	1, 2, 3		
1, 2	3, 4	6	6/16
1, 3	2, 4		
1, 4	2, 3		
2, 3	1, 4		
2, 4	1, 3		
3, 4	1, 2		
1, 2, 3	4	4	4/16
1, 2, 4	3		
1, 3, 4	2		
2, 3, 4	1		
1, 2, 3, 4	0	1	1/16
Усього мікростанів $2^4 = 16$			

Поняття термодинамічної ймовірності дозволяє кількісно виразити зв'язок між ентропією та ступенем наближеності системи до термодинамічної рівноваги, про який говорилося вище. Це зробив Л. Больцман, який довів, що ентропія S системи в стані із термодинамічною ймовірністю Ω визначається, як

$$S = k \ln \Omega. \qquad (8.17)$$

де $k = 1{,}38 \cdot 10^{-23}$ Дж/К – вже відома нам з МКТ "стала Больцмана".

Вираз (8.17), який називається ***формулою Больцмана***, є статистичним означенням поняття ентропії. Формула Больцмана відіграє велику роль у статистичній фізиці та термодинаміці. Зокрема, вона розкриває глибинний зміст закону зростання ентропії. А саме, цей закон указує на те, що всі процеси в природі йдуть у напрямку зростання молекулярного безпорядку в системі. Це в свій час спонукало дискусію про так звану "теплову смерть Всесвіту", тобто, про неминучий перехід нашого Всесвіту до стану термодинамічної рівноваги й припинення в ньому будь-яких процесів. Це, в принципі, не виключено, але сьогодні для однозначної відповіді на це питання наявні знання про походження та еволюцію Всесвіту є недостатніми.

Контрольні запитання

1. Що таке "зведена кількість теплоти"? Чому вона дорівнює в довільному коловому процесі?
2. Яка фізична величина називається ентропією? В яких одиницях вона вимірюється? Які має основні властивості?
3. Чому не можна безпосередньо визначити зміну ентропії в заданому необоротному процесі?
4. Чи можна взагалі визначити зміну ентропії в необоротному процесі? Чому? Як?
5. У чому полягає закон зростання ентропії? Чи може ентропія термодинамічної системи зменшуватись? За яких умов?
6. Дайте означення рівноважного стану системи на основі поняття ентропії.
7. Який статистичний зміст має ентропія?
8. Що таке "термодинамічна імовірність" макроскопічного стану системи. Який зв'язок існує між нею та ентропією?

[1] Див. «Фізика для бакалаврів. Механіка», розділ IV, п.2.2.

[2] Див. розділ V, п. 3.

[3] Нагадаємо, що в необоротному процесі параметри стану не мають визначених значень для всієї системи, тому процес неможливо зобразити якоюсь лінією на графіку.

[4] Конкретним прикладом цього є результат наведеного в п. 2 розрахунку зміни ентропії при розширенні газу в пустоту.

[5] Це означає "удавано" і є одним із способів доведення. До прикладу, у фізиці нерідко розглядають "позірні експерименти" – удавані досліди, які не суперечить фізичним законам, але технічно є неможливими.

ЕЛЕКТРИКА І МАГНЕТИЗМ

Зміст

ЕЛЕКТРИКА І МАГНЕТИЗМ

ПЕРЕДМОВА

ВСТУП

I. ЕЛЕКТРИЧНЕ ПОЛЕ У ВАКУУМІ
1. Напруженість електричного поля
2. Потенціал
3. Зв’язок між напруженістю та потенціалом
4. Електричний диполь
5. Циркуляція і потік електричного поля

II. ЕЛЕКТРИЧНЕ ПОЛЕ В РЕЧОВИНІ
1. Електричне поле в діелектриках
2. Провідники в електричному полі

III. ЕЛЕКТРИЧНИЙ СТРУМ
1. Загальні закони електричного струму
2. Електричне коло
3. Струм у колі з конденсатором
4. Робота і потужність струму

IV. МАГНІТНЕ ПОЛЕ
1. Магнітна взаємодія
2. Потік і циркуляція магнітного поля
3. Магнітне поле в речовині

V. ЕЛЕКТРОМАГНЕТИЗМ
1. Явище електромагнітної індукції
2. Рівняння Максвелла

ПЕРЕДМОВА

У третій частині короткого загального курсу фізики для студентів технічних ЗВО, які навчаються за програмою підготовки бакалаврів, висвітлюються основні положення і закони електромагнітної теорії. При цьому технічним застосуванням приділяється лише мінімальна увага, позаяк це є предметом профільних курсів фізичного спрямування, котрі передбачаються навчальними планами спеціальностей.

Вивчення загальної фізики у вишах звичайно припадає на самий початок навчання і проходить за хронічного дефіциту адекватного математичного забезпечення. З цієї причини значну увагу в посібнику приділено роз'ясненню фізичного змісту математичних понять, шо використовуються. Також, аби запобігти перевантаження математикою, в посібнику розглядаються тільки інтегральні рівняння електромагнітного поля, а головні поняття і співвідношення диференціального апарату теорії векторних полів у застосуванні до рівнянь електромагнітного поля винесені в Додаток як матеріал для не обов'язкового читання.

Слід також зробити наступне зауваження стосовно прийнятих у посібнику позначень: скрізь, де це не тягне неоднозначності, модулі векторів позначаються символом величини без використання модульних дужок і додаткових індексів, а в позначеннях проекцій векторів опускаються індекси осей. Така спрощена система позначень є поширеною у фізиці та технічних науках і робить формули більш прозорими та зручними для викладок. Ще одне зауваження стосується запису числових формул і рівнянь. У числовому рівнянні мають бути вказані одиниці, в яких виражені всі величини, але, коли про це не сказано прямим текстом, то всі числові величини подаються в основних одиницях СІ.

ВСТУП

Електричний заряд і електромагнітне поле

Фундаментальні взаємодії. Вивчаючи механіку, можна переконатися, що зміни стану системи зумовлюються силами — величинами, які визначають взаємодію частинок між собою та із зовнішніми тілами. При цьому спостерігається багато різних сил: сили тяжіння, пружності, тертя та опору середовища, сили натягу нитки або троса, магнітні та електричні сили, тощо. Але всі ці сили є проявом всього чотирьох фундаментальних[1] взаємодій — гравітаційної, електромагнітної, слабкої та сильної[2].

Властивості фундаментальних взаємодій суттєво відрізняються. Зокрема, гравітаційні та електромагнітні сили теоретично існують при будь-якій відстані між взаємодіючими частинками і є *далекодіючими*. Натомість сильна та слабка взаємодії є *короткодіючими* вони існують тільки на відстанях порядку розмірів атомного ядра і менше. Тому всі процеси в навколишньому світі, крім процесів в атомних ядрах та надрах зірок, зумовлені тільки гравітаційною та електромагнітною взаємодіями. При цьому гравітаційна взаємодія є гранично слабкою,[3] вона помітна лише при дуже великій (астрономічній) масі хоча б одного із взаємодіючих тіл. Тому, за винятком тяжіння всі процеси, які ми спостерігаємо навколо себе, мають електромагнітну природу.

Електричний заряд. Гравітаційна взаємодія є властивою для всіх без винятку частинок матерії. Але до електромагнітної взаємодії є здатними лише деякі з них. Так, протони й електрони беруть участь в електромагнітній взаємодії, а нейтрони і фотони — ні. До електромагнітної взаємодії здатними є тільки частинки з певною внутрішньою властивістю, що називається електричним *зарядом*. При цьому, будучи внутрішньою властивістю частинки, заряд є невіддільним від неї і не може бути повністю чи частково переданий іншій частинці, як, скажімо, імпульс або енергія. Тому словом “заряд” задля зручності позначають і саму таку частинку. До прикладу, говорять не “рух частинок, які мають заряд”, а “рух зарядів”.

Природа електричного заряду ще остаточно не встановлена, але його властивості вивчені достатньо повно. Нагадаємо їх.

1. Заряд усіх заряджених елементарних частинок, які можуть існувати у вільному стані, має однакову величину $e = 1,6 \cdot 10^{-19}$ Кл, яка називається елементарним зарядом. Отже, в природі існує найменша порція або квант електричного заряду, тобто він є дискретною величиною на відміну, скажімо, від маси, яка може бути якою завгодно.

2. У природі існує два типи електричного заряду, названі позитивним “+” та негативним “–”. Із добре знайомих частинок носієм негативного заряду $-e$ є електрон, а позитивного $+e$ — протон. Наявність двох типів заряду випливає з того факту, що між зарядженими частинками спостерігаються як сили притягання, так і сили відштовхування.

3. Заряд є величиною, що зберігається: в природі діє *закон збереження електричного заряду*, який гласить:

алгебраїчна сума зарядів усіх частинок електрично ізольованої системи[4] зберігається, тобто не змінюється ні при яких процесах у ній.

Цей закон може здатися тривіальним наслідком збереження кількості частинок в ізольованій системі. Але елементарні частинки здатні до взаємних перетворень: одні частинки можуть зникати і натомість з’являтися інші, у тому числі й заряджені. Тож кількість частинок в замкненій системі може змінюватись, але її електричний заряд лишається незмінним. А це означає, що електрично заряджені частинки можуть зникати й народжуватися не поодинці, а тільки парами протилежно заряджених частинок. Як приклад можна навести народження електрон-позитронної пари (e^-, e^+) при розпаді електрично нейтральної елементарної частинки — γ-кванта, що є носієм електромагнітного випромінювання. Тому закон збереження електричного заряду є одним з важливих законів, які “керують” процесами в мікросвіті.

4. Електричний заряд є інваріантною величиною, тобто не змінюється при переході від однієї системи відліку до іншої. Це природньо, адже заряд є не характеристикою стану, а внутрішньою властивістю частинки.

Електромагнітне поле. За сучасними уявленнями будь-яка силова взаємодія між частинками здійснюється не безпосередньо, а за допомогою відповідного *поля*. Зокрема, електромагнітна взаємодія між зарядженими частинками забезпечується *електромагнітним полем* за схемою “заряд — поле — заряд”: кожна заряджена частинка чи тіло створює у просторі електромагнітне поле, в якому на будь-яку іншу заряджену частинку чи тіло діє відповідна сила. Тому можна вказати на ще одну, первісну, властивість електричного заряду:

електричний заряд є джерелом електромагнітного поля.

На загал властивості електромагнітного поля крім величини зарядів, які його створюють, визначаються просторовим розташуванням і станом руху зарядів. Але в окремому випадку один із указаних факторів може виявитися не істотним. Приміром, якщо поле створюється нерухомими зарядженими тілами, то на його властивості в навколишньому просторі не впливають мікроскопічні теплові рухи зарядів всередині тіла. У такому разі говорять, що в просторі створюється *електричне поле*, і на вміщені в нього заряди діють *електричні сили*. А ось будь-яка ділянка провідника із постійним струмом є електрично нейтральною. Тому характеристики поля поза провідником визначаються тільки рухом його зарядів. У такому випадку говорять *про магнітне поле* та *магнітні сили*. Отож можна говорити, що електромагнітне поле складається з електричного та магнітного полів. Але електричне і магнітне поля не є двома різними фізичними полями — це два різні прояви єдиного електромагнітного поля. В цьому можна переконатися на такому простому прикладі. Досліджуючи поле, створене нерухомою зарядженою кулькою, ми побачимо, що воно є чисто електричним. Але якщо спостерігати поле тієї ж кульки з візка, що проїжджає повз неї, то буде виявлено як електричне, так і магнітне поле. Це показує, що поділ поля на електричне та магнітне є відносним і залежить від системи відліку, в якій воно розглядається. При цьому у випадку змінних полів електричне та магнітне поля не можуть існувати окремо одне від одного. Ці та інші факти переконують, що у природі існує єдине електромагнітне поле, котре поряд з речовиною є окремим видом матерії, а не просто зручним способом розгляду взаємодії між зарядами.

Навчальний матеріал, який розглядається в даному розділі курсу, розподілено по таких темах:

I. ЕЛЕКТРИЧНЕ ПОЛЕ У ВАКУУМІ

II. ЕЛЕКТРИЧНЕ ПОЛЕ В РЕЧОВИНІ

III. ЕЛЕКТРИЧНИЙ СТРУМ

IV. МАГНІТНЕ ПОЛЕ

V. ЕЛЕКТРОМАГНЕТИЗМ

---

[1] Фундаментальною називається взаємодія, що не зводиться ні до якої іншої.

[2] Слід зауважити, що за сучасними уявленнями електромагнітна та слабка взаємодії є проявами єдиної *електрослабкої* взаємодії. Але за конкретними властивостями ці взаємодії настільки різні, що їх розглядають окремо.

[3] Для прикладу, сила гравітаційного притягання між двома протонами приблизно в 10^{36} разів менша за силу їхнього електричного відштовхування.

[4] Мається на увазі система, в яку не можуть заходити та виходити з неї заряджені частинки.

І. ЕЛЕКТРИЧНЕ ПОЛЕ У ВАКУУМІ

Як говорилося вище, переважна більшість явищ, які ми спостерігаємо, має електромагнітну природу і визначається властивостями електричного та магнітного полів.

Далі розглядаються характеристики електричного поля у вакуумі. При цьому реально йдеться і про повітря, позаяк гази при невеликому тиску мають відносно невелику концентрацію молекул і практично не впливають на характеристики поля. Розглянуто наступні питання:

1. Напруженість електричного поля
2. Потенціал
3. Зв’язок між напруженістю і потенціалом
4. Електричний диполь
5. Циркуляція і потік електричного поля

1. Напруженість електричного поля

Електромагнітне поле ніяк не впливає на органи чуття людини, тому його властивості можна встановити лише опосередковано через властивості сил, які воно створює. При цьому електричні сили і поле досліджують за допомогою *“пробних зарядів”* — заряджених частинок з малими розмірами та величиною електричного заряду. Через малі розміри такі частинки відображають властивості електричного поля локально (в точці), а завдяки малій величині заряду не спотворюють його. Цей параграф присвячено основній величині, що характеризує електричне поле — напруженості. Розглянуто такі питання:

1.1. Вектор $\vec{E}$

1.2. Напруженість поля точкового заряду

1.3. Принцип суперпозиції.

1.4. Лінії напруженості

Контрольні запитання

1.1. *Вектор* $\vec{E}$

Експериментальні дослідження сил, які діють на пробні заряди в електричному полі, свідчать про таке:

1. У даній точці поля вектор сили $\vec{F}$, що діє на пробний заряд q, завжди напрямлений вздовж однієї й тієї ж прямій в той чи інший бік, залежно від знаку заряду. Це означає, що електричне поле в кожній точці має визначений напрям.

2. Величина сили, з якою поле діє на пробний заряд, є прямо пропорційною до величини цього заряду:

$$F \sim |q|.$$

Це означає, що відношення

$$E = \frac{F}{|q|}$$

не залежить від заряду і є мірою “потужності” самого поля. Воно називається напруженістю електричного поля і, аби відображало ще й напрям поля, розглядається як вектор:

$$\vec{E} = \frac{\vec{F}}{q}. \tag{1.1}$$

Сказане ілюструє рис. 1.1.

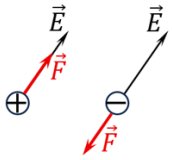


Рис. 1.1

Таким чином, вектор напруженості $\vec{E}$ є *силовою характеристикою* (силовим вектором) електричного поля. За модулем і напрямом він збігається із силою, що діє на одиничний позитивний заряд вміщений в дану точку поля.

За одиницю напруженості природньо було би прийняти 1 Н/Кл, тобто напруженість такого електричного поля, в якому на одиничний заряд діє одинична сила. Але в міжнародній системі одиниць (СІ) одиниця напруженості уводиться інакше і називається “один вольт-на-метр” (1 В/м). Про це буде сказано далі (п. 2.4), а тут лише зауважимо, що 1 В/м = 1 Н/Кл, тобто при E = 1 В/м на заряд 1 Кл діє сила 1 Н.

З означення (1.1) випливає загальна формула сили, що діє в електричному полі на точковий заряд:

$$\vec{F} = q\vec{E}. \tag{1.2}$$

Із цієї формули, зокрема, випливає, що сила, яка діє в полі на позитивний заряд, за напрямком є збіжною, а на негативний — антипаралельною з напрямом поля, як показано на рис. 1.1.

Формула (1.2) є однією з основних у теорії електричного поля, оскільки дозволяє обчислювати силу і відтак визначати рух заряджених частинок в електричних полях. Але для цього треба вміти знаходити напруженість поля незалежно від означення (1.1), тобто без вимірювання сил. Тому встановлення способів визначення напруженості є однією з основних задач теорії електричного поля.

1.2. *Напруженість поля точкового заряду*

Можливість теоретичного розрахунку напруженості різних електричних полів ґрунтується на двох основних законах, які встановлені на досліді — законі Кулона та принципі суперпозиції.

Закон Кулона. Цей закон був установлений ще до виникнення концепції поля і визначає електричну взаємодію точкових зарядів. Він твердить, що

сили взаємодії між двома точковими зарядами є прямо пропорційні величинам цих зарядів і обернено пропорційні квадратові відстані між ними.

Ці сили задовольняють третій закон Ньютона: вони рівні по модулю і діють у протилежних напрямах уздовж прямої, що проходить через заряди, рис. 1.2.

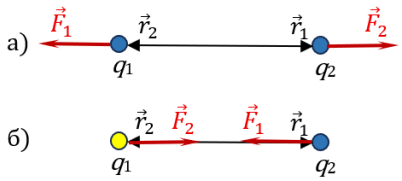


Рис. 1.2

У векторній формі ці сили виражаються, як

$$\vec{F}_1 = k\frac{q_1 q_2}{r^2}\vec{e}_{12}, \quad \vec{F}_2 = k\frac{q_1 q_2}{r^2}\vec{e}_{21}, \tag{1.3}$$

або

$$\vec{F}_1 = k\frac{q_1 q_2}{r^3}\vec{r}_1, \quad \vec{F}_2 = k\frac{q_1 q_2}{r^3}\vec{r}_2, \tag{1.3а}$$

де $\vec{e}_{12} = (\vec{r}_1/r_1)$ і $\vec{e}_{21} = (\vec{r}_2/r_2)$ — одиничні вектори, а $\vec{r}_1$ та $\vec{r}_2$ — радіуси-вектори точки розташування заряду, на який діє сила з боку поля, що створюється в цій точці іншим зарядом; $r = |\vec{r}_1| = |\vec{r}_2|$ — відстань між зарядами; k — коефіцієнт пропорційності.

Позаяк $\vec{r}_1 = -\vec{r}_2$, напрямки цих сил протилежні, а модулі однакові і дорівнюють

$$F = k\frac{|q_1|\cdot|q_2|}{r^2}. \tag{1.3б}$$

Отже, сили кулонівської взаємодії між зарядами задовольняють третій закон Ньютона. Векторні вирази (1.3а) відображають також і характер взаємодії зарядів. Справді, якщо заряди однойменні, то $q_1 \cdot q_2 > 0$, і вектори $\vec{F}_1$ і $\vec{F}_2$ мають напрямки радіусів-векторів $\vec{r}_1$ і $\vec{r}_2$, що відображає відштовхування однойменних зарядів (рис. 1.2*а*). Для різнойменних зарядів $q_1 \cdot q_2 < 0$, що відповідає притяганню між ними (рис. 1.2*б*).

Розмірний коефіцієнт k визначається силою взаємодії між двома одиничними зарядами у вакуумі при одиничній відстані між ними. Його числове значення залежить від системи одиниць і в СІ складає $k = 9\cdot 10^9$ м/Ф (Ф — одиниця електричної ємності “фарад”). Зауважимо, що розмірність [м/Ф] є рівносильною розмірності $[\text{Н}\cdot\text{м}/\text{Кл}^2]$, яка випливає з формули (1.3а). Число k є зручним для обчислень, але в літературі формули прийнято записувати через електричну сталу ε_0, що визначається як

$$\varepsilon_0 = \frac{1}{4\pi k} = 8{,}85\cdot 10^{-12}\ \frac{\text{Ф}}{\text{м}}.$$

Тому розгорнуто вирази (1.3а), (1.3б) записуються, як

$$\vec{F}_1 = \frac{q_1 q_2}{4\pi\varepsilon_0 r^3}\vec{r}_1,$$
$$\vec{F}_2 = \frac{q_2 q_1}{4\pi\varepsilon_0 r^3}\vec{r}_2, \qquad (1.3в)$$
$$F = \frac{|q_1|\cdot|q_2|}{4\pi\varepsilon_0 r^2}.$$

Напруженість поля точкового заряду. Сили (1.3в) зумовлені дією на кожен із двох зарядів з боку електричного поля, створеного іншим. Тому відповідно до означення (1.1), поділивши вектор сили, що діє на будь-який із зарядів на його величину, отримаємо загальний вираз вектора напруженості поля точкового заряду:

$$\vec{E} = \frac{q}{4\pi\varepsilon_0 r^3}\vec{r}, \qquad (1.4)$$

де q — алгебраїчна величина *заряду, що створює поле*, $\vec{r}$ — радіус-вектор точки, в якій визначається його напруженість (рис. 1.3).

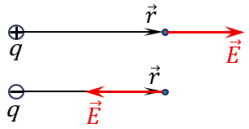


Рис. 1.3

Як видно з формули (1.4) і рис. 1.3, поле точкового заряду є центральним — воно скрізь напрямлене радіально від заряду, якщо $q > 0$, і до заряду при $q < 0$, та має модуль

$$E = \frac{|q|}{4\pi\varepsilon_0 r^2}, \quad \text{або} \quad E = \frac{k|q|}{r^2}. \qquad (1.4а)$$

Зауважимо також, що формули (1.4) і (1.4а) по суті виражають закон Кулона в польовій формі.

1.3. *Принцип суперпозиції*

Поле системи зарядів. Принцип суперпозиції. Уявімо, що вимірюють напруженість поля системи з двох зарядів у досліді, який складається з наступному трьох частин.

1. Спочатку у вільному просторі в якусь точку 1 вміщують заряд q_1 і за допомогою пробного заряду q та формули (1.1) визначають напруженість поля $\vec{E}_1$ у вибраній точці P (рис. 1.4*а*).

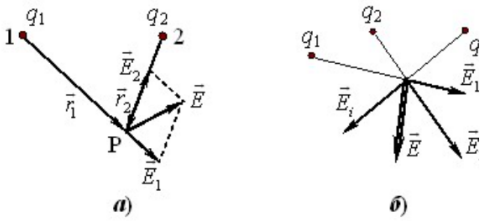


Рис. 1.4

2. Заряд q_1 прибирають, а в точку 2 вносять заряд q_2 і аналогічно визначають у точці P напруженість $\vec{E}_2$.

3. Не прибираючи заряду q_2, в точку 1 повертають заряд q_1 і визначають напруженість поля $\vec{E}$ в точці P, коли присутні обидва заряди.

Цей дослід, проведений реально, дає

$$\vec{E} = \vec{E}_1 + \vec{E}_2.$$

А це означає, що характеристики поля одного заряду не змінюються в присутності іншого. Такий самий результат виходить і при довільній кількості зарядів. Отже, можна зробити висновок, що

електричні поля діють на заряди, але не діють одне на одного.

Указану незалежність полів відображає принцип суперпозиції, який гласить:

вектор напруженості електричного поля, створеного заданою системою зарядів, в будь-якій точці дорівнює сумі векторів напруженості полів, які створюються в цій точці кожним із зарядів системи окремо (тобто за відсутності інших).

Математично принцип суперпозиції виражається загальною формулою

$$\vec{E} = \vec{E}_1 + \vec{E}_2 + \cdots + \vec{E}_i + \cdots = \\ = \sum \vec{E}_i \quad \Rightarrow \quad \vec{E} = \sum \vec{E}_i. \tag{1.5}$$

Згідно з формулою (1.4), для поля системи точкових зарядів q_i (рис. 1.4*б*) можна записати:

$$\vec{E} = \frac{1}{4\pi\varepsilon_0} \sum \frac{q_i}{r_i^3} \vec{r}_i. \tag{1.6}$$

Щодо наведеного формулювання принципу суперпозиції слід зауважити таке. Весь ряд доданків у сумі (1.5) можна згрупувати в декілька окремих блоків (наприклад, у два) і спочатку аналогічно обчислити напруженості $\vec{E}_{\text{I}}$ і $\vec{E}_{\text{II}}$ поля кожного блока. Тоді напруженість поля всієї системи виразиться, як

$$\vec{E} = \vec{E}_{\text{I}} + \vec{E}_{\text{II}}.$$

Це означає, що

напруженість електричного поля довільної системи зарядів дорівнює сумі напруженостей полів будь-яких (не обов'язково точкових) частин цієї системи.

Тому при обчисленні поля за принципом суперпозиції систему можна поділяти на окремі області так, як це найбільш зручно.

Поле неперервного розподілу заряду. За принципом суперпозиції можна розрахувати і поля, створені неточковими тілами, в яких заряд розподілений неперервно. Для цього тіло слід подумки розділити на елементарні області, кожна з яких містить гранично малий заряд $\mathrm{d}q$ і створює в заданій точці гранично слабке поле $\mathrm{d}\vec{E}$ (рис. 1.5*а*).

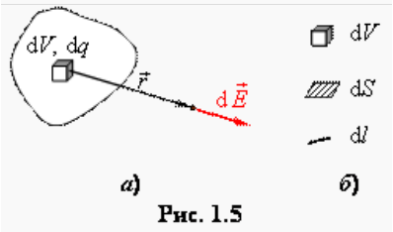


Рис. 1.5

Тоді поле, створюване всім тілом,

$$\vec{E} = \int \mathrm{d}\vec{E}.$$

Якщо елементарні області точкові, то поля $\mathrm{d}\vec{E}$ визначаються кулонівською формулою (1.4):

$$\mathrm{d}\vec{E} = \frac{k\,\mathrm{d}q}{r^3}\vec{r}.$$

У такому разі

$$\vec{E} = \int \frac{k\,\mathrm{d}q}{r^3}\vec{r} = \frac{1}{4\pi\varepsilon_0}\int \frac{\mathrm{d}q}{r^3}\vec{r}. \tag{1.7}$$

Ця інтегральна формула, яка є узагальненим виразом принципу суперпозиції, дає алгоритм розрахунку електричного поля створеного будь-яким заданим розподілом електричного заряду.

На практиці окремо розрізняють об'ємний, поверхневий і лінійний розподіли заряду і розділяють область його локалізації на елементарні *об'ємчики* $\mathrm{d}V$, *площинки* $\mathrm{d}S$, або *відрізки* $\mathrm{d}l$ (рис. 1.5*б*).

Об'ємний розподіл. При розподілі заряду по об'єму тіла величина $\mathrm{d}q$ у виразі (1.7) у кожній ділянці об'єму $\mathrm{d}V$ визначається, як $\mathrm{d}q = \rho(\vec{r})\mathrm{d}V$, через об'ємну густину заряду

$$\rho(\vec{r}) = \frac{\mathrm{d}q}{\mathrm{d}V}\left(\frac{\text{Кл}}{\text{м}^3}\right).$$

У такому разі

$$\vec{E} = k \int\limits_V \frac{\rho(\vec{r})\vec{r}}{r^3}\mathrm{d}V = \\ = \frac{1}{4\pi\varepsilon_0} \int\limits_V \frac{\rho(\vec{r})\vec{r}}{r^3}\mathrm{d}V, \tag{1.8}$$

де інтегрування проводиться по всьому об'єму зарядженого тіла.

Поверхневий розподіл. Аналогічно, якщо поле створюється зарядженою поверхнею, то заряд її довільної елементарної ділянки $\mathrm{d}S$ $\mathrm{d}q = \sigma(\vec{r})\mathrm{d}S$, де

$$\sigma = \frac{\mathrm{d}q}{\mathrm{d}s}\ \left(\frac{\text{Кл}}{\text{м}^2}\right)$$

— поверхнева густина заряду в заданій точці. Тоді

$$\vec{E} = k \int\limits_S \frac{\sigma\left(\vec{r}\right)\vec{r}}{r^3}\mathrm{d}S = \\ = \frac{q}{4\pi\varepsilon_0} \int\limits_S \frac{\sigma\left(\vec{r}\right)\vec{r}}{r^3}\mathrm{d}S, \tag{1.8а}$$

де інтегрування ведеться по всій зарядженій поверхні.

Лінійний розподіл. Нарешті, коли заряд розподілений по заданій лінії («нитці»), то $\mathrm{d}q = \lambda(\vec{r})\mathrm{d}l$, де

$$\lambda(\vec{r}) = \frac{\mathrm{d}q}{\mathrm{d}l}\ \left(\frac{\text{Кл}}{\text{м}}\right)$$

— лінійна густина заряду. В цьому випадку напруженість визначається через криволінійний інтеграл

$$\vec{E} = k \int\limits_l \frac{\lambda\left(\vec{r}\right)\vec{r}}{r^3}\mathrm{d}l = \\ = \frac{1}{4\pi\varepsilon_0} \int\limits_l \frac{\lambda\left(\vec{r}\right)\vec{r}}{r^3}\mathrm{d}l. \tag{1.8б}$$

Інтегральні вирази (1.8)—(1.8б) випливають із загальної формули (1.7), в основі якої лежить закон Кулона. Тому електричні поля, створювані не лише точковими, а й неперервно розподіленими зарядами, називаються кулонівськими полями.

Принагідно зазначимо, що з усіх наведених формул найбільш загальною є формула (1.8) для об'ємного розподілу заряду, бо поверхневий і лінійний розподіли є просто виродженими випадками об'ємного. Дійсно, заряджена поверхня є не поверхнею в математичному сенсі, а тонкою пластиною із якоюсь товщиною $h = \mathrm{const}$, тож $\mathrm{d}q = \rho h\,\mathrm{d}S$. Тому розподіл заряду замість об'ємної густини ρ зручніше задавати поверхневою густиною $\sigma = \rho h$ (Кл/м^2), яка визначає заряд одиниці поверхні пластинки. Так само нитка являє собою прямий чи викривлений тонкий стержень сталого перерізу S, об'єм елементарної ділянки котрого $\mathrm{d}V = S\,\mathrm{d}l$. Тоді $\mathrm{d}q = \rho S\,\mathrm{d}l$, або $\mathrm{d}q = \lambda\,\mathrm{d}l$, де $\lambda = \rho S$ — лінійна густина заряду (Кл/м). Нарешті, формула (1.8) вбирає і дискретний розподіл, тобто систему заряджених "кульок" малого об'єму ΔV_i. У такому разі при інтегруванні у виразі (1.8) в місцях локалізації заряду матимемо $\rho_i\,\Delta V_i = q_i$, а скрізь у решті об'єму $\rho\,\mathrm{d}V = 0$. Відтак інтеграл (1.8) вироджується в дискретну суму (1.6).

Що до практичних розрахунків, то способи обчислення наведених інтегралів розглядаються у вищій математиці. Проте в простих випадках достатньо розуміти фізичний зміст виразів (1.8), (1.8а) і (1.8б). У складних випадках, коли задача взагалі не має аналітичних (формульних) розв'язків, використовують комп'ютерні розрахунки. Але принципово використання закону Кулона та принципу суперпозиції є універсальним алгоритмом, який дозволяє розрахувати будь-яке кулонівське електричне поле.

1.4. *Лінії напруженості*

Поля безпосередньо не впливають на наші органи чуття. Тому, задля отримання наочного уявлення про поле, його в той чи інший спосіб зображують на папері чи екрані монітора. Одним із способів візуалізації поля є лінії напруженості (лінії поля, силові лінії, лінії вектора $\vec{E}$), тобто лінії, в кожній точці яких дотична збігається з вектором напруженості в цій точці (рис. 1.6).

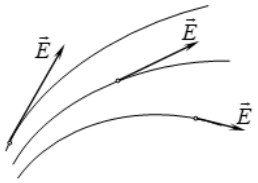
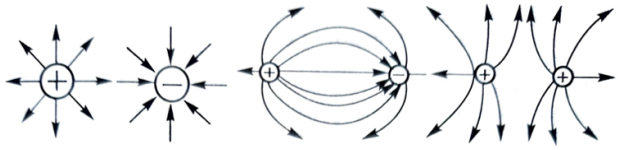


Рис. 1.6

Лінії поля відображають конфігурацію поля, тобто його напрям в різних точках, як це ілюструє рис. 1.7 для полів поодиноких зарядів і пари неподалік один від одного розташованих різнойменних зарядів.



а) б) в) г)

Рис. 1.7

Конфігурація полів інших систем зарядів є складнішою, але їхні силові лінії мають певні спільні властивості. А саме.

Лінії напруженості кулонівського електричного поля є незамкненими; вони виходять із позитивних і сходяться на негативних зарядах.

Також, позаяк в одній точці поле не може мати два різні напрями, лінії поля не можуть перетинатися.

Окрім того, силові лінії відображають і величину напруженості поля в різних місцях. Як видно з наведених на рис. 1.7 прикладів, поблизу зарядів, де поле є більш сильним, силові лінії розташовані щільніше, ніж на віддаленні. Отже, між лініями поля та величиною напруженості є певний зв'язок. Його легко встановити для поля точкового заряду, яке в усіх рівновіддалених на відстань r точках має однакову напруженість

$$E = \frac{q}{4\pi\varepsilon_0 r^2}$$

і радіальні силові лінії (рис. 1.7*а*, *б*). Тож, якщо для зображення цього поля використати N ліній, то крізь поверхню концентричної сфери з радіусом r вони будуть проходити з однаковою щільністю

$$n = \frac{N}{S} = \frac{N}{4\pi r^2}.$$

Якщо в цьому виразі на основі формули (1.4а) зробити заміну

$$4\pi r^2 = \frac{q}{\varepsilon_0 E},$$

то отримаємо:

$$E = \frac{q}{\varepsilon_0 N} n \quad \Rightarrow \quad E \sim n.$$

Тобто на будь-якій відстані від точкового заряду напруженість його поля є прямо пропорційною щільності силових ліній на цій відстані. Такий зв'язок зберігається й для більш складних полів. При цьому щільність ліній n в кожній точці визначається як

$$n = \frac{\mathrm{d}N}{\mathrm{d}S_\perp},$$

де $\mathrm{d}N$ — кількість ліній поля, що перетинають площадку $\mathrm{d}S_\perp$, перпендикулярну до напрямку поля.

Через будь-яку точку поля можна провести відповідну силову лінію, тож при зображенні поля кількість силових ліній можна вибирати на свій розсуд. Тому домовились уважати, що в кожній точці щільність ліній поля дорівнює числовому значенню його напруженості:

$$E = n. \tag{1.9}$$

Відтак можна вважати, що кількість силових ліній $\mathrm{d}N$, які перетинають перпендикулярну до поля елементарну площадку $\mathrm{d}S_\perp$, дорівнює

$$\mathrm{d}N = E\,\mathrm{d}S_\perp. \tag{1.10}$$

Така домовленість у деяких ситуаціях виявляється вельми зручною.

Контрольні запитання

1. Що таке фундаментальні взаємодії? Перелічіть їх.

2. Що є схожого і що відмінного між електромагнітною та гравітаційною взаємодією?

3. Що таке електричний заряд і які його основні властивості?

4. Опишіть польовий механізм взаємодій. Чим визначаються загальні властивості електромагнітного поля та його складових — електричного і магнітного полів?

5. Перелічіть загальні властивості електричних сил.

6. Що називається напруженістю електричного поля?

7. Сформулюйте та запишіть у векторній формі закон Кулона. Чи виконується він для довільних заряджених тіл?

8. Якою формулою визначається та як напрямлений вектор напруженості електричного поля точкового заряду?

9. Як записується та який фізичний зміст має принцип суперпозиції електричних полів?

10. Запишіть і поясніть загальний зміст виразів напруженості електричного поля, що створюється зарядженими не точковими тілами при різних типах розподілу заряду.

11. Чи може модуль напруженості поля системи з двох точкових зарядів дорівнювати: 1) сумі модулів і 2) модулю різниці модулів напруженості поля кожного із зарядів? Якщо так, то коли? Окремо розглянути випадки, коли заряди: а) однойменні і б) різнойменні

12. Запишіть вираз напруженості електричного поля посередині між двома однаковими точковими зарядами q розміщеними не відстані l один від одного, коли заряди: а) різнойменні і б) однойменні.

13. Чому дорівнює напруженість електричного поля в центрі правильного трикутника, якщо в його вершинах розміщені однойменні заряди однакової величини.

14. Електричне поле створюється системою з чотирьох однакових зарядів q розміщених у вершинах квадрата зі стороною a. Чому дорівнює напруженість цього поля на відстані $r \gg a$ від центра квадрата?

15. Чи можна говорити, що силові лінії збігаються з траєкторіями руху заряджених частинок в електричному полі? Чому?

16. Як за картиною силових ліній електричного поля встановити, де воно сильніше?

17. Чи можуть перетинатися або дотикатися силові лінії електричного поля? Чому?

2. Потенціал

У вступі говорилося, що в природі існує єдине електромагнітне поле, яке “складається” з електричного та магнітного полів. Але ці поля не є двома окремими полями і за відповідних умов можуть породжувати (індукувати) одне одного. Через це в природі є два види електричних полів: “кулонівські”, що безпосередньо створюються зарядженими частинками, і так звані “вихрові”, які породжуються магнітними полями. Силова дія обох полів, зокрема, їхня робота при переміщенні зарядів визначається вектором напруженості. Але кулонівське поле до того ж є потенціальним, тобто його робота не залежить від траєкторії руху зарядженої частинки та пройденого нею шляху. Тому кулонівське поле, крім напруженості, характеризується ще й скалярною величиною — потенціалом.

Далі розглядаються такі питання:

2.1. Потенціальність електричного поля зарядів

2.2. Різниця потенціалів і потенціал

2.3. Еквіпотенціальні поверхні

Контрольні запитання

2.1. *Потенціальність електричного поля зарядів*

У потенціальності електричного поля зарядів можна переконатися прямими розрахунками. Розглянемо спочатку роботу поля точкового заряду q при переміщенні іншого точкового заряду q' вздовж довільної траєкторії між заданими точками 1 і 2 (рис. 2.1).

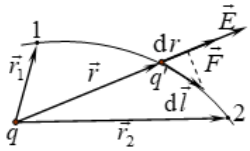


Рис. 2.1

На гранично малому переміщенні $\mathrm{d}\vec{l}$ поле виконує елементарну роботу $\delta A = \vec{F}\,\mathrm{d}\vec{l} = q'\vec{E}\,\mathrm{d}\vec{l}$. Робота на всьому шляху складається з усіх елементарних робіт δA і визначається криволінійним інтегралом по траєкторії:

$$A = \int\limits_{1\to 2} E\,|\mathrm{d}\vec{l}\,|\cos\alpha. \tag{2.1}$$

З рис. 2.1 видно, що величина $|\mathrm{d}\vec{l}\,|\cos\alpha = \mathrm{d}r$ є приростом відстані r при переміщенні $\mathrm{d}\vec{l}$. Тому, врахувавши формулу (1.4а), отримаємо:

$$\begin{aligned} A = \frac{q'q}{4\pi\varepsilon_0}\int\limits_{r_1}^{r_2}\frac{\mathrm{d}r}{r^2} \quad \Rightarrow \quad A = \\ = q'\left(\frac{q}{4\pi\varepsilon_0 r_1} - \frac{q}{4\pi\varepsilon_0 r_2}\right). \end{aligned} \tag{2.2}$$

Очевидно, знайдена робота залежить тільки від відстаней r_1 і r_2 від заряду q' до q, а не від форми траєкторії та пройденого ним шляху. Це вірно і для поля створеного довільною системою зарядів q_i. Дійсно, напруженість поля системи $\vec{E} = \sum\limits_i \vec{E}_i$, тому

$$\begin{aligned} A = \int\limits_{1\to 2} q'\left(\sum_i \vec{E}_i\right)\mathrm{d}\vec{l} = \\ = \sum_i\left(\int\limits_{1\to 2} q'\vec{E}_i\,\mathrm{d}\vec{l}\right) = \sum_i A_i, \end{aligned}$$

де роботи A_i складових полів $\vec{E}_i$ визначаються виразом (2.2). Тож

$$A = q' \left(\sum_i \frac{q_i}{4\pi\varepsilon_0 r_{i1}} - \sum_i \frac{q_i}{4\pi\varepsilon_0 r_{i2}} \right), \tag{2.2а}$$

де r_{i1} і r_{i2} — відстані від початкової та кінцевої точок переміщення заряду q_i до i-го заряду системи.

Таким чином, при переміщенні заряду в будь-якому кулонівському полі робота поля не залежить від траєкторії. Отже,

електричне поле, створене зарядженими частинками (кулонівське поле), є потенціальним полем.

Слід зауважити, що це не так для електричних полів, які породжується не зарядами, а змінним магнітним полем.

2.2. *Різниця потенціалів і потенціал*

Означення. У виразах (2.2) і (2.2а) величина в дужках не залежить від заряду q' і визначається тільки величиною поля у початковій та кінцевій точках переміщення. Тому вираз роботи поля при переміщенні заряду q між заданими двома точками по будь-якому шляху можна подати, як

$$A = q(\varphi_1 - \varphi_2) = -q\,\Delta\varphi, \tag{2.3}$$

а при нескінченно малому переміщенні — як

$$\delta A = -q\,\mathrm{d}\varphi. \tag{2.3а}$$

Величина

$$U = \varphi_1 - \varphi_2 \tag{2.4}$$

є однією з характеристик електричного поля і називається різницею потенціалів або напругою. Отже:

різницею потенціалів (напругою) між заданими двома точками називається величина, що чисельно дорівнює роботі поля при перенесенні між цими точками одиничного позитивного заряду.

Відтак формулу (2.3) можна записати і як

$$A = qU. \tag{2.5}$$

Напруга вимірюється у вольтах (В), причому різниця потенціалів (напруга) двома між точками дорівнює $1\ \text{В} = 1\ \text{Дж}/\text{Кл}$, якщо при перенесенні між ними заряду 1 Кл поле виконує роботу 1 Дж.

Таким чином, потенціальне поле визначається на лише силовою характеристикою — вектором напруженості $\vec{E}$, а й скалярною алгебраїчною величиною — потенціалом φ, який є енергетичною характеристикою поля. Але вираз (2.4) визначає тільки різницю потенціалів у двох точках. Що ж до потенціалу φ, то він сам по собі не є однозначно визначеним — якщо в усіх точках поля до величини φ додати будь-яке задане число, то це ніяк не вплине ні на результати обчислень роботи за формулою (2.3), ні на значення напруженості поля. Тому, аби потенціал у всіх точках був визначений однозначно, треба спочатку вибрати так звану нульову точку (або нульовий рівень) потенціалу — якусь точку P_0 (або множину точок), де потенціал за домовленістю приймається рівним нулю: $\varphi(P_0) = 0$. У такому разі формула (2.3) для роботи поля A_0 при переміщенні заряду з довільної точки P у нульову точку P_0 запишеться, як

$$A_0 = q\varphi,$$

і

$$\varphi = \frac{A_0}{q}. \tag{2.6}$$

Цей вираз по суті є означенням потенціалу:

потенціал електричного поля є визначеним тільки по відношенню до вибраної нульової точки і чисельно дорівнює роботі поля при перенесенні одиничного позитивного заряду з даної точки в нульову точку.

Слід зазначити, що вибір нульової точки, в принципі, є довільним і лімітується лише міркуваннями зручності. Наприклад, у випадку окремої зарядженої частинки чи системи зарядів обмежених розмірів поле послаблюється з відстанню, тому у фізиці нульову точку потенціалу вибирають “на нескінченності”, тобто в довільній точці віддаленій від системи на необмежено велику відстань. Виняток становлять тільки ідеалізовані поля необмежених розподілів заряду, таких, як нескінченна зарядженою площина, або нескінченний заряджений циліндр, тощо.

На практиці корпуси вузлові (“нульові”) точки схем електричних приладів сполучають із корпусом, а сам корпус із міркувань техніки безпеки “заземлюють”, тобто з’єднують із земною поверхнею. Тому в електротехніці за нуль приймається потенціал Землі.

При виборі нульової точки на нескінченності з виразів (2.2), (2.2а) і (2.3) випливають такі формули потенціалу поля окремого точкового заряду q та системи зарядів q_i:

$$\varphi = \frac{q}{4\pi\varepsilon_0 r}, \tag{2.7}$$

$$\varphi = \sum_i \frac{q_i}{4\pi\varepsilon_0 r_i}. \tag{2.7а}$$

Остання формула показує, що потенціал, як і напруженість поля, задовольняє принцип суперпозиції:

$$\varphi = \sum_i \varphi_i. \tag{2.8}$$

Це дозволяє розраховувати і потенціали полів неперервно розподілених зарядів за тою самою схемою, що й для напруженості (п. 1.3). А саме, весь заряд поділяють на окремі точкові ділянки $\mathrm{d}q$, потім за формулою (2.7) записують потенціал поля окремої ділянки $\mathrm{d}\varphi = k\,\mathrm{d}q/r$, і у виразі (2.8) дискретне додавання замінюють інтегруванням:

$$\varphi = k\int \frac{\mathrm{d}q}{r}. \tag{2.9}$$

Відтак, записавши $\mathrm{d}q$ через густину заряду та розміри ділянки, отримуємо наступні інтегральні формули для обчислення потенціалу при кожному можливому типі розподілу:

$$\varphi = k\int\limits_V \frac{\rho(\vec r)}{r}\mathrm{d}V = \frac{1}{4\pi\varepsilon_0}\int\limits_V \frac{\rho(\vec r)}{r}\mathrm{d}V. \tag{2.9а}$$

$$\varphi = k\int\limits_S \frac{\sigma(\vec r)}{r}\mathrm{d}S = \frac{1}{4\pi\varepsilon_0}\int\limits_S \frac{\sigma(\vec r)}{r}\mathrm{d}S, \tag{2.9б}$$

$$\varphi = k\int\limits_L \frac{\lambda(\vec r)}{r}\mathrm{d}\vec l = \frac{1}{4\pi\varepsilon_0}\int\limits_L \frac{\lambda(\vec r)}{r}\mathrm{d}\vec l, \tag{2.9в}$$

де $\rho(\vec r)$, $\sigma(\vec r)$ та $\lambda(\vec r)$ — функція об’ємної, поверхневої та лінійної густини заряду.

Наостанок згадаємо про таке. З механіки відомо, що при переміщенні тіла в потенціальному полі будь-якої природи робота сил поля дорівнює спаду потенціальної енергії тіла:

$$A = -\Delta W = W_1 - W_2. \tag{2.10}$$

Тож, зважившись на формулу (2.3), маємо такий загальний вираз потенціальної енергії зарядженої частинки в електричному полі:

$$W = q\varphi. \tag{2.10а}$$

2.3. *Еквіпотенціальні поверхні*

Наочне уявлення про розподіл потенціал поля в просторі можна скласти за допомогою еквіпотенціальних поверхонь, тобто таких,

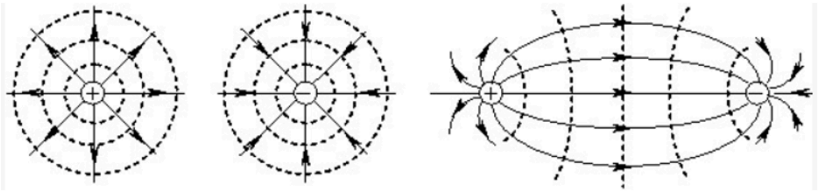
для кожної з яких потенціал поля в усіх точках однаковий: $\varphi = \mathrm{const}$.

Форма еквіпотенціальних поверхонь визначається видом залежності потенціалу від координат і для кожного значення потенціалу φ_0 встановлюється з рівняння $\varphi(\vec r) = \varphi_0$. Зокрема, для поля точкового заряду, згідно з виразом (2.7), виходить:

$$\frac{q}{4\pi\varepsilon_0 r} = \varphi_0 \quad\Rightarrow\quad r = \frac{q}{4\pi\varepsilon_0\varphi_0} = \mathrm{const}.$$

Отже, еквіпотенціальними поверхнями такого поля є сфери з центром у заряді, як показано штриховими лініями на рис. 2.2. (Зрозуміло, що на папері зображуються не самі еквіпотенціальні поверхні, а лінії їхнього перетину із площиною рисунка.)

Аби за допомогою еквіпотенціальних поверхонь наочно показати не тільки розподіл, а й зміну потенціалу при переміщенні в полі від точки до точки, їх (поверхні) проводять з однаковим кроком $\Delta\varphi = \text{const}$. У такому разі чим швидше змінюється потенціал, тим ближче одна до одної розташовуються еквіпотенціальні поверхні. Це стосується як поля точкового заряду, так і інших полів, наприклад, поля двох різнойменних зарядів.



а) б) в)

Рис. 2.2

Форма еквіпотенціальних поверхонь, як і форма ліній поля, в загальному випадку є складною і може взагалі не виражатись якимось рівнянням. Але за будь-яких умов у всіх точках силові лінії та еквіпотенціальні поверхні є взаємно ортогональними, тобто в точці перетину дотична до силової лінії (напрям $\vec{E}$) і площина, дотична до еквіпотенціальної поверхні, є взаємно перпендикулярними. Це випливає з того, що при будь-якому переміщенні $\mathrm{d}\vec{l}$ заряду q по еквіпотенціальній поверхні, згідно з (2.3а), робота поля $\delta A = q\vec{E}\,\mathrm{d}\vec{l} = -q\,\mathrm{d}\varphi = 0$, отже кут між векторами $\vec{E}$ і $\mathrm{d}\vec{l}$ дорівнює 90°.

Контрольні запитання

1. Що свідчить про потенціальність електричного поля зарядів?

2. Що називається різницею потенціалів у двох точках електричного поля? Потенціалом у заданій точці?

3. Потенціал у деякій точці електричного поля дорівнює 1 кВ. Що це означає?

4. Запишіть формулу потенціалу поля точкового заряду відносно землі, якщо земна поверхня має потенціал φ_0 відносно нескінченності.

5. Електричне поле нескінченної рівномірно зарядженої площини є однорідним (скрізь однаковим). Чи можна записати для нього формулу потенціалу? Чому?

6. Електричне поле створюється двома закріпленими однаковими різнойменними зарядами. Яку форму має і що собою являє еквіпотенціальна поверхня, котра проходить через середину відрізка, що з'єднує заряди?

7. Як, маючи картину еквіпотенціальних поверхонь електричного поля, можна показати вигляд його силових ліній?

3. Зв’язок між напруженістю та потенціалом

Із механіки відомо, що при переміщенні частинки в потенціальному полі роботу поля можна визначити як через силу, так і через потенціальну енергію. Тому роботу при переміщенні заряду в кулонівському електричному полі можна знайти як через напруженість поля, так і через потенціал. Отже, між напруженістю і потенціалом є органічний зв’язок, що дозволяє визначати одну з цих величин за відомою іншою. Далі розглянуто:

3.1. Визначення потенціалу поля через напруженість

3.2. Визначення напруженості поля через потенціал

3.3. Градієнт

Контрольні запитання

3.1. *Визначення потенціалу поля через напруженість*

Задачу про зв’язок між напруженістю та потенціалом можна розглядати двояко: як установлення способу визначення потенціалу поля $\varphi(\vec{r})$ через задану напруженість $\vec{E}(\vec{r})$, та як установлення способу визначення напруженості поля $\vec{E}(\vec{r})$ через заданий потенціал $\varphi(\vec{r})$.

Перша задача по суті вже була розглянута в п. 2.2. Справді, підставляючи у формулу (2.3) вираз роботи поля (2.1), отримуємо загальний вираз різниці потенціалів між заданими точками 1 і 2 через напруженість поля:

$$\varphi_1 - \varphi_2 = \int\limits_{1\to 2} \vec{E}\,\mathrm{d}\vec{l}\,. \tag{3.1}$$

Так само, попередньо обравши нульову точку P_0, за формулою (2.6) можна знайти й потенціал в будь-якій точці P:

$$\varphi = \int\limits_{P\to P_0} \vec{E}\,\mathrm{d}\vec{l}\,, \tag{3.2}$$

Як було показано раніше (п. 2.1), ці інтеграли не залежать від траєкторії, що дозволяє при розрахунках обирати найзручніший шлях інтегрування. Часто це силова лінія поля.

3.2. *Визначення напруженості поля через потенціал*

Проекція напруженості $\vec{E}$ на заданий напрям. Розглянемо роботу поля ΔA при невеликому переміщенні $\Delta\vec{l}$ заряду q у заданому напрямку між точками з потенціалами φ та $\varphi + \Delta\varphi$ (рис. 3.1).

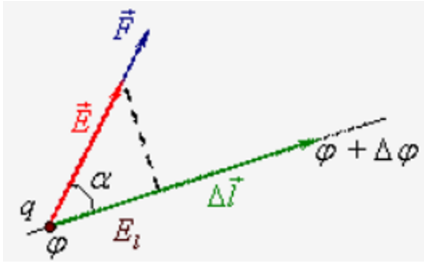


Рис.3.1¶

Згідно з (2.3), ця робота складає

$$\Delta A = -q\,\Delta\varphi.$$

На невеликому переміщенні зміни напруженості поля $\vec{E}$ та сили $\vec{F} = q\vec{E}$, що діє на заряд є незначними. Тому, вказану роботу можна наближено визначити через напруженість в початковій точці:

$$\Delta A \approx q\vec{E}\cdot\Delta\vec{l}\,.$$

Прирівнюючи обидва вирази роботи і враховуючи, що $\vec{E}\,\Delta\vec{l} = E\,\Delta l\cos\alpha = E_l\,\Delta l$ (див. рис. 3.1), отримаємо для проекції вектора $\vec{E}$ на напрям переміщення заряду q:

$$E_l \approx -\frac{\Delta\varphi}{\Delta l}.$$

При зменшенні відстані Δl похибка буде зменшуватись, і в границі $\Delta l \to 0$ цей вираз стане точним. Отже,

$$E_l = -\lim_{\Delta l \to 0} \frac{\Delta\varphi}{\Delta l}.$$

Границя у правій частині — то є швидкість зміни величини φ при переміщенні в заданому напрямку $\vec{l}$. Вона позначається як $\partial\varphi/\partial l$ і називається похідною функції φ за напрямом $\vec{l}$:

$$\frac{\partial\varphi}{\partial l} = \lim_{\Delta l \to 0} \frac{\Delta\varphi}{\Delta l}.$$

Таким чином,

$$E_l = -\frac{\partial\varphi}{\partial l}, \tag{3.3}$$

тобто

проекція вектора напруженості електричного поля на заданий напрям дорівнює взятій із зворотнім знаком похідній потенціалу в цьому напрямі.

Зв'язок між $\vec{E}$ та φ. Попри схожість в означенні та позначенні, похідна за напрямом не є звичною похідною, і в (3.3) символ ∂l означає не приріст якогось аргументу l, а величину нескінченно малого переміщення у напрямі l. Виняток становить рух уздовж координатних осей, коли елементарні переміщення $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}z$ є приростами аргументів x, y, z функції $\varphi(x, y, z)$. У такому випадку похідні у напрямках координатних осей вироджуються в частинні похідні по відповідних координатах, і згідно з (3.3)

$$E_x = -\frac{\partial\varphi}{\partial x}, \quad E_y = -\frac{\partial\varphi}{\partial y} \quad \text{і} \quad E_z = -\frac{\partial\varphi}{\partial z}.$$

Це дозволяє виразити через потенціал і вектор напруженості $\vec{E} = E_x\vec{i} + E_y\vec{j} + e_z\vec{k}$:

$$\vec{E} = -\left(\frac{\partial\varphi}{\partial x}\vec{i} + \frac{\partial\varphi}{\partial y}\vec{j} + \frac{\partial\varphi}{\partial z}\vec{k}\right).$$

Величина в дужках називається *градієнтом потенціалу* і позначається як $\operatorname{grad}\varphi$:

$$\operatorname{grad}\varphi = \frac{\partial\varphi}{\partial x}\vec{i} + \frac{\partial\varphi}{\partial y}\vec{j} + \frac{\partial\varphi}{\partial z}\vec{k}. \tag{3.4}$$

Отже,

$$\vec{E} = -\operatorname{grad}\varphi. \tag{3.5}$$

Вектор напруженості електричного поля в кожній точці дорівнює взятому з протилежним знаком градієнту потенціалу в цій точці.

Ще раз наголосимо, що йдеться тільки про потенціальне (кулонівське) поле.

3.3. *Градієнт*

Вирази (3.4) і (3.5) показують, як слід обчислювати напруженість поля через потенціал, заданий в декартових координатах, і не дають наочного уявлення про зв'язок напруженості з потенціалом та зміст поняття "градієнт". Аби висвітлити його, згадаймо, що в кожній точці вектор напруженості $\vec{E}$ напрямлений уздовж нормалі $\vec{n}$ до еквіпотенціальної поверхні, і, згідно з (3.3), проекція E_n і модуль E напруженості поля визначаються похідною вздовж нормалі:

$$E_n = \frac{\partial\varphi}{\partial n} \qquad \text{і} \qquad E = \left|\frac{\partial\varphi}{\partial n}\right|. \tag{3.6}$$

У випадку однорідного поля ($\vec{E} = \text{const}$), потенціал у будь-якому напрямку змінюється лінійно, тому

$$E = \frac{U}{d}, \tag{3.6а}$$

де $U = \varphi_1 - \varphi_2$ — різниця потенціалів (напруга) між двома точками, розміщеними одна від одної на відстані d у напрямку поля. На основі цієї формули встановлена одиниця напруженості поля “вольт-на-метр” (В/м):

1 В/м — це напруженість однорідного електричного поля, в якому між точками, розташованими на відстані 1 м у напрямку поля, напруга становить 1 В.

Відповідно до співвідношень (3.6) вектор $\vec{E}$ виражається, як

$$\vec{E} = -\frac{\partial \varphi}{\partial n}\vec{n},$$

де $\vec{n}$ — орт (одиничний вектор) нормалі до еквіпотенціальної поверхні, напрямлений в бік зростання потенціалу.

Зіставляючи цей вираз із виразом (3.5), бачимо, що

$$\operatorname{grad}\varphi = \frac{\partial \varphi}{\partial n}\vec{n}. \tag{3.7}$$

Цей результат можна розглядати як інваріантне (не залежне від системи координат) означення поняття градієнта потенціалу. Аби зрозуміти його зміст, згадаймо, що модуль вектора більший за модуль його проекції на будь-який інший напрям: $|\vec{E}| > |E|$. Це означає, що $\left|\frac{\partial \varphi}{\partial n}\right| > \left|\frac{\partial \varphi}{\partial l}\right|$, тобто напрям градієнта є напрямом найшвидшої зміни φ. Відтак можна дати наступне формулювання:

градієнтом потенціалу у заданій точці називається вектор, який спрямований в бік найшвидшого зростання потенціалу і має модуль, рівний швидкості його зміни в цьому напрямі.

Примітка. У математиці градієнт трактують як указану міру швидкості зміни будь-якої функції координат.

Відповідно, позаяк у виразі (3.5) фігурує знак “$-$”, вектор напруженості електричного поля у кожній точці напрямлений в бік найшвидшого зменшення потенціалу.

Контрольні запитання

1. Яка відміна існує між похідною за напрямом і частинною похідною?

2. Що є спільним і що відмінним у поняттях “похідна за напрямом” і “градієнт”?

3. Як відрізняється похідна потенціалу у напрямку силової лінії поля від похідної в якомусь іншому напрямку?

4. Чому не можна сказати, що вектор напруженості електричного поля є похідною потенціалу в напрямку силової лінії?

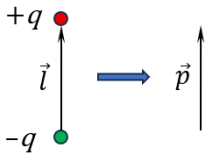
5. Часто в репортажі про велоперегони коментатори вживають вислів “градієнт траси”. Що вони мають на увазі і чи відповідає це математичному означенню поняття “градієнт”?

6. Як за картиною еквіпотенціальних поверхонь узнати, де поле сильніше, а де слабше?

7. Електричне поле створюється двома паралельними нескінченними площинами, що різноіменно заряджені з однаковими густинами заряду σ і $-\sigma$. Що можна сказати про напруженість і потенціал поля поза площинами?

4. Електричний диполь

Електричним диполем (двополюсником) називаються два різнойменні заряди (полюси) $+q$ і $-q$ однакової величини, що розміщені на малій відстані l один від одного (див. рис.).



Диполь у цілому є електронейтральним, але через рознесеність полюсів у просторі він створює власне електричне поле та зазнає дії зовнішніх полів. При цьому власне поле диполя та його поведінка у зовнішньому полі визначаються не безпосередньо зарядом полюса, а електричним (дипольним) моментом $\vec{p}$ — величиною, що визначається як

$$\vec{p} = q\vec{l}, \tag{4.1}$$

де $q > 0$ — заряд диполя, а вектор $\vec{l}$, проведений від негативного до позитивного полюса, називається плечем диполя. На рисунках диполь зазвичай, зображують струлкою, напрямленою від негативного до позитивного заряду, і вказують позначення вектора дипльного моменту $\vec{p}$.

Варто завважити, що розглянута система з двох заряджених кульок майже не має прямого практичного застосування. Але диполь є цінною моделлю, що використовується в різних розділах електродинаміки. Зокрема, уявлення про електричний диполь є базовим у теорії електричного поля в діелектриках. Далі розглядаються такі питання:

4.1. Електричне поле диполя

4.2. Диполь у зовнішньому електричному полі

4.1. *Електричне поле диполя*

Розглянемо електричне поле, що створюється диполем із моментом $\vec{p}$ на відстанях $r \gg l$ від його центра (у так званій дипольній зоні). Очевидно, що це поле має осьову симетрію: його характеристики однакові у всіх точках кола з центром на осі диполя й перпендикулярного до неї. Тому напруженість і потенціал поля диполя у будь-якій точці P залежать тільки від відстані r та кута ϑ (рис. 4.1*а*):

$$\vec{E} = \vec{E}(r, \vartheta), \qquad \varphi = \varphi(r, \vartheta).$$

Потенціал поля диполя. Розрахунок поля диполя почнемо з потенціалу, який за принципом суперпозиції дорівнює

$$\varphi = \varphi_- + \varphi_+,$$

де потенціали φ_- і φ_+ полів кожного з полюсів визначаються формулою (2.7). Отже,

$$\varphi = \frac{-q}{4\pi\varepsilon_0 r_-} + \frac{q}{4\pi\varepsilon_0 r_+} = \\ = \frac{q}{4\pi\varepsilon_0}\left(\frac{1}{r_+} - \frac{1}{r_-}\right) = \frac{q}{4\pi\varepsilon_0}\cdot\frac{r_- - r_+}{r_+ r_-}.$$

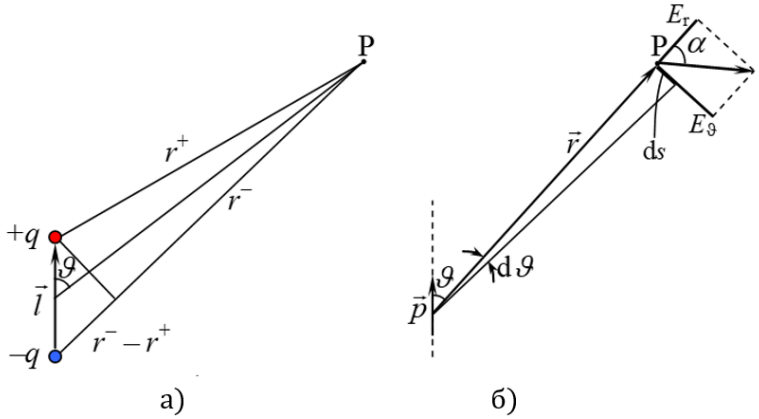


Рис. 4.1

Оскільки за умовою $r \gg l$, різниця довжин відрізків r_- і r_+ та кут між ними є малими (рис. 4.1), можна прийняти $r_- - r_+ = l\cos\vartheta$ і $r_+ r_- = r^2$. Тоді

$$\varphi(r,\vartheta) = \frac{p\cos\vartheta}{4\pi\varepsilon_0 r^2}. \tag{4.2}$$

Отже, як і говорилося, потенціал поля диполя визначається не окремо величиною зарядів і відстанню між ними, а дипольним моментом.

Напруженість поля диполя. Вектор напруженості теж можна знайти за принципом суперпозиції як $\vec{E} = \vec{E}_+ + \vec{E}_-$. Але корисніше зробити це за допомогою функції потенціалу (4.2) через проекції E_r на радіус-вектор $\vec{r}$ точки P і E_ϑ — на дотичну до кола радіуса r із центром у центрі диполя (рис. 4.1*б*). Ці проекції, відповідно до формули (3.3), визначаються як

$$E_r = -\frac{\partial\varphi}{\partial r}, \qquad E_\vartheta = -\frac{\partial\varphi}{\partial s} = -\frac{\partial\varphi}{r\,\partial\vartheta}.$$

Тут враховано, що елементарне переміщення $\mathrm{d}s$ у напрямку дотичної збігається з дугою кола: $\mathrm{d}s = r\,\mathrm{d}\vartheta$. Відтак, після указаних диференціювань функції (4.2), отримуємо:

$$\begin{cases} E_r = \dfrac{1}{4\pi\varepsilon_0}\dfrac{2p\cos\vartheta}{r^3}, \\ E_\vartheta = \dfrac{1}{4\pi\varepsilon_0}\dfrac{p\sin\vartheta}{r^3}. \end{cases} \tag{4.3}$$

Звідси для модуля напруженості $E = \sqrt{E_r^2 + E_\vartheta^2}$ маємо:

$$E = \frac{p}{4\pi\varepsilon_0 r^3}\sqrt{1 + 3\cos^2\vartheta}. \tag{4.4}$$

Напрям вектора $\vec{E}$ задамо кутом α між ним і радіусом-вектором $\vec{r}$ точки P (рис. 4.1*б*). Тоді, згідно з (4.3),

$$\operatorname{tg}\alpha = \frac{1}{2}\operatorname{tg}\vartheta. \tag{4.5}$$

Формули (2.23) і (2.24) повністю визначають вектор $\vec{E}$ у будь-якій точці дипольної зони ($r \gg l$), хоч і не є дуже наочними. Але для точок на осі диполя та на перпендикулярі до осі, що проходить через центр диполя, вектор напруженості можна визначити однією формулою. В першому випадку кут ϑ складає або 0, або π, тож відповідно до (4.5), такі самі значення має і кут α. Це означає, що на осі напрям поля скрізь збігається з напрямком електричного моменту диполя $\vec{p}$. Врахувавши це та підставивши в (4.4) значення $\cos^2\vartheta = 1$, отримуємо:

$$\vec{E}_\parallel = \frac{1}{4\pi\varepsilon_0}\cdot\frac{2\vec{p}}{r^3}. \tag{4.6}$$

Прийнявши $\vartheta = \pi/2$, так само одержимо вираз вектора напруженості в другому випадку:

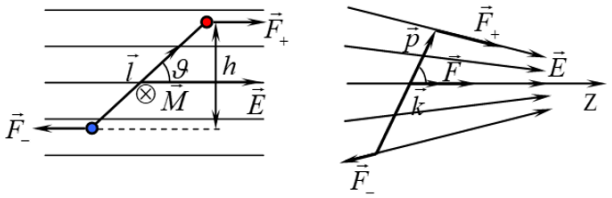
$$\vec{E}_\perp = -\frac{1}{4\pi\varepsilon_0}\cdot\frac{\vec{p}}{r^3}. \tag{4.7}$$

З приводу цього виразу корисно зауважити, що в усіх таких точках $\cos\vartheta = 0$, тож відповідно до (4.2), $\varphi = 0$. Це наочно ілюструє специфіку зв'язку між напруженістю та потенціалом електричного поля (формула (3.5)). А саме, відсутність потенціалу в заданій точці ще не означає відсутності самого поля. Про відсутність поля свідчить не величина потенціалу, а його незмінність при зміщенні з даної в будь-яку сусідню точку.

Підводячи підсумок, зазначимо, що в цілому електрично нейтральна система зарядів, якою є диполь, навіть на великих порівняно з її розмірами відстанях створює електричне поле. Але його потенціал і напруженість змінюються з відстанню швидше, ніж у випадку поля точкового заряду. Це є характерним і для полів мультиполів — нейтральних систем зарядів складнішої конфігурації. Причому, чим складніша конфігурація, тим сильніша залежність поля від відстані. Тому електричні поля таких систем, як атоми та молекули, існують практично тільки на відстанях порядку розмірів цих частинок.

4.2. *Диполь у зовнішньому електричному полі*

Момент сил, що діють на диполь. Якщо диполь із зарядом q і плечем $\vec{l}$ вміщено в однорідне електричне полі $\vec{E} = \text{const}$, то на полюси діють протилежно напрямлені сили $\vec{F}_+$ і $\vec{F}_-$ (рис. 4.2*а*), які намагаються повернути диполь за годинниковою стрілкою.



а) б)

Рис. 4.2

Обертальна дія сил визначається вектором моменту сил $\vec{M}$. У нашому випадку його легко розрахувати, позаяк сили є однакові за величиною та протилежні за напрямом, тож утворюють так звану “пару сил”. А з механіки відомо, що модуль моменту $\vec{M}$ пари дорівнює

$$M = Fh,$$

де відстань між лініями дії сил, яка називається “плечем” пари.

Напрям вектора моменту сил визначається за правилом правого гвинта. Тут це напрям “від нас”.

(На рис. 4.2*а* це позначено як $\otimes$. Аналогічно, напрям “до нас” позначають як $\odot$.)

З рисунка видно, що $h = l \sin\vartheta$, отже

$$M = qEl \sin\vartheta = pE \sin\vartheta, \tag{4.8}$$

де $p = ql$ — модуль електричного моменту диполя.

З урахуванням напрямків момент сил визначається векторним добуток векторів $\vec{p}$ і $\vec{E}$:

$$\vec{M} = \left[\vec{p}\vec{E}\right]. \tag{4.9}$$

У неоднорідному полі $\vec{E} \neq \text{const}$ сили $\vec{F}_+$, $\vec{F}_-$ не утворюють пари (див. рис. 4.2*б*), і формула (4.9) строго є не дійсною. Але для елементарного (із дуже малим плечем) диполя вона зберігає чинність, бо в межах диполя поле лишається практично однорідним. Тому можна сказати, що

будь-яке електричне поле намагається зорієнтувати елементарний диполь так, аби напрям дипольного моменту збігався із напрямом поля.

Сила, що діє на диполь у неоднорідному полі. В однорідному полі результуюча сила, що діє на диполь, $\vec{F} = \vec{F}_+ + \vec{F}_- = 0$. Тому поле лише повертає диполь навколо нерухомого центра мас. Але в неоднорідному полі (рис. 4.2*б*) $\vec{F}_+ \neq -\vec{F}_-$, і на диполь діє ще й деяка результуюча сила $\vec{F}_+ + \vec{F}_- \neq 0$. Тому вільний диполь буде не лише повертатися навколо центра мас, а й рухатись як ціле.

Указана сила визначається величиною та орієнтацією електричного моменту диполя $\vec{p}$ і конфігурацією та ступенем неоднорідності зовнішнього поля. Її загальний вираз легко знайти:

$$\vec{F} = \vec{F}_+ + \vec{F}_- = q\big(\vec{E}_+ - \vec{E}_-\big) = q\,\Delta\vec{E}. \tag{4.10}$$

де $\Delta\vec{E}$ — зміна напруженості поля при переміщенні на відстань l у напрямку плеча диполя. При малій величині l зміна поля є незначною, тож можна записати:

$$\Delta\vec{E} = \frac{\partial\vec{E}}{\partial l}l,$$

де $\partial\vec{E}/\partial l$ — похідна вектора $\vec{E}$ у напрямку $\vec{l}$, яка визначає швидкість зміни напруженості поля в напрямку плеча диполя. Відтак вираз (4.10) набуває вигляду:

$$\vec{F} = p\frac{\partial\vec{E}}{\partial l}, \tag{4.11}$$

де $p = ql$ — електричний момент диполя.

Слід зауважити, що попри позірну простоту формули (4.11), властивості сили $\vec{F}$ у деталях є досить складними. Але ситуація спрощується у випадку елементарного диполя в осесиметричному неоднорідному полі, напруженість якого змінюється тільки у напрямку самого поля, як схематично показано на рис. 4.2*б*. В такому випадку якщо спрямувати координатну вісь Z (орт $\vec{k}$) у напрямку поля, то $\vec{E} = E\vec{k}$, а елементарне переміщення вздовж осі диполя, що фігурує у виразі (4.11), $\partial l = (\partial z/\cos\vartheta)$. Отже,

$$\vec{F} = p_z \frac{\partial E}{\partial z}\vec{k}, \tag{4.12}$$

де $p_z = p\cos\vartheta$ — проекція моменту диполя на напрям поля, а вектор $(\partial E/\partial z)\vec{k}$ — градієнт напруженості поля, в якому знаходиться диполь.

За вказаних умов вектор сили $\vec{F}$ є колінеарним до осі Z, тож вираз сили зручно записати через проекцію $F_z = \pm F$

$$F_z = p_z \frac{\partial E}{\partial z}. \tag{4.13}$$

Оскільки вісь Z на рис. 4.2 спрямована в бік підсилення поля, то у виразах (4.12) і (4.13) $\partial E/\partial z > 0$. Це означає, що при орієнтації по полю $\vartheta < 90^\circ$ диполь втягується в область сильнішого поля, а при протилежній орієнтації ($\vartheta > 90^\circ$) — виштовхується в бік слабшого поля. Якщо ж у неоднорідне поле внести диполь у довільній орієнтації та вивільнити, то він почне одночасно повертатися і рухатись у той чи інший бік, але врешті зорієнтується по полю і рухатиметься в бік зростання напруженості.

Енергія диполя в зовнішньому полі. Електричне поле діє не на диполь як ціле, а на кожен заряд окремо. Тому, попри електричну нейтральність, диполь у зовнішньому полі має відповідну потенціальну енергію, яку можна визначити через роботу поля при зміні орієнтації диполя.

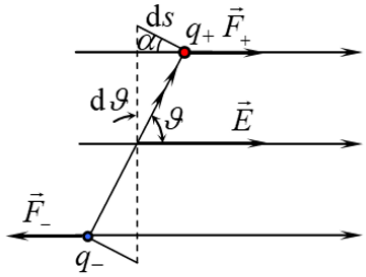


Рис. 4.3

Нехай диполь із зарядом q і плечем $\vec{l}$ є зорієнтований під кутом ϑ до напрямку електричного поля $\vec{E}$, рис. 4.3. Тоді при повороті диполя проти годинникової стрілки на елементарний кут $\mathrm{d}\vartheta$ заряди диполя переміщуються на відстань $\mathrm{d}s$, а сили поля $F = qE$ виконують сумарну роботу $\delta A = 2F\cos\vartheta\,\mathrm{d}s$. Із рис. 4.3 видно, що $\mathrm{d}s = (l/2)\,\mathrm{d}\vartheta$, і $\alpha = 90^\circ - \vartheta$. Отже, $\cos\alpha = -\sin\vartheta$, і

$$\delta A = -qEl\sin\vartheta\cdot\mathrm{d}\vartheta = -pE\sin\vartheta\cdot\mathrm{d}\vartheta,$$

де $p = ql$ — модуль електричного моменту диполя.

Звідси, згідно із співвідношенням (2.10), можна знайти приріст потенціальної енергії диполя при заданій зміні кута ϑ від ϑ_1 до ϑ_2:

$$W_2 - W_1 = -A = pE\int\limits_{\vartheta_1}^{\vartheta_2}\sin\vartheta\cdot\mathrm{d}\vartheta = \\ = pE\big(\cos\vartheta_1 - \cos\vartheta_2\big).$$

Переписавши цей вираз як

$$W_2 - W_1 = \big(-pE\cos\vartheta_2\big) - \big(-pE\cos\vartheta_1\big),$$

бачимо, що потенціальна енергія диполя у зовнішньому полі визначається формулою

$$W = -pE\cos\vartheta, \tag{4.14}$$

або

$$W = -\vec{p}\vec{E}. \tag{4.14а}$$

При цьому $W = 0$, коли $\cos\vartheta = 0$, тобто за нульовий рівень енергії приймається енергія диполя в положенні коли його вісь перпендикулярна до напрямку поля. Відповідно, при $\vartheta = 0^\circ$ енергія диполя є мінімальною і складає $W_{\min} = -pE$, а при $\vartheta = 180^\circ$ вона максимальна і дорівнює $W_{\max} = pE$. У відповідності до законів механіки в обох цих положеннях диполь перебуває в рівновазі, але в першому вона є стійкою, тоді як в другому — нестійкою.

5. Циркуляція і потік електричного поля

Окрім напруженості $\vec{E}$, що визначає електричне поле локально (в кожній точці), існують ще й інтегральні величини, котрі характеризують його не в точці, а на заданій лінії або поверхні. Ці величини називаються “циркуляцією” та “потоком” поля і розглядаються в наступних питаннях:

5.1. Циркуляція потенціального електричного поля

5.2. Електростатична теорема Гаусса

5.3. Обчислення полів за допомогою теореми Гаусса

Контрольні запитання

5.1. *Циркуляція потенціального електричного поля*

Співвідношення (3.5) дозволяє розраховувати напруженість не прямо, а через потенціал. Це може спростити задачу, бо визначати скалярні інтеграли (2.9) зазвичай легше, ніж векторні (1.7). Але такий шлях є прийнятним тільки для потенціальних полів. Тому треба мати формальний (математичний) критерій, який дозволяє розрізняти потенціальні та непотенціальні поля.

Для встановлення такого критерію розглянемо роботу A_{12} кулонівського поля при переміщенні заряду з точки 1 у точку 2 по двох різних траєкторіях I і II, рис. 5.1.

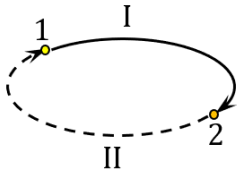


Рис. 5.1

Позаяк таке поле є потенціальним, то за означенням ці роботи однакові:

$$A_{12}^{(\mathrm{I})} = A_{12}^{(\mathrm{II})}.$$

Тому, якщо заряд перемістити з точки 1 в точку 2 по шляху (I) і потім повертати назад по шляху (II), то на зворотньому шляху робота $A_{21}^{(\mathrm{II})}$ буде відрізнятися від величини $A_{12}^{(\mathrm{I})}$ тільки знаком. Отже, на всьому замкненому шляху робота поля

$$A_0 = A_{12}^{(\mathrm{I})} + A_{21}^{(\mathrm{II})} = 0. \tag{5.1}$$

Із механіки відомо, що робота сили $\vec{F}$ на заданій замкненій траєкторії L виражається криволінійним інтегралом по цій траєкторії від скалярного добутку сили на переміщення:

$$A_0 = \oint\limits_L \vec{F}\,\mathrm{d}\vec{l}\,.$$

В електричному полі $\vec{F} = q\vec{E}$, отже, враховуючи вираз (5.1), для потенціального (кулонівського) електричного поля маємо:

$$\oint\limits_L \vec{E}\,\mathrm{d}\vec{l} = \oint\limits_L E_l\,\mathrm{d}l = 0, \tag{5.2}$$

де E_l — проекція напруженості на напрям обходу траєкторії інтегрування.

Інтеграл

$$C = \oint\limits_L \vec{E}\,\mathrm{d}\vec{l} = \oint\limits_L E_l\,\mathrm{d}l \tag{5.2а}$$

має спеціальну назву — циркуляція. Тобто,

к***риволінійний інтеграл від скалярного добутку вектора напруженості поля на вектор елементарного переміщення вздовж заданого замкненого контуру називається циркуляцією поля по цьому контуру.***

Можна сказати, що циркуляція дорівнює роботі поля при переміщенні по замкненому контуру одиничного заряду.

Поняття циркуляції і рівність (5.2) дозволяють сформулювати наступний інтегральний критерій потенціальності поля:

умовою потенціальності поля є нульова величина його циркуляції по будь-якому замкненому контуру, що повністю розташований у полі.

Саме таким є кулонівське поле — електричне поле, що створюється зарядженими частинками.

Зауважимо також, що, позаяк циркуляція по суті є питомою роботою поля на замкненому контурі, вона може мати той чи інший знак залежно від обраного напрямку обходу контуру при інтегруванні.

5.2. *Електростатична теорема Гаусса*

Зміст, властивості та значення другої з інтегральних характеристик електричного поля — *потоку* розглядаються далі в наступних питаннях:

5.2.1. Потік електричного поля

5.2.2. Теорема Гаусса

5.2.1. *Потік електричного поля*

Поняття та означення потоку, які використовуються нижче, походить із механіки суцільних середовищ.

Потік рідини. У повсякденному вжитку слово “потік” означає характерний рух рідини чи газу, скажімо, гірський чи повітряний потік. Але в науковій і технічній лексиці цим терміном позначають також кількісну міру різних процесів переносу: перенесення рідини, газу, чи сипучої речовини при транспортуванні трубопроводами, тощо.

Нехай в якійсь області простору переноситься речовина, приміром, тече вода в річці. Кількісна міра перенесення води через вибрану на її шляху задану поверхню називається *потоком* крізь цю поверхню і визначається як *об’єм води, що проходить крізь неї за одиницю часу* ($м^3/с$):

$$\Phi = \frac{\mathrm{d}V}{\mathrm{d}t}. \tag{5.3}$$

Зрозуміло, що потік рідини залежить від швидкості течії та характеристик поверхні. Аби з’ясувати, як саме, виділимо на вибраній поверхні S (рис. 5.2*а*) довільну елементарну площинку $\mathrm{d}S$ і покажемо орт (одиничний вектор) нормалі до неї $\vec{n}$ та вектор швидкості течії $\vec{v}$ в цьому місці.

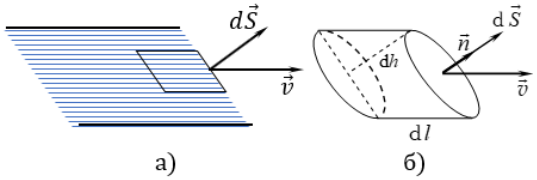


Рис. 5.2

Відтак побудуємо на площинці $\mathrm{d}S$ як на основі косий циліндрик із твірною $\mathrm{d}l$ та об’ємом $\mathrm{d}V = \mathrm{d}S \cdot \mathrm{d}h = \mathrm{d}S \cdot \mathrm{d}l \cos\alpha$, як показано на рис. 5.2*б*. Очевидно, що вся рідина в об’ємі циліндрика проходить крізь площинку $\mathrm{d}S$ за час $\mathrm{d}t = \mathrm{d}l/v$. Тому, згідно з (5.3), елементарний потік $\mathrm{d}\Phi$ крізь площинку $\mathrm{d}S$ дорівнює

$$\mathrm{d}\Phi = v\,\mathrm{d}S\cos\alpha. \tag{5.4}$$

Потік Φ крізь поверхню S складається з елементарних потоків $\mathrm{d}\Phi$ через усі її ділянки: $\Phi = \int\limits_S \mathrm{d}\Phi$, отже

$$\Phi = \int\limits_S v\,\mathrm{d}S\cos\alpha, \tag{5.4а}$$

де індекс S показує, що інтегрування проводиться по всій поверхні S.

Як бачимо, потік крізь задану поверхню визначається як її величиною, так і орієнтацією її ділянок відносно напрямку перенесення рідини. Тому в теорії елементарну площинку трактують як вектор $\mathrm{d}\vec{S} = \vec{n}\,\mathrm{d}S$. Тим самим вектор $\mathrm{d}\vec{S}$ відображає як величину, так і орієнтацію площинки в просторі. Це дозволяє лаконічно записати вирази (5.4) і (5.4а) у векторній формі:

$$\mathrm{d}\Phi = \vec{v}\,\mathrm{d}\vec{S}$$

і

$$\Phi = \int\limits_S \vec{v}\,\mathrm{d}\vec{S}. \tag{5.5}$$

Швидкість руху частинок $\vec{v}$ є скрізь визначеною і такою, що закономірно змінюється від точки до точки, тобто є заданою функцією координат: $\vec{v} = \vec{v}(\vec{r})$. У математиці векторну функцію координат, визначену в заданій області простору, називають полем вектора, або векторним полем. Отже, мовою математики об'єм рідини, що проходить крізь задану поверхню за одиницю часу, дорівнює потокові поля вектора швидкості течії крізь цю поверхню, або коротко — потокові вектора $\vec{v}$.

Потік електричного поля. Поняття потоку і означення (5.4) і (5.4а) або (5.5) зберігають чинність та корисність і для векторних полів будь-якої іншої природи. Зокрема,

інтеграл по заданій поверхні S від скалярного добутку вектора напруженості електричного поля $\vec{E}$ на вектор елемента поверхні $\mathrm{d}\vec{S}$ називається потоком напруженості (або потоком поля) крізь цю поверхню:

$$\Phi = \int\limits_S \vec{E}\,\mathrm{d}\vec{S}. \tag{5.6}$$

Підінтегральний вираз $\mathrm{d}\Phi = \vec{E}\,\mathrm{d}\vec{S}$, що визначає потік напруженості крізь довільну елементарну ділянку $\mathrm{d}S$ поверхні S, можна записати й інакше, розкривши скалярний добуток:

$$\begin{aligned} \mathrm{d}\Phi = \vec{E}\,\mathrm{d}\vec{S} = E\,\mathrm{d}S\cos\alpha = \\ = E_n\,\mathrm{d}S = E\,\mathrm{d}S_\perp, \end{aligned} \tag{5.7}$$

де E і $\mathrm{d}S$ — модулі векторів $\vec{E}$ і $\mathrm{d}\vec{S}$, E_n — проекція $\vec{E}$ на нормаль до поверхні, і $\mathrm{d}S_\perp$ — проекція площинки $\mathrm{d}\vec{S}$ на площину перпендикулярну до напрямку поля (рис. 5.3).

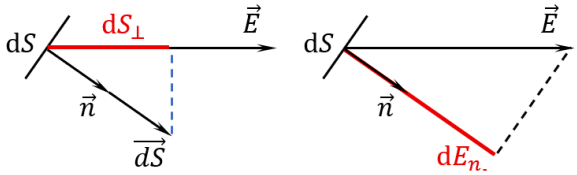


Рис. 5.3

Відповідно, потік поля крізь усю поверхню теж можна подати в різний спосіб:

$$\begin{aligned} \Phi = \int\limits_S \vec{E}\,\mathrm{d}\vec{S} = \int\limits_S E\,\mathrm{d}S\cos\alpha = \\ = \int\limits_S E_n\,\mathrm{d}S = \int\limits_S E\,\mathrm{d}S_\perp. \end{aligned} \tag{5.7а}$$

Говорячи про потік поля, слід зауважити таке. Потік є алгебраїчною величиною і, залежно від кута α між векторами $\vec{E}$ і $\mathrm{d}\vec{S}$, може бути додатнім чи від'ємним. З іншого боку, напрям векторів $\mathrm{d}\vec{S}$, відтак і знак Φ, є невизначеними, позаяк орти $\vec{n}$ можна провести як по один, так і по інший бік поверхні, причому для незамкнених поверхонь — все одно як. Проте у випадку замкнених поверхонь знак потоку набуває суттєвого змісту, тому за домовленістю в усіх точках замкненої поверхні вектори $\mathrm{d}\vec{S}$ напрямляють назовні (рис. 5.4).

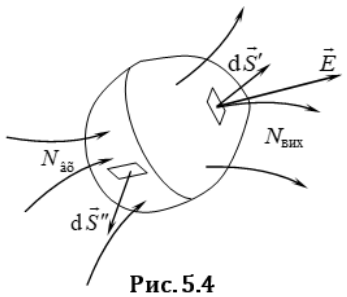


Рис. 5.4

Геометрична інтерпретація потоку. Раніше (п. 1.4) говорилося, що електричне поле можна зображувати за допомогою ліній вектора $\vec{E}$ (силових ліній), і що такі лінії відображають не лише конфігурацію, а й величину поля в різних областях простору. При цьому за умовою (1.9а) $E\,\mathrm{d}S_\perp = \mathrm{d}n$. Отже, відповідно до означень (5.7), (5.7а), виконуються числові рівності:

$$|\mathrm{d}\Phi| = \mathrm{d}N \quad \Rightarrow \quad |\Phi| = N, \tag{5.8}$$

тобто, можна вважати, що

чисельно потік поля крізь задану поверхню дорівнює кількості ліній поля, що її перетинають.

У випадку замкненої поверхні (оболонки) лінії поля в одних точках заходять усередину, а в інших виходять назовні (рис. 5.4). При цьому в першому випадку $\alpha > 90^\circ$ і створюється від'ємний потік, а в другому $\alpha < 90^\circ$ і потік є додатнім. Тому згідно з (3.6) потік крізь замкнену поверхню $\Phi_0 = \Phi_1 + \Phi_2$ можна подати, як

$$\Phi_0 = N_{\text{вих}} - N_{\text{вх}}, \tag{5.8а}$$

де $N_{\text{вих}}$ — кількість ліній поля, що виходять із поверхні назовні, а $N_{\text{вх}}$ — тих, які входять усередину.

Таке геометричне трактування потоку є зручним у багатьох випадках.

5.2.2. *Теорема Гаусса*

Потік є важливою характеристикою поля, котра разом із циркуляцією відображає специфічні риси ("обличчя") кожного силового поля. Теорема Гаусса трактує про властивості потоку електричного поля крізь замкнену поверхню.

Потік поля точкового заряду. Визначимо потік Φ напруженості крізь довільну замкнену поверхню S для поля точкового заряду q у вакуумі (рис. 5.5*а*).

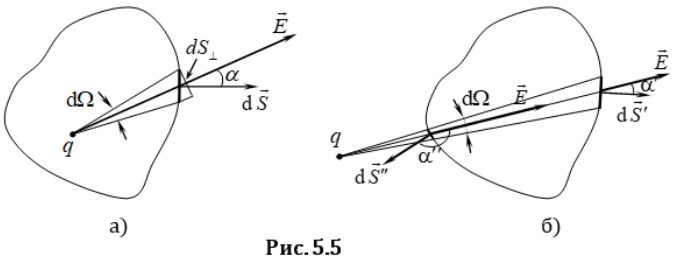


Рис. 5.5

Згідно з виразами (5.7а),

$$\Phi = \oint\limits_S E\,\mathrm{d}S\cos\alpha = \oint\limits_S E\,\mathrm{d}S_\perp,$$

де $\mathrm{d}S$ — площа якоїсь елементарної ділянки замкненої поверхні S.

Проекцію цієї ділянки на площину, перпендикулярну до напрямку поля, можна розглядати як елементарну площадку на поверхні сфери радіуса r і виразити величину $\mathrm{d}S_\perp = \mathrm{d}S\cos\alpha$ через відповідний тілесний кут $\mathrm{d}\Omega$:

$$\mathrm{d}S_\perp = r^2\,\mathrm{d}\Omega. \tag{5.9}$$

Довідка. Плоский кут φ визначає частину простору між двома променями, які починаються в одній точці — вершині кута. Його мірою є довжина дуги s, що вирізається сторонами кута на колі заданого радіуса r із центром у вершині кута. Одиниця плоского кута називається "радіан" (рад) і дорівнює кутові, для якого $(s/r) = 1$. За цим означенням усій площині відповідає плоский кут $\varphi_0 = (2\pi r/r) = 2\pi$ рад. Окрім радіанної, використовують й інші міри, зокрема градусну, де з усією площиною (або довжиною кола) зіставляють кут 360°. Тілесний або просторовий кут Ω — це частина простору, обмежена певною незамкненою конічною поверхнею.

Його мірою є площа сегмента, що вирізається конічною поверхнею на сфері заданого радіуса r із центром у вершині кута. Одиницею тілесного кута є стерадіан (1 ср) — тілесний кут, для якого $(S/r^2) = 1$. Відповідно, всьому тривимірному простору відповідає тілесний кут $\Omega_0 = (4\pi r^2/r^2) = 4\pi$ ср.

З іншого боку, величина E визначається формулою (1.4а), тому

$$\oint_S \vec{E}\,\mathrm{d}\vec{S} = \oint_S \frac{q}{4\pi\varepsilon_0 r^2} r^2\,\mathrm{d}\Omega = \\ = \frac{q}{4\pi\varepsilon_0}\oint_S \mathrm{d}\Omega.$$

Останній інтеграл виражає тілесний кут усього простору і дорівнює 4π ср, отже,

$$\oint_S \vec{E}\,\mathrm{d}\vec{S} = \frac{q}{\varepsilon_0}. \tag{5.10}$$

Тепер розглянемо другий можливий випадок, коли заряд q знаходиться поза поверхнею S (рис. 5.5*б*). В такому разі конічні промені, спрямовані від усіх точок контуру якоїсь площадки $\mathrm{d}S'$ до заряду й утворюють тілесний кут $\mathrm{d}\Omega$, виріжуть на протилежному боці відповідну площадку $\mathrm{d}S''$ з тим самим тілесним кутом. Тому, врахувавши формули (5.9) і (1.4а) та те, що $\cos\alpha < 0$, для потоків крізь указану пару ділянок отримаємо:

$$\mathrm{d}\Phi'' = -\mathrm{d}\Phi' \quad \Rightarrow \quad \mathrm{d}\Phi' + \mathrm{d}\Phi'' = 0.$$

Очевидно, що в такий спосіб можна розчленувати всю замкнену поверхню. Тож сумарний потік крізь замкнену поверхню, що не охоплює заряд,

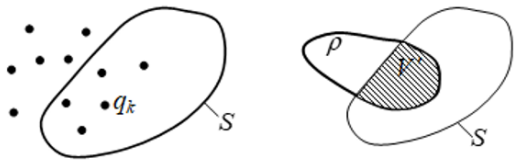
$$\oint_S \vec{E}\,\mathrm{d}\vec{S} = 0. \tag{5.10а}$$

Такого висновку можна дійти і на основі геометричної інтерпретації потоку (5.8а), позаяк у даному випадку кількості ліній поля, що входять усередину поверхні та виходять назовні є однакові.

Відтак, об'єднуючи вирази (5.10) і (5.10а), маємо:

$$\oint_S \vec{E}\mathrm{d}\vec{S} = \begin{cases} \dfrac{q}{\varepsilon_0} & (q \text{ всередині } S); \\ 0 & (q \text{ назовні } S). \end{cases} \tag{5.11}$$

Теорема Гаусса. Отриманий результат легко узагальнити і на поле, створене довільною системою зарядів (рис. 5.6*а*).



а) б)

Рис. 5.6

За принципом суперпозиції напруженість поля системи в будь-якій точці

$$\vec{E} = \sum_i \vec{E}_i.$$

Тому потік результуючого поля крізь довільну замкнену поверхню дорівнює сумі потоків полів кожного із зарядів системи:

$$\oint_S \vec{E}\,\mathrm{d}\vec{S} = \oint_S \left(\sum_i \vec{E}_i\right)\mathrm{d}\vec{S} = \\ = \sum_i \oint_S \vec{E}_i\,\mathrm{d}\vec{S}.$$

Але згідно з (5.11) потік створюють лише заряди, що знаходяться всередині вибраної поверхні. Тому, позначивши такі заряди індексом k, можемо записати:

$$\oint\limits_S \vec{E}\,\mathrm{d}\vec{S} = \frac{1}{\varepsilon_0}\sum_k q_k,$$

або

$$\oint\limits_S \vec{E}\,\mathrm{d}\vec{S} = \frac{q}{\varepsilon_0}, \tag{5.12}$$

де $q = \sum\limits_k q_k$.

Це рівняння й виражає інтегральну теорему Гаусса для електричного поля у вакуумі:

потік напруженості електричного поля у вакуумі крізь довільну замкнену поверхню дорівнює сумарному зарядові, що знаходиться всередині цієї поверхні, діленому на φ_0.

Рівняння (5.12) є чинним і для заряджених тіл, тобто для полів створених неперервним розподілом заряду в заданому об'ємі (рис. 5.5*б*). Тільки й того, що величина q визначається не додаванням зарядів , а інтегруванням заданої густини заряду ρ по об'єму V', який займають заряди всередині поверхні S. Але область інтегрування формально поширюють на весь об'єм V обмежений замкненою поверхнею і записують

$$q = \int\limits_V \rho\,\mathrm{d}V. \tag{5.13}$$

Таким чином,

$$\oint\limits_S \vec{E}\,\mathrm{d}\vec{S} = \frac{1}{\varepsilon_0}\int\limits_V \rho\,\mathrm{d}V. \tag{5.14}$$

Це рівняння є найбільш загальним виразом теореми Гаусса, оскільки інтеграл у правій частині включає як окремі випадки і поверхневий, і лінійний, і дискретний розподіли заряду (див. п. 1.3).

Фізичний зміст теореми Гаусса. Переходячи до аналізу отриманих результатів, одразу звернемо увагу на те, що при незмінній величині заряду всередині замкненої поверхні потік поля крізь неї, тобто значення інтеграла (5.7), не залежить від форми та розмірів поверхні, тим самим і від величини та напрямку напруженості $\vec{E}$ в різних точках поверхні. Це означає, що на практиці

при визначенні потоку крізь задану складну поверхню можна замінити її зручнішою для обчислень простою поверхнею, що охоплює ті самі заряди.

Друга особливість полягає в тому, що потік поля крізь замкнену поверхню визначається не всіма зарядами, що створюють це поле, а тільки тими з них, які опинилися всередині поверхні. Тому, якщо всередині поверхні зарядів немає, або їхня сума дорівнює нулю, то потік $\Phi = 0$. Це може спровокувати на хибний висновок, що й напруженість поля $\vec{E} = 0$. Але відсутність потоку ще не означає відсутності поля. Просто в такій ситуації крізь різні ділянки замкненої поверхні створюються потоки *різного знаку* і такої величини, що сумарний потік виявляється нульовим. Очевидним прикладом є однорідне поле $\vec{E} = \mathrm{const}$, потік якого крізь будь-яку замкнену поверхню дорівнює нулю.

Походження вказаних особливостей теореми Гаусса можна зрозуміти, повернувшись до її виведення. При цьому легко бачити, що вже у викладці, що передує виразу (5.10), щезає змінна r, яка визначається положенням точок на вибраній поверхні, тож залежить від її розмірів і форми. Це стається через закон Кулона, згідно з яким поле точкового заряду $E \sim (1/r^2)$. При будь-якій іншій залежності $E(r)$ потік був би залежним від r. За праву частину рівняння (5.12) чи (5.14) "відповідальним" є принцип суперпозиції, який відображає незалежність електричних полів. Завдяки цьому заряди, розташовані назовні замкненої поверхні, не впливають на поле внутрішніх зарядів, отже, й на потік крізь цю поверхню.

Таким чином, теорема Гаусса органічно об'єднує обидва основні закони теорії електричного поля — закон Кулона і принцип суперпозиції. Тому в класичній теоретичній електродинаміці теорема Гаусса розглядається як один із вихідних постулатів (основних законів), а закон Кулона — як один із його наслідків.

Теорема Ірншоу. Позаяк між зарядами спостерігається і притягання, і відштовхування, то можна так підібрати величину та розташування зарядів у системі, що вона буде перебувати в стані механічної рівноваги. Прикладом може слугувати система з двох однакових зарядів q і розташованого посередині між ними заряду $q' = -q/4$. Прямим обчисленням можна переконатися, що за таких умов рівнодійна сила, котра діє на кожен заряд з боку двох інших, дорівнює нулю, отже всі заряди перебувають в рівновазі. Але ця рівновага є нестійкою, і зміна положення будь-якого із зарядів призведе до її руйнації. Це відображає *теорема Ірншоу*, згідно з якою жодна система частинок, між якими діють тільки кулонівські сили, принципово не може перебувати в стані стійкої рівноваги.

Справедливість цього твердження прямо випливає з теореми Гаусса. Доведемо це від зворотнього, припустивши, що якийсь заряд $-q$ (знак не істотний) системи знаходиться в положенні стійкої рівноваги. Тоді при всякому його зміщенні має з'являтися напрямлена до положення рівноваги сила $\vec{F} = -q\vec{E}$ (рис. 5.7), що створюється полем $\vec{E}$ решти зарядів.

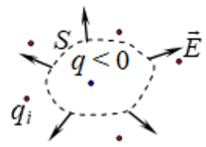


Рис. 5.7

Тому, якщо охопити заряд маленькою замкненою поверхнею так, аби всі інші заряди опинилися поза нею, то вектори $\vec{E}$ на цій поверхні скрізь мають бути напрямлені назовні і створювати крізь неї потік $\Phi > 0$. Але так не може бути, бо це суперечить теоремі Гаусса: всередині S немає жодного із зарядів, які створюють указане поле $\vec{E}$, отже $\Phi = 0$. Таким чином, стійка рівновага зарядів під дією тільки кулонівських сил є неможливою.

5.3. *Обчислення полів за допомогою теореми Гаусса*

Напруженість електричного поля, створеного будь-якою заданою системою зарядів, принципово завжди можна прямо розрахувати на основі закону Кулона і принципу суперпозиції, хоча такий розрахунок може виявитися досить трудомістким. Але для високосиметричних полів задачу можна розв'язати набагато легше за допомогою теореми Гаусса. Далі розглянуто декілька прикладів такого розрахунку для полів, що створюються однорідно (із сталою густиною) розподіленим зарядом.

Приклад 1. Куля

Приклад 2. Сфера

Приклад 3. Циліндр

Приклад 4. Циліндрична поверхня

Приклад 5. Нитка

Приклад 6. Площина

Поле однорідно зарядженої кулі. Поставимо таку задачу: розрахувати напруженість $\vec{E}(\vec{r})$ поля, що створюється у вакуумі зарядом розподіленим із густиною $\rho(\vec{r}) = \text{const}$ по об'єму кулі радіуса R (рис. 5.8).

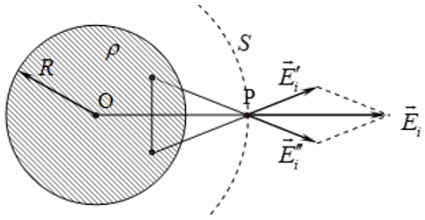


Рис. 5.8

Приступаючи до обчислень, найперше встановимо *симетрію поля* — проаналізуємо, який воно має напрям і від чого залежить його величина в довільній точці P. Це можна зробити за допомогою закону Кулона та принципу суперпозиції наступним чином. Якщо провести через точку P радіальну вісь, то при $\rho = \text{const}$ всю кулю можна поділити на пари однакових малих ділянок із зарядами q_i, розташованих симетрично до цієї осі. Поля цих зарядів $\vec{E}_i'$ і $\vec{E}_i''$ симетрично орієнтовані і однакові за модулем. Тож зрозуміло, що сумарне поле кожної пари зарядів $\vec{E}_i = \vec{E}_i' + \vec{E}_i''$ і поле всієї кулі $\vec{E} = \sum \vec{E}_i$ скрізь мають радіальний напрям і однаковий модуль у рівновіддалених від центра кулі точках.

Таким чином, шукане поле є сферично симетричним (центральним): його лінії напруженості $\vec{E}$ є радіальними променями, які починаються (або закінчуються) в центрі зарядженої кулі. Тому на концентричній сферичній поверхні довільного радіуса r проекція напруженості $E_n = \text{const}$. Через це інтеграл потоку (5.7) крізь таку поверхню легко обчислюється:

$$\Phi = \oint\limits_S E_n \, \mathrm{d}S = E_n \oint\limits_S \mathrm{d}S,$$

або

$$\Phi = E_n S, \tag{5.15}$$

де $S = \oint\limits_S \mathrm{d}S$ — площа всієї поверхні, крізь яку обчислюється потік.

З іншого боку за теоремою Гаусса потік крізь усяку замкнену поверхню визначається рівнянням (5.14). Тому підставивши в нього вираз (5.13), отримаємо:

$$E_n = \frac{q}{\varepsilon_0 S}, \tag{5.16}$$

де q — величина заряду всередині сфери.

Зауважимо, що позаяк вектор $\vec{E}$ напрямлений по нормалі до поверхні S, вона є однією з еквіпотенціальних поверхонь поля (див. п. 3.3). Тому зрозуміло, що формула (5.16) годиться не лише для центрального, а й для всякого іншого симетричного поля, що має еквіпотенціальні поверхні простої форми, на яких $E_n = \text{const}$. Що ж до поля кулі, то у формулах (5.15) і (5.16) $S = 4\pi r^2$, отже

$$E_n = \frac{q}{4\pi\varepsilon_0 r^2}. \tag{5.17}$$

Таким чином, для визначення напруженості поля $E_n(r)$ треба лишень обчислити заряд зосереджений всередині сфери радіуса r. В загальному випадку він визначається інтегралом (5.11), але при $\rho = \text{const}$

$$q = \rho V',$$

де V' — частина об'єму всередині сферичної поверхні, що *заповнена зарядом*.

Об'єм V' залежить від того, охоплює сферична поверхня всю заряджену кулю, чи тільки її частину. Тому для визначення напруженості всередині кулі та поза нею розрахунок величини q треба проводити нарізно.

Поле всередині кулі ($0 \leqslant r \leqslant R$). У цьому випадку весь обмежений сферичною поверхнею об'єм є заповнений зарядом, і

$$q = \rho \frac{4}{3} \pi r^3.$$

Отже, згідно з формулою (5.14), всередині зарядженої кулі напруженість електричного поля

$$E_n = \frac{\rho r}{3\varepsilon_0}, \tag{5.18}$$

або у векторному вигляді

$$\vec{E} = \frac{\rho \vec{r}}{3\varepsilon_0}. \tag{5.18а}$$

Поле назовні кулі ($r > R$). За такої умови всередині сферичної поверхні незалежно від її радіуса r знаходиться весь заряд кулі

$$q = \rho \frac{4}{3} \pi R^3.$$

Відповідно, напруженість електричного поля у зовнішній області простору

$$E_n = \frac{\rho R^3}{3\varepsilon_0 r^2}. \tag{5.19}$$

Аби записати цей вираз у векторній формі, помножимо праву частину на одиничний вектор $(\vec{r}/r)$. Тоді

$$\vec{E} = \frac{\rho R^3}{3\varepsilon_0 r^3}\vec{r}. \qquad (5.19а)$$

Отримані формули (5.18) і (5.19) показують, що заряджена куля поділяє простір на дві області, в кожній з яких електричне поле залежить від відстані r до центра кулі за своїм законом. Графік залежності $E(r)$ показано на рис. 5.9-1.

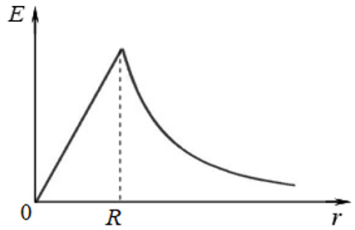


Рис. 5.9-1

Зауважимо також, що коли у формулі (5.18а) виразити величину ρ через заряд всієї кулі, як $\rho = q/(4/3)\pi R^3$, то вийде:

$$\vec{E}(\vec{r}) = \frac{q}{4\pi\varepsilon_0 r^3}\vec{r}. \qquad (5.19б)$$

Отже, поле назовні рівномірно зарядженої кулі збігається з полем точкового заряду такої самої величини, що розміщений у центрі кулі, рис. 5.9*а*.

Так виходить і при неоднорідному, але сферично симетричному розподілі заряду $\rho(r)$.

Поле однорідно зарядженої сфери. Розрахуємо далі поле, що створюється сферичною оболонкою радіуса R, зарядженою з поверхневою густиною заряду $\sigma = \text{const}$. Це поле теж є сферично симетричним, тому його обчислення не відрізняється від попереднього. Відміна лише в тому, що всередині сферичної "гаусової" поверхні при $r < R$ у формулі (5.16) $q = 0$, а при $r \geqslant R$ $q = \sigma \cdot 4\pi r^2$. Тому всередині зарядженої сфери ($r < R$) електричного поля немає:

$$\vec{E} = 0.$$

Відповідно, назовні зарядженої сфери ($r \geqslant R$) напруженість електричного поля дорівнює

$$E(r) = \frac{\sigma R^2}{\varepsilon_0 r^2}, \quad \vec{E}(\vec{r}) = \frac{\sigma R^2}{\varepsilon_0 r^3}\vec{e}, \qquad (5.20)$$

і, як говорилося, збігається з напруженістю поля точкового заряду $q = \sigma \cdot 4\pi r^2$. Графік залежності $E(r)$ показаний на рис. 5.9-2.

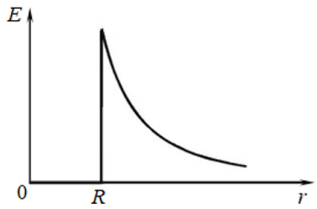


Рис. 5.9-2

Поле однорідно зарядженого циліндра. Тепер знайдемо електричне поле, що створюється зарядом розподіленим із об'ємною густиною $\rho = \text{const}$ по нескінченному циліндру радіуса R (рис. 5.10).

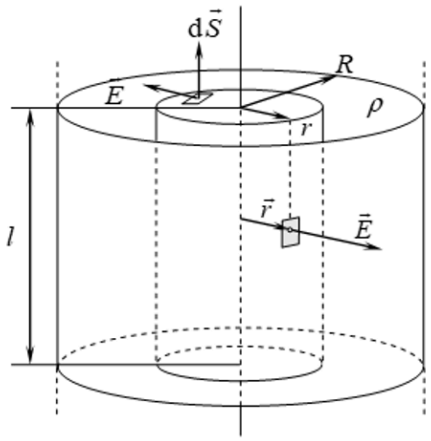


Рис. 5.10

Міркуючи, як раніше, встановлюємо, що шукане має осьову симетрію : в будь-якій точці вектор $\vec{E}$ напрямлений радіально, і його модуль залежить тільки від відстані до осі циліндра. Отже, якщо охопити частину заряду коаксіальною циліндричною поверхнею довільного радіуса r і висоти l, то на його основах $E_n = 0$. Тому потік створюється тільки через бічну поверхню $S_{\text{біч}} = 2\pi rl$, на якій скрізь $E_n = \text{const}$. Відтак

$$\Phi = \oint\limits_S E_n \, \mathrm{d}S = E_n \oint\limits_{S_{\text{біч}}} \mathrm{d}S = E_n \cdot 2\pi rl.$$

Тож за формулою (5.14) для проекції напруженості на радіальний напрям маємо:

$$E_r = \frac{q}{2\pi l \varepsilon_0 r}. \tag{5.21}$$

Далі, як і в попередніх прикладах, обчислення величини q виконуємо нарізно для областей всередині та назовні зарядженого циліндра.

Поле всередині циліндра ($r < R$). У цьому випадку вся замкнена циліндрична поверхня радіуса r і висоти l заповнена зарядом. Тому

$$q = \rho \pi r^2 l,$$

і, згідно з формулою (5.19),

$$E_r = \frac{\rho r}{2\varepsilon_0}, \tag{5.22}$$

а у векторному вигляді

$$\vec{E} = \frac{\rho \vec{r}}{2\varepsilon_0}. \tag{5.22а}$$

Поле назовні циліндра ($r > R$). За такої умови всередині замкненої поверхні незалежно від її радіуса r зосереджений весь заряд охопленої частини циліндра:

$$q = \rho \pi R^2 l.$$

Відповідно, формула (5.19):

$$E_r = \frac{\rho R^2}{2\varepsilon_0 r}. \tag{5.23}$$

Помноживши праву частину цієї формули на одиничний вектор $(\vec{r}/r)$, отримаємо вираз для вектора напруженості:

$$\vec{E} = \frac{\rho R^2}{2\varepsilon r^2}\vec{r}. \tag{5.23а}$$

Порівнюючи вирази (5.22а) і (5.18а) та (5.23а) і (5.19а), бачимо, що залежності $\vec{E}(\vec{r})$ для поля циліндра та кулі є схожими. Відповідно, і графік залежності $E(r)$ для поля циліндра якісно є таким, як на рис. 5.9-1.

Поле однорідно зарядженої циліндричної поверхні. Якщо заряд розподілений із густиною $\sigma = \text{const}$ по нескінченній циліндричній поверхні радіуса R, то, як і у випадку сфери, для точок всередині у формулі (5.21) $q = 0$. Тому всередині зарядженої циліндричної поверхні ($r < 0$) поля немає:

$$\vec{E} = 0;$$

Назовні ($r \geqslant R$) $q = \sigma \cdot 2\pi Rl$, тож

$$E_r = \frac{\sigma R}{\varepsilon_0 r} \tag{5.24}$$

і

$$\vec{E} = \frac{\sigma R}{\varepsilon_0 r^2}\vec{r}. \tag{5.24а}$$

Графік залежності $E(r)$ має вигляд рис. 5.9-2.

Поле однорідно зарядженої нитки. Нитку будемо розглядати як нескінченний циліндр гранично малого радіуса, заряджений з лінійною густиною $\lambda = \text{const}$. Тому напруженість поля визначається формулою (5.21), в якій $q = \lambda l$. Отже,

$$E_r = \frac{\lambda}{2\pi\varepsilon_0 r}. \tag{5.25}$$

У векторній формі

$$\vec{E} = \frac{\lambda}{2\pi\varepsilon_0 r^2}\vec{r}. \tag{5.25а}$$

Графік $E(r)$ наведено на рис. 5.11.

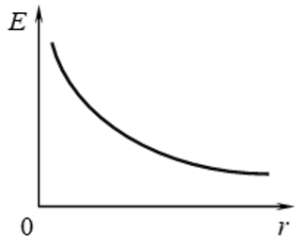


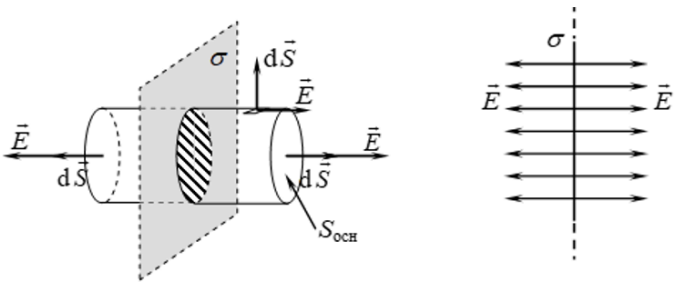
Рис. 5.11

Поле однорідно зарядженої площини. Нарешті визначимо напруженість електричного поля, що створюється нескінченною плоскою поверхнею, зарядженою з поверхневою густиною заряду $\sigma = \text{const}$.

Для встановлення симетрії поля теж можна скористатися законом Кулона та принцип суперпозиції. Але симетрію як цього так і розглянутих полів можна з'ясувати й на основі загальних логічних міркувань, які спираються на однорідність та ізотропність простору. А саме. Уявімо, що вектор $\vec{E}$ має певний напрям, що складає з площиною кут $\alpha \neq 90^\circ$. Але тоді виникає запитання, на яке немає відповіді: а чому в такому разі вектор $\vec{E}$ не може бути спрямований в будь-якому іншому з безлічі напрямків під тим самим кутом α до площини, адже всі вони є *фізично еквівалентними* (рівноправними). У той же час подібне запитання не виникає щодо перпендикулярного ($\alpha = 90^\circ$) до площини напрямку, бо він є один. Отже, у випадку нескінченної однорідно зарядженої площини вектор $\vec{E}$ скрізь напрямлений перпендикулярно до неї в той чи інший бік залежно від знаку заряду. З таких самих міркувань очевидно, що фізично еквівалентними є і всі точки будь-якої площини паралельної до тієї, що створює поле. Тому на такій площині не тільки напрям, а й модуль напруженості E скрізь однаковий, тож $\vec{E} = \text{const}$. Іншими словами, дане поле є *плоско симетричним і однорідним*.

Зауваження. Наведені міркування є придатними для аналізу симетрії будь-якої задачі і мають таку саму доказову силу, як і аналітичні викладки. Але до подібного аналізу слід підходити з уважністю. Наприклад, поле однорідно зарядженої плоскої поверхні у формі диска теж симетричне, але його симетрія не є плоскою. Зокрема, перпендикулярно до площини диска воно напрямлене тільки на осі. Модуль напруженості теж однаковий не скрізь, а тільки на колі з центром на осі й площиною паралельною до площини диска.

Завдяки високій симетрії поле площини теж легко розраховується за допомогою теореми Гаусса. Для цього оточимо частину площини якоюсь замкненою поверхнею, до прикладу, прямим циліндром із паралельними до площини і рівновіддаленими від неї основами, рис. 5.12*а*.



а) б)

Рис. 5.12

У такому випадку вектори $\vec{E}$ на основах циліндра мають однакову величину $E_n = \text{const}$, а на бічній поверхні поле є скрізь паралельним до неї і не створює потоку. За таких умов інтеграл потоку (5.5) крізь увесь замкнений циліндр виражається як

$$\Phi = 2E_n \int\limits_{S_{\text{осн}}} \mathrm{d}S = 2E_n S_{\text{осн}}.$$

При цьому заряд всередині циліндра $q = \sigma S_{\text{осн}}$. Отже, з рівняння (5.10) маємо:

$$E_n = \frac{\sigma}{2\varepsilon_0}. \tag{5.26}$$

У векторному вигляді

$$\vec{E} = \frac{\sigma}{2\varepsilon_0}\vec{n}, \tag{5.26а}$$

де $\vec{n}$ — орт зовнішньої нормалі до площини.

Таким чином, заряджена з густиною $\sigma = \text{const}$ нескінченна площина створює по обидва боки однорідне електричне поле, що напрямлене від площини при $\sigma > 0$ і до неї при $\sigma < 0$. Лінії такого поля являють систему паралельних променів, які розміщені на однаковій відстані один від одного (рис. 5.12*б*).

Контрольні запитання

1. Що називається циркуляцією силового поля $\vec{F}$ по заданому контуру L. Який фізичний зміст має ця величина та від чого залежить її знак для заданого контуру?

2. Яку ознаку поля відображає його циркуляція по довільному контуру? Чому вона дорівнює для електричного поля зарядів?

3. Поміркуйте, в яких одиницях вимірюється та що собою являє потік електричного заряду через поверхню поперечного перерізу провідника, по якому тече електричний струм.

4. Потік електричного поля крізь задану замкнену поверхню дорівнює нулю. Чи означає це, що всередині неї немає поля? Немає зарядів?

5. Чому потік поля точкового заряду q крізь якусь замкнену поверхню дорівнює нулю, коли заряд перебуває поза нею, і не дорівнює нулю, коли він знаходиться всередині?

6. Електричне поле створюється двома паралельними площинами, що різноіменно заряджені з однаковою густиною заряду σ (Кл/м^2). Визначте потік електричного поля крізь якусь замкнену поверхню, що розташована: а) поза площинами; б) між площинами. Що спричинює кожну відповідь?

7. Електричне поле створюється двома паралельними площинами, що заряджені з однаковою густиною заряду σ (Кл/м^2). Визначте потік електричного поля крізь внутрішню та зовнішню частину поверхні невеликої сфери радіуса R із центром на одній з пластин, якщо площини заряджені: а) різнойменно і б) однойменно.

8. Із теореми Гаусса випливає, що всередині будь-якої зарядженої провідної оболонки (замкненої поверхні) електричне поле відсутнє. Як це довести?

II. ЕЛЕКТРИЧНЕ ПОЛЕ В РЕЧОВИНІ

Добре відомо, що середовище істотно впливає на електричне поле. Це пояснюється тим, що всяка речовина є сукупністю великої кількості атомів, що складаються із заряджених частинок – ядер та електронів, які створюють мікроскопічне електричне поле (“мікрополе”) атома $\vec{E}_{\text{мікро}}$. Через загальну електричну нейтральність атома та швидкий рух зарядів, мікрополе має помітну величину лише в околі атома і хаотично змінюється в часі внаслідок теплового руху атомів. Тому при описі поля в речовині використовують усереднене мікроскопічне поле $\vec{E}' = \langle E_{\text{мікро}} \rangle$, яке називають внутрішнім або власним макроскопічним полем (макрополем) речовини. При цьому, аби величина $\vec{E}'$ визначала поле локально (в кожній точці середовища), усереднення проводиться по так званому фізично нескінченно малому об’єму — ділянці, яка за розмірами практично є точковою, але містить макроскопічну кількість атомів. При такому усередненні згладжуються всі нерегулярності та швидкі зміни поля на відстанях порядку міжатомних, але зберігаються плавні зміни макроскопічного поля.

Через хаотичність теплового руху атомів, мікрополя зазвичай є компенсованими, і власне макроскопічне поле в речовині є відсутнім. Але при вміщенні зразка речовини в зовнішнє електричне поле воно спричинює переміщення негативних і позитивних зарядів атомів у взаємно протилежних напрямках, що порушує компенсацію мікрополів. Як наслідок, у речовині виникне власне макроскопічне поле, що разом із зовнішнім утворює результуюче поле з іншою напруженістю.

Описане переміщення зарядів під дією поля відбувається в усіх речовинах, але його особливості і ступінь впливу на поле істотно залежить від внутрішньої будови речовини. З огляду на це всі речовини поділяються на два класи — діелектрики і провідники. За електричними властивостями розрізняють іще один клас речовин — напівпровідники. Але за характером впливу на електричне поле вони є просто недосконалими діелектриками.

Електричні властивості діелектриків і провідників та їхній вплив на електричне поле розглядаються в наступних розділах:

1. Електричне поле в діелектриках
2. Провідники в електричному полі

1. Електричне поле в діелектриках

Вплив діелектриків на характеристики електричного поля розглядається в наступних питаннях:

1.1. Поляризація діелектриків

1.2. Теорема Гаусса для діелектриків. Вектор $\vec{D}$

1.3. Діелектричні сприйнятливість і проникність

1.4. Умови на межі двох діелектриків

Контрольні запитання

1.1. Поляризація діелектриків

Полярні та неполярні молекули. Діелектриками називаються речовини, в молекулах яких усі електрони сильно зв'язані з ядрами так, що навіть у сильних електричних полях вони ледь зміщуються відносно ядер. Через це в діелектрику є неможливим електричний струм, тобто переміщення зарядів по всьому об'єму. З цієї причини суцільна діелектрична оболонка ізолює область, яку оточує, від проникнення в неї зарядів. Тому діелектрики інакше називаються "ізоляторами".

В атомах і молекулах електрони рухаються так швидко, що ядра є немов оточені суцільною хмаринкою від'ємного заряду, або вкриті "електронною оболонкою". При цьому, залежно від конфігурації електронних оболонок, молекули і діелектричні речовини поділяють на неполярні та полярні. У молекулах *неполярних діелектриків*, прикладами яких є гелій He та молекулярний водень $\mathrm{H_2}$, "центри ваги" зарядів електронів і ядер збігаються, як можна зрозуміти із схематичного на рис. 1.1*а*.

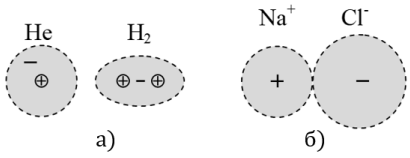
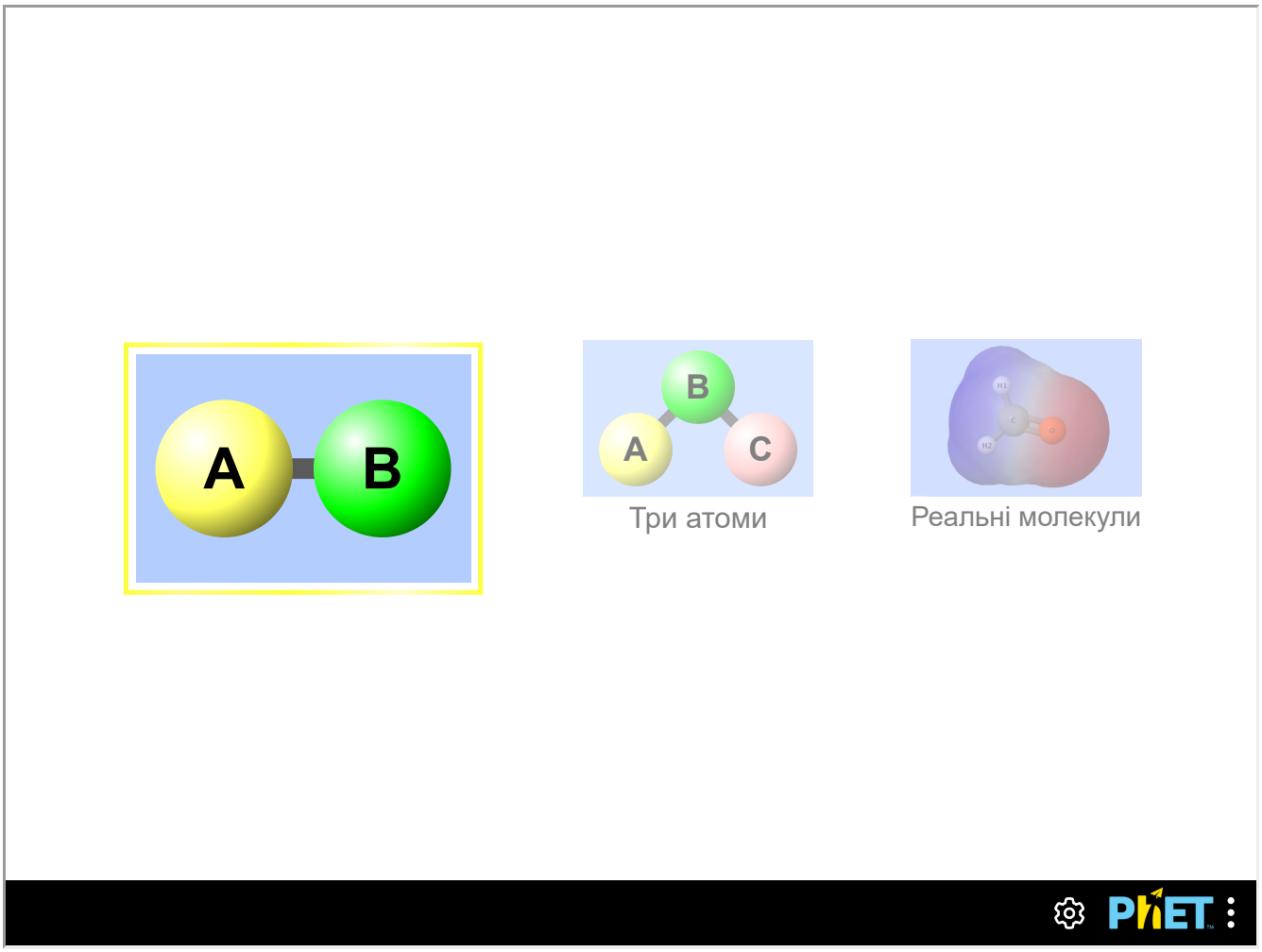


Рис. 1.1

Тому неполярні молекули самі по собі не мають електричного моменту і не створюють власного електричного поля. В молекулах *полярних діелектриків*, таких як NaCl, заряди електронів і ядер рознесені у просторі (рис. 1.1*б*) й утворюють елементарні диполі з відповідним власним електричним моментом. Тому в околі полярної молекули існує власне мікрополе $\vec{E}_{\text{мікро}}$.

Нижче ви можете бачити анімацію, як демонструє поведінку різних молекул в електричному полі.



Поляризація і поляризованість. Унаслідок теплового руху молекул їхні мікрополя в полярному діелектрику орієнтовані хаотично, так що середня величина $\langle \vec{E}_{\text{мікро}} \rangle = 0$. Тому не тільки в неполярних, а й у полярних діелектриках зазвичай ніякого *макроскопічного* поля немає. Але при внесенні діелектрика в зовнішнє електричне поле $\vec{E}_0$ ситуація змінюється внаслідок поляризації.

Сутність цього явища з'ясуємо на прикладі діелектрика з неполярними молекулами, які будемо розглядати як нейтральні кульки. На перший погляд такі молекули через високий ступінь просторової компенсації зарядів узагалі не мають реагувати на зовнішнє електричне поле. Але поле діє не на молекулу як ціле, а на електрони і на ядра окремо. Через це відбувається деформація електронних оболонок — електрони зміщуються відносно ядер і в кожної молекули з'являється так званий *індукований* (або *наведений*) електричний момент $\vec{p}_i$, орієнтований у напрямку електричного поля (рис. 1.2*а*).

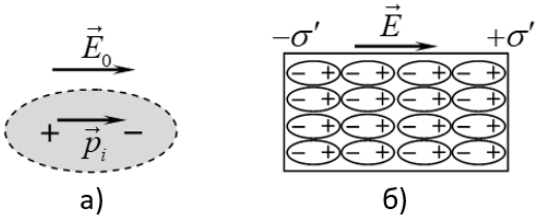


Рис. 1.2

Рознесення негативних і позитивних зарядів частинок діелектрика у взаємно протилежних напрямках під дією електричного поля називається поляризацією.

Примітка. У деяких діелектриках, які називаються п'єзоелектриками, поляризація спостерігається і при механічних деформаціях.

У полярних діелектриках головну роль при поляризації відіграє не деформація електронних оболонок молекул, а переорієнтація їхніх власних електричних моментів під дією поля. Відповідно, поляризацію неполярних діелектриків називають *деформаційною*, а полярних — *орієнтаційною*. Слід відмітити, що власні моменти полярних молекул є значно більшими, ніж індуковані. Але через їхню хаотичну орієнтацію та інтенсивний тепловий рух молекул при орієнтаційній поляризації відбувається лише часткове впорядкування власних моментів, так що в обох цих типах діелектриків здатність до поляризації є приблизно однаковою.

При будь-якому виді поляризації молекули під дією електричного поля утворюють систему елементарних диполів із частково впорядкованими (ефективними) моментами $\vec{p}_i$, фспрямованими по полю, як на рис. 1.2. Тому кожна область діелектрика набуває електричного моменту $\sum \vec{p}_i \neq 0$. Як наслідок порушується компенсація мікроскопічних полів молекул і в діелектрику виникає внутрішнє (власне) макроскопічне поле $\vec{E}'$. Це поле разом із зовнішнім полем $\vec{E}_0$ формує результуюче ***макроскопічне поле в діелектрику***:

$$\vec{E} = \vec{E}_0 + \vec{E}'. \tag{1.1}$$

Зрозуміло, що чим сильніша поляризація, тим більша величина $\sum \vec{p}_i$. Тому її використовують для кількісної характеристики ступеня поляризації діелектрика в кожній точці. Відповідна фізична величина називається поляризованістю або вектором поляризації $\vec{P}$**.** За означенням

$$\vec{P} = \frac{1}{\Delta V} \sum \vec{p}_i, \tag{1.2}$$

де ΔV — фізично нескінченно малий об'єм діелектрика в околі даної точки, $\vec{p}_i$ — електричні моменти окремих молекул у цьому об'ємі.

Примітка. Як говорилося на початку, фізично нескінченно малим об'ємом називається область у речовині, яка може вважатися точковою, але вміщує макроскопічну кількість молекул. Такі вимоги не є суперечливими через гранично малі розміри молекул. Наприклад, у кубику твердої речовини з ребром 10^{-6} м міститься $10^{10} \div 10^{11}$ молекул.

Отже, можна сказати, що поляризованість $\vec{P}$ дорівнює електричному моменту одиниці об'єму поляризованого діелектрика.

Одиницею поляризованості є 1 Кл/м^2, що збігається (і не випадково) з одиницею поверхневої густини заряду.

В однорідному неполярному діелектрику індуковані дипольні моменти всіх молекул однакові, і вираз (1.2) можна записати у вигляді:

$$\vec{P} = n\vec{p}, \tag{1.2а}$$

де n $(1/\text{м}^3)$ — концентрація молекул (кількість в одиниці об'єму), $\vec{p}$ — дипольний момент однієї молекули. Якщо ж вираз (1.2) помножити і поділити на кількість молекул в об'ємі ΔV, то таким робом можна виразити і поляризованість неоднорідного діелектрика:

$$\vec{P} = \frac{N}{\Delta V} \cdot \frac{\sum \vec{p}_i}{N} \quad \Rightarrow \quad \vec{P} = n\langle \vec{p}_i \rangle, \tag{1.2б}$$

де $\langle \vec{p}_i \rangle$ — дипольний момент, який у середньому припадає на одну молекулу в об'ємі ΔV.

Поляризаційні та сторонні заряди. В усіх діелектриках поляризація супроводжується появою макроскопічних поляризаційних зарядів.

Механізм виникнення поляризаційних зарядів в однорідному неполярному діелектрику, вміщеному в однорідне електричне поле, ілюструє рис. 1.2*б*. Унаслідок деформації електронних оболонок молекул електрична нейтральність діелектрика в цілому не порушується, але на його поверхнях з'являються *макроскопічні* заряди протилежного знаку, що розподіляються з деякою густиною σ' і $-\sigma'$ і тому називаються поверхневими поляризаційними зарядами. В той же час всередині різнойменні заряди сусідніх молекул розміщені щільно, і в об'ємі діелектрик лишається електрично нейтральним у кожній точці. У полярних діелектриках поверхневі поляризаційні заряди виникають унаслідок повороту дипольних моментів молекул у напрямку поля.

У неоднорідних діелектриках, а також при вміщенні діелектриків у неоднорідне електричне поле, крім поверхневих, виникають і об'ємні поляризаційні заряди, розподілені з деякою об'ємною густиною ρ'. Це легко зрозуміти на прикладі поляризації умовного неоднорідного діелектрика, в якому концентрація молекул зростає в напрямку зовнішнього поля. В такому разі на поверхню зразка, що орієнтована по полю, вийде більше «плюсів» ніж вийде «мінусів» на протилежну поверхню (рис. 1.3).

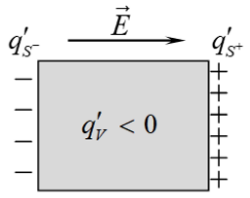


Рис. 1.3

Тож на поверхнях з'явиться сумарний поляризаційний заряд $(q'_{s^+} + q'_{s^-}) > 0$. Тому, враховуючи загальну нейтральність зразка, в об'єм теж стане заряджеником:

$$q'_{s^+} + q'_{s^-} + q'_V = 0 \quad \Rightarrow \quad q'_V = -(q'_{s^+} + q'_{s^-}) < 0.$$

Цей об'ємний поляризаційний заряд, розподіляється з відповідною об'ємною густиною $\rho' < 0$. Зрозуміло, що при зміні напрямку поля на протилежний зміниться і знак ρ'.

При поляризації електрони в молекулах діелектрика зміщуються тільки на дуже малу відстань, лишаючись зв'язаними із своїми ядрами. Тому поляризаційні заряди не можна відокремити один від одного і перемістити в інше місце. Приміром, якщо зразок поляризованого діелектрика розрізати навпіл, то обидві частини лишаться електрично нейтральними і поляризованими так само, як вихідний зразок. Тому поляризаційні заряди інакше називають зв'язаними зарядами.

Принагідно зазначимо, що в теорії електричного поля в діелектриках оперують також поняттям вільних зарядів. Так називають заряди, що, в принципі, можуть довільно розташовуватися і переміщуватися. Вони не належать атомам середовища, тому ще називаються сторонніми зарядами. Такі заряди створюють зовнішнє поле $\vec{E}_0$, яке є першопричиною поляризації. Прикладом сторонніх зарядів можуть бути заряди заряджених кульок, вміщених у діелектричну рідину, або протилежні за знаком заряди на пластинах зарядженого конденсатора, між якими знаходиться діелектрик.

Зв'язок між поляризованістю та поляризаційними зарядами. Із сказаного вище зрозуміло, що величина і розподіл поляризаційних зарядів у діелектрику залежать від ступеня та характеру поляризації.

Установимо спочатку зв'язок між поляризованістю $\vec{P}$ і густиною поверхневих поляризаційних зарядів σ'. Для цього в шарі однорідно поляризованого діелектрика ($\vec{P} = \text{const}$) товщиною h виділимо ділянку у вигляді косого циліндра із твірною l та основами ΔS, як показано на рис. 1.4.

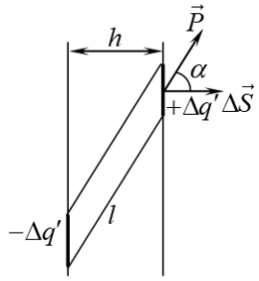


Рис. 1.4

На основах циліндра зосереджені різнойменні поляризаційні заряди однакової величини $q = \sigma' \Delta S$, отже, він є диполем із плечем $l = h/\cos\alpha$ і величиною електричного моменту

$$p = \sigma' \Delta S l = \sigma' \Delta S \frac{h}{\cos\alpha}.$$

Поділивши цю величину на об'єм циліндра $\Delta V = \Delta S h = \Delta S l \cos\alpha$ згідно з означенням (1.2), отримаємо для поляризованості:

$$P = \frac{\sigma'}{\cos\alpha} \quad \Rightarrow \quad P\cos\alpha = \sigma',$$

або

$$P_n = \sigma', \tag{1.3}$$

де $P_n = P\cos\alpha$ — нормальна до площадки ΔS складова поляризованості.

Такий самий результат дістанемо і при неоднорідній поляризації, взявши такі малі значення h і ΔS, аби в межах ΔV неоднорідність була не істотною. Тому формула (1.3) є загальною й означає, що на поверхні діелектрика нормальна складова поляризованості в будь-якій точці дорівнює поверхневій густині зв'язаних зарядів у цій точці.

Зокрема, в прямому циліндрі з діелектрика, що однорідно поляризований уздовж осі, поверхнева густина зв'язаних зарядів на основах $\sigma' = P$, а повна величина заряду $q'_s = PS$ (S — площа основи). На бічній поверхні $P_n = 0$, отже, $\sigma' = 0$ і $q'_S = 0$.

У загальному випадку зв'язаний заряд на якійсь поверхні S (рис. 1.5*а*) визначається інтегруванням зарядів $\mathrm{d}q' = \sigma'\,\mathrm{d}S$ на всіх її елементарних ділянках:

$$q'_S = \int\limits_S P_n \, \mathrm{d}S = \int\limits_S P \, \mathrm{d}S \cos\alpha = \int\limits_S \vec{P} \, \mathrm{d}\vec{S}.$$

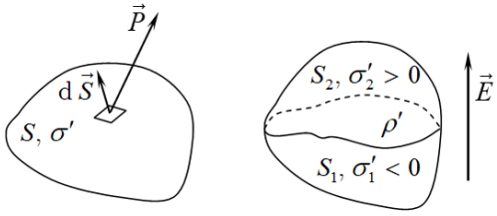


Рис. 1.5

Іншими словами,

поляризаційний заряд на поверхні діелектрика дорівнює потокові вектора $\vec{P}$ крізь цю поверхню.

За допомогою виразу (1.3) можна знайти зв'язок і між поляризованістю та об'ємними поляризаційними зарядами. Для цього розглянемо потік вектора $\vec{P}$ крізь довільну замкнену поверхню S (рис. 1.5*б*). Подумки поділимо її на дві частини S_1 і S_2. Тоді потік крізь усю поверхню

$$\oint\limits_S \vec{P} \, \mathrm{d}\vec{S} = \int\limits_{S_1} \vec{P} \, \mathrm{d}\vec{S} + \int\limits_{S_2} \vec{P} \, \mathrm{d}\vec{S},$$

і, враховуючи (1.3),

$$\oint\limits_S \vec{P} \, \mathrm{d}\vec{S} = \int\limits_{S_1} \sigma'_1 \, \mathrm{d}S + \int\limits_{S_2} \sigma'_2 \, \mathrm{d}S = q'_{S1} + q'_{S2} = q'_S, \tag{1.4}$$

де $q'_S = q'_{S1} + q'_{S2}$ — поляризаційний заряд на всій замкненій поверхні S.

При однорідній поляризації q'_{S1} і q'_{S2} відрізняються тільки знаком, так що $q'_S = 0$. Але в неоднорідно поляризованому діелектрику $|q'_{S1}| \neq q'_{S2}$, тож при його загальній електричній нейтральності зразка заряд на всій поверхні $q'_S \neq 0$. Це означає, що при неоднорідній поляризації в об'ємі теж виникає поляризаційний заряд $q'_V = -q'_S$. Відтак вираз (1.4) можна подати у вигляді:

$$\oint\limits_S \vec{P} \, \mathrm{d}\vec{S} = -q'_V. \tag{1.5}$$

або

$$\oint\limits_S \vec{P} \, \mathrm{d}\vec{S} = -\int\limits_V \rho' \, \mathrm{d}V, \tag{1.5а}$$

де ρ' — об'ємна густина поляризаційного заряду.

Рівняння (1.5), (1.5а) встановлюють загальний зв'язок між поляризованістю та поляризаційними зарядами в діелектрику і по суті виражають теорему Гаусса для вектора $\vec{P}$:

потік вектора поляризованості крізь довільну замкнену поверхню в діелектрику дорівнює величині поляризаційного заряду всередині цієї поверхні взятому з протилежним знаком.

1.2. Теорема Гаусса для діелектриків. Вектор $\vec{D}$

Як говорилося на початку, поле в діелектрику складається з поля сторонніх зарядів $\vec{E}_0$ і внутрішнього макроскопічного поля $\vec{E}'$:

$$\vec{E} = \vec{E}_0 + \vec{E}'.$$

При цьому внутрішнє макроскопічне поле $\vec{E}'$ є усередненим мікроскопічним полем всіх молекул поляризованого діелектрика. Але речовина діелектрика скрізь, окрім місць розташування поляризаційного заряду, лишається електрично нейтральною і як така участі у створенні поля $\vec{E}'$ не бере. Тому при розгляді

поля в діелектрику можна вважати, що поляризаційні заряди, як і сторонні, розташовані у вакуумі. Відтак на основі теореми Гаусса ((5.14), розділ I) для результуючого поля $\vec{E}$ можна записати:

$$\oint\limits_S \vec{E}\,\mathrm{d}\vec{S} = \frac{1}{\varepsilon_0}\int\limits_V (\rho + \rho')\,\mathrm{d}V, \tag{1.6}$$

де ρ і ρ' — об'ємна густина сторонніх і поляризаційних зарядів, відповідно.

Рівняння (1.6) є подібним до ((5.14), розділ I) але трактувати його як вираз теореми Гаусса для поля в діелектриках не можна, бо воно містить невідомий параметр ρ', який не визначає поле $\vec{E}$, а сам залежить від нього. Тому з огляду на теорію рівняння (1.6) не виражає фундаментальних (загальних) властивостей поля в діелектриках, і не може трактуватись як самодостатній фізичний закон. Воно лише відображає той тривіальний факт, що у створенні електричного поля в діелектрику окрім сторонніх зарядів беруть участь і заряди самого діелектрика. Практичної цінності рівняння (1.6) теж не має, бо шукати з нього невідоме поле $\vec{E}$ можна було би тільки знаючи розподіл поляризаційних зарядів ρ', який визначається тим таки невідомим полем $\vec{E}$.

Але вихід з цього хибного логічного кола існує, позаяк поляризаційні заряди є пов'язані з поляризованістю. Розкривши у виразі (1.6) дужки та зробивши заміну (1.5а) матимаємо:

$$\oint\limits_S \vec{E}\,\mathrm{d}\vec{S} = \int\limits_V \rho\,\mathrm{d}V - \oint\limits_S \vec{P}\,\mathrm{d}\vec{S} \quad \Rightarrow \quad \oint\limits_S (\varepsilon_0\vec{E} + \vec{P})\,\mathrm{d}\vec{S} = \int\limits_V \rho\,\mathrm{d}V. \tag{1.7}$$

Величина в дужках, як і напруженість $\vec{E}$, є визначеною в усіх точках простору і утворює відповідне векторне поле. Тому вона позначається одним символом $\vec{D}$ і називається "вектором $\vec{D}$" або "вектором електричного зміщення". Таким чином, за означенням

вектором електричного зміщення називається вектор

$$\vec{D} = \varepsilon_0\vec{E} + \vec{P}. \tag{1.8}$$

Відтак рівняння (1.7) набуває вигляду:

$$\oint\limits_S \vec{D}\,\mathrm{d}\vec{S} = \int\limits_V \rho\,\mathrm{d}V \tag{1.9}$$

або

$$\oint\limits_S \vec{D}\,\mathrm{d}\vec{S} = q \tag{1.9а}$$

і виражає теорему Гаусса для електричного поля в діелектриках:

потік вектора $\vec{D}$ крізь довільну замкнену поверхню дорівнює сумарній величині вільного заряду всередині цієї поверхні.

Рівняння (1.9), на відміну від (1.7), не містить невідомих параметрів і є чинним як для всіх діелектриків, так і для вакууму. Справді, у вакуумі через відсутності речовини $\rho' = 0$, тож $\vec{D} = \varepsilon_0\vec{E}$, і рівняння (1.9) трансформується в рівняння (5.14) теореми Гаусса для поля у вакуумі:

$$\oint\limits_S \varepsilon_0\vec{E}\,\mathrm{d}\vec{S} = \int\limits_V \rho\,\mathrm{d}V \quad \Rightarrow \quad \oint\limits_S \vec{E}\,\mathrm{d}\vec{S} = \frac{1}{\varepsilon_0}\int\limits_V \rho\,\mathrm{d}V.$$

Теорема Гаусса виконується і в нестаціонарних полях при не надто швидкій їхній зміні з часом. Тому рівняння (1.9) є одним з основних інтегральних рівнянь класичної теорії електромагнітного поля, створеної Максвеллом. Разом із тим слід усвідомити, що вектор $\vec{D}$ є синтетичним утворенням і не має прямого фізичного змісту, як, скажімо, вектор $\vec{E}$. Справді, згідно з означенням (1.8), вектор $\vec{D}$ є комбінацією двох різних за змістом векторів: вектора $\varepsilon_0\vec{E}$, що визначається станом поля, та вектора $\vec{P}$, який відображає стан речовини. Тому

вектор $\vec{D}$ не є польовим вектором,

тобто кількісною характеристикою електричного поля.

1.3. Діелектричні сприйнятливість і проникність

Уведення вектора $\vec{D}$ є доцільним і в практичному сенсі, бо часто полегшує обчислення електричного поля в діелектриках. Це зумовлено тим, що в деяких випадках поле вектора $\vec{D}$ можна розрахувати, знаючи розподіл тільки сторонніх зарядів. З іншого боку, за поляризацію відповідає повне поле в діелектрику, тож поляризованість $\vec{P}$ є функцією напруженості $\vec{E}$: $\vec{P} = \vec{P}(\vec{E})$. Тому при відомому зміщенні, вираз (1.8) із означення перетворюється на функціональне рівняння

$$\vec{D}(\vec{r}) = \varepsilon_0 \vec{E}(\vec{r}) + \vec{P}(\vec{E}(\vec{r})) \tag{1.10}$$

з якого можна визначити напруженість поля, створеного в діелектрику заданим розподілом сторонніх зарядів $\rho(\vec{r})$.

Такий розрахунок включає два етапи:

— визначення допоміжного вектора $\vec{D}(\vec{r})$ через заданий розподіл сторонніх зарядів $\rho(\vec{r})$;

— визначення основного вектора поля $\vec{E}(\vec{r})$ за допомогою рівняння (1.10).

У загальному випадку це може бути не простою задачею через складну залежність поляризованості від напруженості поля $\vec{P} = \vec{P}(\vec{E})$. Але ізоляційні матеріали, що використовуються в електротехніці та електроніці, є ізотропними діелектриками, і залежність $\vec{P}(\vec{E})$ в них має простий вигляд:

$$\vec{P} = \varepsilon_0 \varkappa \vec{E}, \tag{1.11}$$

де безрозмірне число $\varkappa$ називається діелектричною сприйнятливістю діелектрика. Ця величина показує, наскільки сильно, чи слабко поляризується діелектрик в полі заданої напруженості, тобто дійсно характеризує його сприйнятливість до впливу електричного поля.

Підставивши співвідношення (1.11) у рівняння (1.10), отримаємо:

$$\vec{D} = \varepsilon_0 (1 + \varkappa) \vec{E}. \tag{1.12}$$

У цьому виразі число в дужках має той самий, що й $\varkappa$, загальний зміст. Тому його позначають одним символом ε і називають діелектричною проникністю речовини. Отже, за означенням

$$\varepsilon = 1 + \varkappa. \tag{1.13}$$

Діелектрична проникність є табличною характеристикою речовини і для звичайних діелектриків складає декілька одиниць за винятком води, для якої $\varepsilon = 81$. Але існують також тверді діелектрики, що звуться сегнетоелектриками, і мають аномально високу (до $\sim 10^4$) проникність. Це зумовлено особливостями їхньої внутрішньої будови.

Із виразів (1.12) і (1.13) випливає важлива формула, що напряму пов'язує вектори $\vec{D}$ і $\vec{E}$ в ізотропному діелектрику:

$$\vec{D} = \varepsilon_0 \varepsilon \vec{E}. \tag{1.14}$$

1.4. Умови на межі двох діелектриків

Співвідношення (1.14) пов'язує напруженість $\vec{E}$ поля в діелектрику з допоміжною величиною $\vec{D}$. Але реально напруженість поля в заданому діелектрику визначається полем сторонніх зарядів $\vec{E}_0$, в яке він уміщений і яке є першопричиною його поляризації. Аби встановити зв'язок між векторами $\vec{E}$ і $\vec{E}_0$, для початку розглянемо поле у плоскопаралельній пластині однорідного ізотропного діелектрика з проникністю ε, яка вміщена в однорідне зовнішнє поле $\vec{E}_0 = \text{const}$, що перпендикулярне до її поверхонь, рис. 1.6*а*.

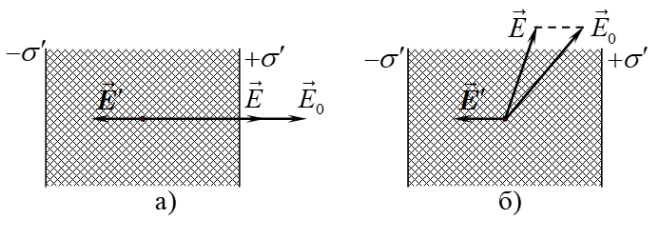


Рис. 1.6

За таких умов внутрішнє поле $\vec{E}'$ можна розглядати як поле двох заряджених з густиною заряду σ' і $-\sigma'$ паралельних площин (поверхонь діелектрика). Воно напрямлене протилежно до зовнішнього поля $\vec{E}_0$ і, згідно з принципом суперпозиції і формулою ((5.23), розділ I), має величину

$$E' = \frac{\sigma'}{\varepsilon_0}.$$

Тому модуль напруженості результуючого поля в діелектрику, відповідно до (1.1), дорівнює

$$E = E_0 - \frac{\sigma'}{\varepsilon_0}.$$

Звідси, врахувавши співвідношення (1.3) і (1.11), знаходимо зв'язок між полями:

$$E = E_0 - \varkappa E \quad \Rightarrow \quad (1+\varkappa)E = E_0 \quad \Rightarrow \quad \varepsilon E = E_0 \Rightarrow \vec{E} = \frac{\vec{E}_0}{\varepsilon}. \tag{1.15}$$

Отже, пластина послаблює поле в ε разів без зміни його напрямку. Тому може здатися, що проникність зразка діелектрика ε — то є просто коефіцієнт послаблення ним електричного поля. Але насправді діелектрична проникність не має такого прозорого змісту, а простота виразу (1.15) зумовлена не властивостями речовини діелектрика, а формою та орієнтацією його поверхонь. Це стає зрозумілим, якщо розглянути ту саму пластину вміщену в електричне поле, що спрямоване навкіс, як на рис. 1.6*б*. Очевидно, що тепер вектори $\vec{E}_0$, $\vec{E}'$ і $\vec{E}$ вже не є колінеарними, тож і $\vec{E} \neq (\vec{E}_0/\varepsilon)$. Тому формула (1.15) без застережень виконується лише в абстрактному ізотропному діелектрику, що не має поверхонь, тобто займає увесь безмежний простір. Що ж до обмеженого діелектрика, що має певні розміри та поверхню, співвідношення (1.15) виконується тільки у відповідних випадках, які визначаються так званими умовами на межі — співвідношеннями між векторами $\vec{D}_1$ і $\vec{D}_2$ та $\vec{E}_1$ і $\vec{E}_2$ на поверхні поділу двох діелектриків.

Визначимо указані співвідношення для двох ізотропних діелектриків з проникностями ε_1 і ε_2 у випадку, коли на межі поділу немає сторонніх зарядів. Для цього спочатку розглянемо потік вектора $\vec{D}$ крізь поверхню невеликого замкненого циліндра з паралельними до межі поділу основами ΔS і висотою h, рис. 1.7*а*.

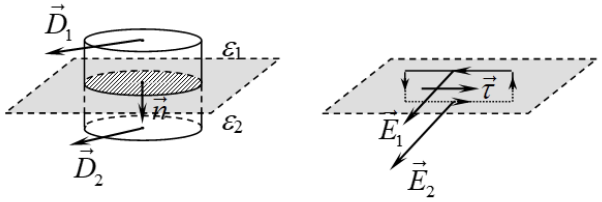


Рис. 1.7

Цей потік складається з потоків через основи циліндра $\Delta\Phi_1, \Delta\Phi_2$ і бічну поверхню $\Delta\Phi'$ і, згідно з теоремою Гаусса (1.9), за відсутності сторонніх зарядів дорівнює нулю:

$$\Delta\Phi_1 + \Delta\Phi_2 + \Delta\Phi' = 0.$$

Позаяк циліндр малий, потоки через основи можна наближено записати як $\Delta\Phi_1 \approx -D_{1n}\Delta S$ і $\Delta\Phi_2 \approx D_{2n}\Delta S$, спроектувавши вектори на спільну нормаль $\vec{n}$ до межі поділу. Тоді виходить:

$$-D_{1n}\Delta S + D_{2n}\Delta S + \Delta\Phi' \approx 0.$$

Неточність цієї рівності зумовлена тим, що в кожному з діелектриків у різних точках площинки ΔS величина D_n не однакова. Тому для отримання точного виразу треба перейти до границі, "стягнувши" циліндр у якусь точку на межі поділу. Задля більшої прозорості зробимо це в два етапи. Спочатку будемо подумки зменшувати висоту циліндра, поступово наближаючи основи до межі поділу. Тоді у границі $h \to 0$ зникне потік крізь бічну поверхню:

$$-D_{1n}\Delta S + D_{2n}\Delta S \approx 0 \quad \Rightarrow \quad D_{1n}\Delta S \approx D_{2n}\Delta S.$$

А тепер будемо зменшувати основи, зменшуючи й похибку в отриманій рівності, аж поки в границі $\Delta S \to 0$ вона не стане математично точною:

$$D_{1n} = D_{2n}. \tag{1.16а}$$

Відтак за допомогою співвідношення (1.14) встановлюємо зв'язок між нормальними складовими для векторів напруженості поля:

$$\varepsilon_1 E_{1n} = \varepsilon_2 E_{2n} \quad \Rightarrow \quad \frac{E_{1n}}{E_{2n}} = \frac{\varepsilon_2}{\varepsilon_1}. \tag{1.16б}$$

Далі за допомогою теореми про циркуляцію ((5.2), розділ I) і формули (1.14) встановимо межові умови для тангенціальних складових векторів. Для цього розглянемо циркуляцію C вектора $\vec{E}$ по малому прямокутнику, що охоплює деяку точку на межі поділу діелектриків і орієнтований так, як показано на рис. 1.7*б*. Ураховуючи малість контуру, інтеграл у рівнянні ((5.2), розділ I) наближено виразимо через вклади окремих сторін:

$$-E_{1\tau} + E_{2\tau} + \Delta C \approx 0.$$

де ΔC — спільний вклад від поперечних сторін, а проекції векторів поля на напрям переміщення вздовж поздовжніх сторін записані через спільний орт $\vec{\tau}$ дотичної до межі поділу.

Далі, стягаючи спочатку поперечні, а потім поздовжні сторони прямокутника в вибрану точку на межі поділу діелектриків, отримаємо в границі точні співвідношення для напруженості

$$E_{1\tau} = E_{2\tau} \tag{1.17а}$$

і, за допомогою формули (1.14), — для електричного зміщення.

$$\frac{D_{1\tau}}{\varepsilon_1} = \frac{D_{2\tau}}{\varepsilon_2}. \tag{1.17б}$$

Зведемо всі отримані результати, які й складають умови на межі двох діелектриків:

$$\begin{aligned} D_{1n} &= D_{2n}, & \frac{E_{1n}}{E_{2n}} &= \frac{\varepsilon_2}{\varepsilon_1}; \\ \frac{D_{1\tau}}{D_{2\tau}} &= \frac{\varepsilon_1}{\varepsilon_2}, & E_{1\tau} &= E_{2\tau}. \end{aligned} \tag{1.18}$$

З умов (1.18) випливає, що при переході через межу поділу діелектриків змінюються не тільки модулі, а й напрямки векторів $\vec{D}$ і $\vec{E}$. При цьому напрямки векторів змінюються однаково, а модулі — ні. Це наочно ілюструє рис. 1.8: коли $\varepsilon_2 < \varepsilon_1$ (рис. 1.8*а*)), то $D_2 < D_1$, а $E_2 > E_1$, тоді як при $\varepsilon_2 > \varepsilon_1$ (рис. 1.8*б*)) навпаки $D_2 > D_1$, а $E_2 < E_1$.

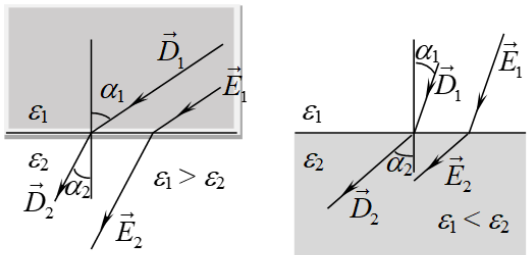


Рис. 1.8

Для межі поділу вакууму ($\varepsilon_1 = 1$) і діелектрика ($\varepsilon_2 = \varepsilon$) умови (1.18) мають вигляд:

$$\begin{aligned} D_n &= D_{0n}, & \frac{D_\tau}{D_{0\tau}} &= \varepsilon; \\ \frac{E_n}{E_{0n}} &= \frac{1}{\varepsilon}, & E_\tau &= E_{0\tau}. \end{aligned} \tag{1.18а}$$

Звідси випливає, що $E = \sqrt{(E_{0n}/\varepsilon)^2 + E_{0\tau}^2}$. Отже, на загал $E \neq (E_0/E_0\varepsilon)$, як і говорилося на початку. Виняток становить тільки випадок, коли поле напрямлене по нормалі до поверхні діелектрика. Тоді $E_\tau = E_{0\tau} = 0$, і поверхня діелектрика співпадає з однією з еквіпотенціальних поверхонь поля (див. розділ I, п. 2). У такому разі нормальні проекції векторів напруженості збігаються з модулями, і поле в діелектрику

$$E = \frac{E_0}{\varepsilon} \tag{1.19}$$

так, немов він є безмежним. Відповідно, для двох діелектриків, межа поділу яких є еквіпотенціальною,

$$\frac{E_1}{E_2} = \frac{\varepsilon_2}{\varepsilon_1}. \tag{1.20}$$

У загальному випадку при переході з одного діелектрика в інший змінюється і напрямок поля. Через це на межі поділу спостерігається ефект заломлення ліній векторів $\vec{D}$ і $\vec{E}$, рис. 1.9

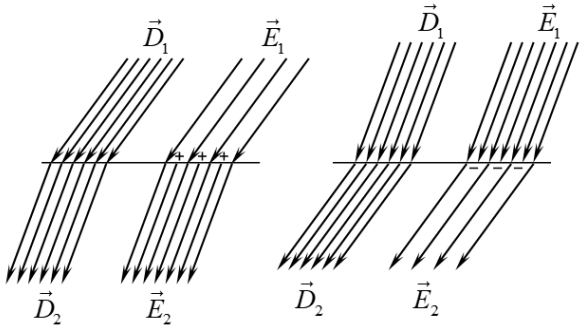


Рис. 1.9

При цьому з умов (1.18) випливає наступний “закон заломлення”:

$$\frac{\operatorname{tg}\alpha_1}{\operatorname{tg}\alpha_2} = \frac{\varepsilon_1}{\varepsilon_2}. \tag{1.21}$$

Прийнявши $\varepsilon_1 = 1$ і $\varepsilon_2 = \varepsilon$ та відповідно позначивши кути як α_0 і α, отримаємо зв’язок між напрямками векторів на межі вакууму і діелектрика з проникністю ε:

$$\frac{\operatorname{tg}\alpha_0}{\operatorname{tg}\alpha} = \frac{1}{\varepsilon}. \tag{1.21а}$$

Зрозуміло, що лінії векторів $\vec{D}$ і $\vec{E}$ заломлюються однаково, але в їхній поведінці є й істотна відмінність. Якщо розглянути потоки цих векторів крізь замкнену поверхнею у формі диска з основами ΔS, що прилягають до поверхні поділу діелектриків, як на рис. 1.7, то згідно з (1.18) отримаємо: $D_{1n}\Delta S = D_{2n}\Delta S$. З урахуванням геометричного змісту потоку це означає, що при переході через межу поділу кількість ліній вектора $\vec{D}$ не змінюється, тобто вони є неперервними. Натомість для напруженості $E_{1n}\Delta S \neq E_{2n}\Delta S$, і лінії вектора $\vec{E}$ терплять розрив, як показано на рис. 1.9.

Це не важко пояснити. Через різну проникність діелектриків їхня поляризованість не однакова, і поляризаційні заряди протилежного знаку, що виникають на прилеглих поверхнях мають різну величину. Через це на поверхні дотику існують некомпенсовані зв’язані заряди того, чи іншого знаку. Якщо $\varepsilon_2 > \varepsilon_1$, то ці заряди від’ємні, і на них певна кількість ліній $\vec{E}$ закінчується, як на рис. 1.9*а*. При $\varepsilon_1 > \varepsilon_2$ указані заряди мають знак «$+$», і на межі поділу виникають додаткові лінії напруженості, рис. 1.9*б*.

Наостанку вкажемо наступний корисний для практики висновок, що випливає з проведеного розгляду. Позаяк у безмежному діелектричному середовищі та в діелектрику, обмеженому еквіпотенціальними поверхнями, електричне поле порівняно з вакуумом просто ослаблюється в ε разів без зміни напрямку, то будь-яку необхідну формулу можна отримати з відповідної формули для вакууму заміною ε_0 на $\varepsilon_0\varepsilon$. До прикладу, при наявності діелектрика замість формул (1.4) чи (5.26) з розділу I маємо:

$$E = \frac{q}{4\pi\varepsilon\varepsilon_0 r^2}, \tag{1.22}$$

та

$$E = \frac{\sigma}{\varepsilon_0\varepsilon}. \tag{1.23}$$

Контрольні запитання

1. Які речовини називаються діелектриками та як вони впливають на електричне поле?

2. Що таке внутрішнє (власне) макроскопічне поле в діелектрику? Як воно виникає та від чого залежить?

3. Що таке поляризація та поляризованість діелектрика?

4. Як утворюються поляризаційні заряди. Яка відміна існує між ними та сторонніми зарядами?

5. Який зв’язок існує між поляризаційними зарядами та поляризованістю діелектрика?

6. Як формулюється теорема Гаусса для електричного поля в діелектрику? Чи можна вважати, що вектор $\vec{D}$ визначається тільки розподілом сторонніх зарядів?

7. Чи можна вважати $\vec{D}$ вектором електричного поля в діелектрику?

8. Який зв'язок існує між поляризованістю та напруженістю поля в ізотропних діелектриках?

9. Установіть зв'язок між вектором $\vec{D}$ та напруженістю поля $\vec{E}$ в ізотропному діелектрику.

10. Чи можна трактувати діелектричну проникність ε як коефіцієнт послаблення напруженості поля в ізотропному діелектрику?

11. Запишіть умови на межі двох ізотропних діелектриків за відсутності на межі сторонніх зарядів. У якому випадку поле в діелектрику просто послаблюється в ε разів порівняно з вакуумом?

12. У якому випадку при переході з одного діелектрика в інший лінії векторів $\vec{D}$ і $\vec{E}$ відхиляються до межі поділу діелектриків, а в якому від неї? Покажіть картину ліній цих векторів у випадку однорідного поля.

13. Поясніть, як змінюється щільність ліній векторів $\vec{E}$ і $\vec{D}$ при переході з одного діелектрика в інший, якщо вектори напрямлені а) по нормалі і б) під кутом до межі поділу діелектриків.

2. Провідники в електричному полі

Провідниками називаються речовини, що мають у своєму складі багато вільних заряджених частинок, які здатні рухатися по всьому об'єму зразка. Через це вони за впливом на електричне поле дуже відрізняються від діелектриків. Далі розглядаються властивості та поведінка провідників у статичних електричних полях. А саме:

2.1. Електростатичні властивості провідників

2.2. Електрична ємність. Конденсатори

2.3. Енергія електричного поля

Контрольні запитання

2.1. *Електростатичні властивості провідників*

Електростатична індукція. Класичними провідниками є метали — речовини, в атомах яких валентні електрони є слабко зв'язані з ядром. Тому при об'єднанні атомів у твердий метал валентні електрони взагалі вивільняються від "своїх" ядер і поводяться подібно до частинок газу в посудині, утворюючи "електронний газ". Але при внесенні металу в якесь електричне поле $\vec{E}_0$ вільні електрони, окрім хаотичного теплового руху, починають зміщуватися проти напрямку поля під дією сили $\vec{F} = -e\vec{E}_0$. Як наслідок, на поверхні зразка з одного боку з'являється не компенсований негативний заряд $-q_i$, а з протилежного — такий самий позитивний заряд q_i, рис. 2.1.

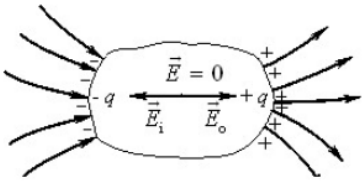


Рис. 2.1

Виникнення макроскопічних зарядів на поверхні незарядженого провідника під дією електричного поля називається електростатичною індукцією, а самі заряди — індукованими зарядами.

Індуковані заряди створюють у провіднику власне поле $\vec{E}_i$, яке напрямлене протилежно до зовнішнього поля $\vec{E}_0$, що нагадує ситуацію в діелектриках. Але, коли в діелектрику на поверхню здатні виходити лише електрони молекул поверхневого шару, то в металі це можуть робити всі вільні електрони, кількість яких є сумірною із кількістю атомів у зразку. Тому можлива величина індукованих зарядів у провіднику по суті є нічим не обмеженою, і їхнє накопичення триває, доки напруженість внутрішнього поля не досягне величини $\vec{E}_i = -\vec{E}_0$, а напруженість результуючого поля — значення $\vec{E} = \vec{E}_0 + \vec{E}_i = 0$. Відтак переміщення електронів припиняється і в провіднику встановлюється такий рівноважний розподіл зарядів, при якому електричне поле існує тільки на поверхні провідника та поза ним. При цьому поле в кожній точці поверхні спрямоване по нормалі (рис. 2.1), бо інакше вільні заряди продовжували би рухатися по поверхні.

Через кулонівське відштовхування так само поводяться й сторонні (надлишкові) заряди в провіднику: вони виходять на поверхню і розміщуються так, що всередині провідника поля відсутнє, а на поверхні воно напрямлене по нормалі.

Із сказаного випливає загальний висновок про те, що

всередині провідника існування статичного електричного поля є неможливим.

Наостанку зазначимо, що описаний тут розгорнуто в часі процес зникнення поля в провіднику насправді відбувається практично миттєво.

Еквіпотенціальність провідника. Вектор напруженості електричного поля пов'язаний з потенціалом співвідношенням ((3.5), розділ I). Тому відсутність статичного електричного поля в провіднику означає, що

$$\operatorname{grad}\varphi = 0 \quad \Rightarrow \quad \varphi = \text{const},$$

тобто, всі точки в об'ємі провідника мають однаковий потенціал φ. Такий самий потенціал мають і всі точки поверхні. Це теж випливає із співвідношення ((3.5), розділ I), за яким функція потенціалу не може мати розривів, інакше на поверхні було би $E \to \infty$.

Таким чином,

при рівновазі зарядів усі точки об'єму та поверхні провідника мають однаковий потенціал.

Тому можна говорити про потенціал усього провідника, тоді як в діелектрику — лише про потенціал заданої точки.

Зауваження. Еквіпотенціальність є важливою властивістю провідників, про яку не можна забувати, розв'язуючи задачі. Приміром, при з'єднанні між собою заряджених металевих куль перерозподіл зарядів на них буде тривати до вирівнювання їхніх потенціалів, а не величини зарядів.

Поле біля поверхні провідника. Вістря. Характеристики електричного поля на поверхні провідника визначаються розподілом індукованих і сторонніх зарядів, який задається поверхневою густиною заряду σ. Для визначення напруженості $\vec{E}$ поля провідника оточимо якусь точку на поверхні невеликою замкненою циліндричною оболонкою, як показано на рис. 2.2.

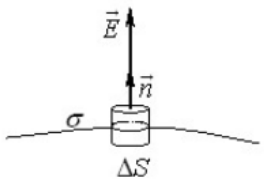


Рис. 2.2

Позаяк поля всередині провідника немає, потік напруженості створюється тільки крізь зовнішню основу циліндрика і наближено дорівнює

$$\Delta\Phi \approx E\,\Delta S.$$

Тому, згідно з теоремою Гаусса (розділ I, п. 5.2.2), можна записати:

$$E\,\Delta S \approx \frac{1}{\varepsilon_0}\sigma\,\Delta S,$$

де $\sigma\,\Delta S = \Delta q$ — заряд ділянки провідника, що охоплена циліндриком. Стягнувши його у вибрану точку на поверхні провідника, отримаємо точний вираз:

$$E = \frac{\sigma}{\varepsilon_0}, \quad \text{або} \quad \vec{E} = \frac{\sigma}{\varepsilon_0}\vec{n}, \tag{2.1}$$

де $\vec{n}$ — орт зовнішньої нормалі до поверхні провідника.

Якщо провідник оточений ізотропним діелектриком з проникністю ε, то напруженість на поверхні виражається формулою

$$\vec{E} = \frac{\sigma}{\varepsilon_0\varepsilon}\vec{n}. \tag{2.1a}$$

Оскільки напруженість в кожній точці поверхні визначається густиною заряду, з'ясуємо, від чого залежить величина σ в різних точках поверхні. Для цього розглянемо провідник, який складається з двох металевих кульок із радіусами R_1 і R_2, що з'єднані довгою тонкою дротиною (рис. 2.3).

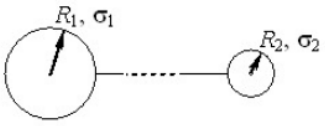


Рис. 2.3

За таких умов кулонівська взаємодія між кульками та заряд, який осідає на дротині, є незначними і можна вважати, що кульки заряджені рівномірно з деякими густинами σ_1 і σ_2.

Кульки є частинами єдиного провідника і мають однакові потенціали $\varphi_1 = \varphi_2$. Тож згідно з формулою ((2.7), розділ I),

$$\frac{q_1}{4\pi\varepsilon_0 R_1} = \frac{q_2}{4\pi\varepsilon_0 R_2}.$$

Підставивши сюди вирази заряду $q = \sigma S$ і площі поверхні кулі $S = 4\pi R^2$, отримаємо:

$$\sigma_1 R_1 = \sigma_2 R_2 \quad \Rightarrow \quad \frac{\sigma_1}{\sigma_2} = \frac{R_2}{R_1}. \tag{2.2}$$

Таким чином, поверхнева густина зарядів на кульках є обернено пропорційною радіусам, які визначають кривизну $K = 1/R$ їхніх поверхонь.

Для провідника довільної форми співвідношення (2.2), звісно, точно не виконується. Але воно відображає загальну тенденцію:

> ***сторонні заряди у провіднику розподіляються відповідно до форми його поверхні так, що поверхнева густина заряду в кожній точці є тим більшою, чим більша кривизна поверхні в цій точці.***

З цієї причини біля вістер і гострих виступів або ребер на поверхні зарядженого провідника електричне поле може виявитися настільки сильним, що викличе електричний пробій у прилеглому шарі діелектрика. Тому в електричному устаткуванні поверхні елементів, що знаходяться під високою напругою, як скажімо ізолятори повітряних ліній передач, роблять гладенькими і закругленими.

Але накопичення зарядів на нерівностях поверхні провідника має і корисні застосування. Одним із прикладів є *громовідвід* — простий і ефективний пристрій для захисту споруд чи устаткування від ураження блискавкою. Громовідвід являє собою загострений металевий штир, який встановлюється на даху будівлі чи поруч із обладнанням і з'єднується металевою шиною із закопаним у землю листом металу. Громовідвід "працює" так. Перед і під час грози в хмарах і повітрі накопичується багато заряджених частинок — іонів, які й спричинюють блискавки, тобто гігантські іскрові розряди між хмарами та наземними об'єктами. Частина вказаних зарядів осідає на загостреному кінці громовідводу і створює навколо нього сильне неоднорідне поле. Під його дією інші атмосферні заряди притягаються до вістря і далі стікають у землю. Внаслідок цього повітря навколо громовідводу стає збідненим на заряди, і ймовірність потрапляння блискавки в споруду значно зменшується. Тож усупереч поширеній думці, громовідвід не приймає блискавку на себе, хоча таке інколи й трапляється, а "відводить" її від споруди.

На властивостях вістря ґрунтується і робота іонного проектора — надпотужного мікроскопа, який дозволяє спостерігати навіть розташування окремих атомів у металі. Проектор, спрощена схема якого наведена на рис. 2.4, складається із скляного балона та дуже гострої голки з досліджуваного металу, яка є одним з електродів.

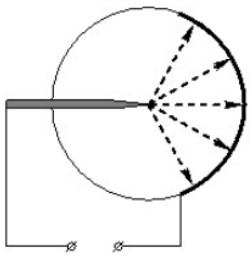


Рис. 2.4

Внутрішня поверхня балона вкрита тонким шаром провідної люмінесцентної речовини, що є другим електродом. Балон відкачують і натомість впускають невелику кількість атомарного газу (зазвичай гелію).

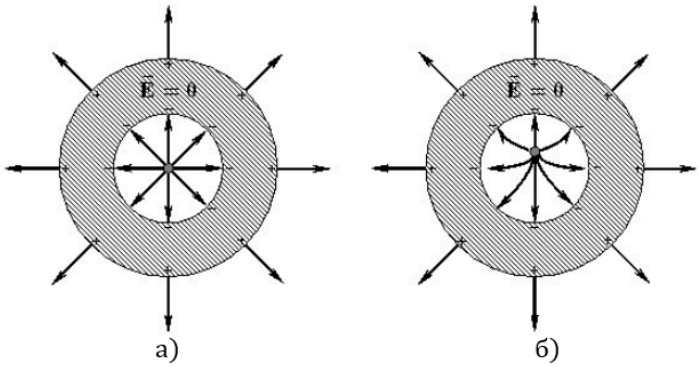
Прилад працює так. При подаванні на електроди високої напруги навколо вістря створюється дуже сильне неоднорідне електричне поле, в якому атоми газу поляризуються (див. п. 1.1) і починають рухатися до вістря (див. розділ I, п. 4.2). Радіус кривизни кінчика добре заточеної голки може складати всього 0,1 мкм, а напруженість поля досягати $\sim 5 \cdot 10^7$ В/м. У такому надсильному полі атоми газу біля вістря іонізуються — розділяються на електрони та іони. Електрони падають на голку, а іони відбиваються і рухаються по силових лініях до балона, де вдаряючи в покриття, викликають його свічення, як у телевізійній трубці. Поверхня вістря складається з окремих атомів і є зернистою, тому у місцях виходу атомів напруженість і щільність ліній поля вища, ніж у проміжках між атомами. Відповідно, до таких точок підлітає і відбивається відносно більша кількість іонів, і на поверхні колби утворюється система світлих цяток, положення яких відображає розташування атомів на поверхні вістря.

Із властивостями вістер пов'язані й інші ефекти, зокрема, електризація тертям.

Електростатичне екранування. Якщо провідник є порожнистим, то статичного електричного поля не може бути не лише в самому провіднику, а й у порожнині. Це легко довести. Якщо суцільний провідник зарядити або внести у зовнішнє поле, то заряди розташуються на поверхні, але в об'ємі речовина лишиться електрично нейтральною в кожній точці. Тож, якщо з об'єму видалити частину речовини, утворивши порожнину, то це не призведе до зміни розподілу зарядів на поверхні та появи в порожнині поля. Звідси, до речі, випливає, що у порожнистому провіднику за відсутності зарядів у порожнині сторонні заряди розміщуються тільки на зовнішній поверхні.

Таким чином, ніякі заряди, розміщені на поверхні провідної оболонки чи поза нею, не можуть створити електричного поля всередині. Тому провідна оболонка забезпечує повний захист (екранування) всього, що знаходиться всередині, від впливу будь-яких зовнішніх електростатичних полів. Це широко використовується на практиці. Зокрема, в електроніці задля захисту чутливих елементів схеми їх вміщують у закриті металеві корпуси, а проводи — в металеву обплітку.

Провідна оболонка “працює” і в зворотному напрямку. Наприклад, якщо в порожнині розмістити точковий заряд q (рис. 2.5*а*), то його електричне поле індукує на поверхнях провідника заряди $q_i^+ = q$ і $q_i^- = -q$, які компенсують поле в самому провіднику.



а) б)

Рис. 2.5

Тому поле назовні провідника буде визначатися тільки величиною та розподілом заряду q_i^+ на зовнішній поверхні, причому зміна поля всередині внаслідок зміщення заряду q і перерозподіл заряду q_i^- на внутрішній поверхні порожнини ніяк не вплине на зовнішнє поле (рис. 2.5*б*). Якщо зовнішню поверхню оболонки заземлити, то ніякого поля назовні взагалі не буде. Таким чином,

> ***замкнена провідна оболонка поділяє весь простір на дві електрично незалежні області.***

Фундаментальна причина цього полягає в тому, що у випадку провідної оболонки на шляху із середини назовні обов’язково трапиться область (сама оболонка), в якій немає електричного поля — “передавача” взаємодії між зарядами, розташованими всередині та назовні оболонки.

2.2. *Електрична ємність. Конденсатори*

Ємність відокремленого провідника. Напруженість електричного поля на поверхні зарядженого провідника в кожній точці визначається густиною заряду (формула (2.1а)) і має різну величину та напрям. Натомість потенціал провідника є скрізь однаковий, тож у кожній точці визначається не густиною заряду в цій точці, а зарядом усього провідника. При цьому дослід свідчить, що потенціал провідника є прямо пропорційним до його заряду:

$$\varphi = \frac{1}{C} q \quad \text{або} \quad q = C\varphi.$$

Величина

$$C = \frac{q}{\varphi} \tag{2.3}$$

називається електричною ємністю провідника, **електроємністю**, або просто **ємністю**.

Одиницею ємності є фарад (Ф). Згідно з (2.3) можна сказати, що

> ***фарад дорівнює ємності провідника, котрому для збільшення потенціалу на*** 1 В ***треба надати заряд*** 1 Кл***. Отже,*** 1 Ф $= 1$ Кл/В**.**

Ємність не залежить ані від заряду, ані від потенціалу і для відокремленого провідника визначається тільки його геометричною конфігурацією та властивостями діелектричного середовища, в якому він знаходиться.

Примітка. Відокремленим називається провідник, який є єдиним у всьому просторі. Звичайно, це ідеалізація. Реально провідник можна вважати відокремленим, якщо вплив на нього з боку інших тіл є неістотним.

Задля ілюстрації визначимо ємність провідної кулі радіуса R, що знаходиться в безмежному однорідному ізотропному діелектрику з проникністю ε. За таких умов заряд розподіляється по поверхні кулі рівномірно, і її електричне поле збігається з полем точкового заряду (1.22). Тож потенціал

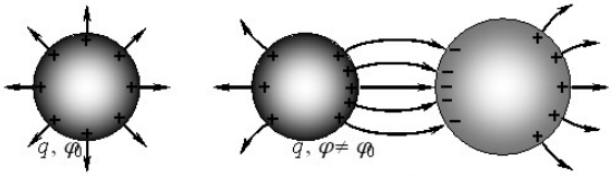
$$\varphi = \frac{q}{4\pi\varepsilon_0\varepsilon R}.$$

Тоді за означенням (2.3) ємність кулі

$$C = 4\pi\varepsilon_0\varepsilon R, \tag{2.4}$$

як і сказано, визначається тільки геометрією та властивостями середовища, в якому вона знаходиться.

Конденсатори, ємність конденсатора. Електрична ємність як фізична властивість провідників має важливі застосування в електротехніці та електроніці. Але окремі провідники в цьому сенсі мають принципові вади. По-перше, ємність окремо взятого провідника є дуже малою. Наприклад, згідно з формулою (2.4), ємність провідної кулі з розмірами Землі ($R \approx 6400$ км) у вакуумі складає всього біля $7 \cdot 10^{-4}$ Ф. А по-друге, в реальних умовах ємність окремого провідника не має визначеної величини, бо істотно залежить від оточення. Справді, якщо заряджена металева куля є відокремленою (рис. 2.6*а*), то її заряд q розподілений рівномірно, і вона має відповідний потенціал φ_0 та ємність $C_0 = q/\varphi_0$. Але коли наблизити до кулі якесь тіло, наприклад, незаряджений кусок металу, то під дією поля кулі на його поверхні з'являться індуковані заряди q_i^- і q_i^+, як показано на рис. 2.6*б*. Це призведе до зміщення зарядів кулі ближче до металу, бо індуковані заряди q_i^- будуть притягати заряди кулі сильніше, ніж заряди q_i^+ — відштовхувати їх. Такий перерозподіл заряду кулі q змінить її потенціал φ, тож і ємність $C = q/\varphi$. При цьому, як показує глибший аналіз, присутність інших тіл завжди зменшує потенціал і збільшує ємність даного тіла. Крім того, ємність буде змінюватися при будь-якій зміні взаємного розташування кулі та сусіднього тіла.



а) б)

Рис. 2.6

Указана нестабільність може негативно впливати на роботу апаратури і вимагає спеціального захисту (екранування) чутливих вузлів електронних пристроїв. Але оточення практично не впливає на системи провідників, які називаються конденсаторами. Найпростіший і найуживаніший ***плоский конденсатор*** складається з двох паралельних металевих пластин — обкладок, простір між якими заповнений діелектриком, зокрема, повітрям (рис. 2.7).

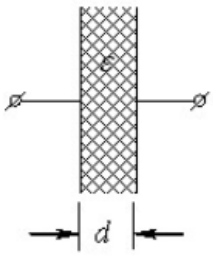


Рис. 2.7

Обкладки розміщують близько одна від одної, тому електричне поле в конденсаторі, за винятком країв пластин, є однорідним, як у випадку нескінченних площин. Спотворення поля на краях пластин (так званий "*крайовий ефект*") є незначними, тому в теорії ними нехтують і розглядають конденсатор як ідеальний.

Нечутливість конденсаторів до впливу зовнішніх тіл і полів пояснюється ефектом електростатичного екранування. Це неважко зрозуміти. При підключенні конденсатора до полюсів джерела напруги частина електронів з однієї обкладки переміщується на іншу, й на них з'являться різнойменні заряди однакової величини q і $-q$ та відповідна різниця потенціалів $U = \varphi_+ - \varphi_-$.

(Принагідно зазначимо, що величина $q > 0$ називається зарядом, а U — напругою конденсатора.)

Заряди в конденсаторі через кулонівське притягання розміщуються на внутрішніх поверхнях пластин. Тож створюване між ними електричне поле є відокремлене від зовнішнього простору самими пластинами, бо всередині них електричного поля не може бути (див. п. 2.1, *електростатичне екранування*). Тому поле конденсатора не може вийти за його межі, а зовнішнє поле не може потрапити всередину. З цієї причини зовнішні тіла та поля не можуть впливати на конденсатор, а він — на них.

Іншою цінною властивістю конденсаторів є те, що вони при невеликих розмірах можуть мати велику ємність і акумулювати значну кількість електричної енергії. Тому конденсатори мають практичне застосування як накопичувачі електричної енергії.

Типи конденсаторів. Електрична ємність конденсатора визначається відношенням його заряду до напруги:

$$C = \frac{q}{U}. \tag{2.5}$$

Тому, аби визначити ємність конденсатора, треба за допомогою відомих методів розрахувати напругу U на конденсаторі при заданій величині заряду q і скористатися формулою (2.5). Проілюструємо це на прикладі ідеального плоского конденсатора з площею пластин S, відстанню між ними d і проникністю діелектрика ε. Напруженість поля в такому конденсаторі визначається формулою (1.23):

$$E = \frac{\sigma}{\varepsilon_0 \varepsilon} = \frac{q}{\varepsilon_0 \varepsilon S}.$$

Це поле є однорідним, тому відповідно до п. 3.1, розділ I, напруга на конденсаторі

$$U = Ed = \frac{qd}{\varepsilon_0 \varepsilon S},$$

і, згідно з означенням (3.5), ***ємність плоского конденсатора*** визначається формулою:

$$C = \frac{\varepsilon_0 \varepsilon S}{d}. \tag{2.6}$$

Отже, чим більша проникність діелектрика та площа обкладок і менша відстань між ними, тим більша ємність конденсатора. У цьому зв'язку слід зазначити, що сучасні технології дозволяють виготовляти конденсатори ємністю ~ 1 Ф розмірами з монету.

На практиці використовують переважно плоскі конденсатори. Але в теорії розглядають ще два типи простих конденсаторів — сферичні та циліндричні. Обкладками сферичного конденсатора є дві концентричні металеві оболонки з якимись радіусами R_1 і R_2 (рис. 2.8*а*), а циліндричного — коаксіальні циліндри з радіусами R_1 і R_2 і довжиною l (рис. 2.8*б*).

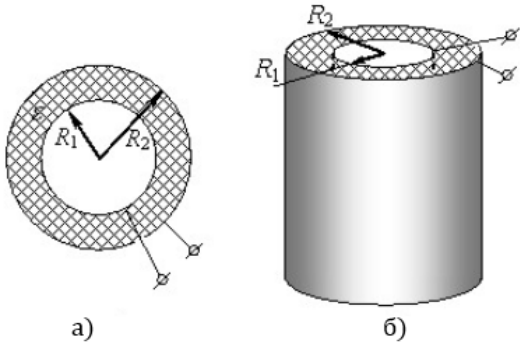


Рис. 2.8

Формули ємності таких конденсаторів розраховуються так само, як для плоского, і мають вигляд:

$$C_{\text{сфер}} = \frac{4\pi\varepsilon_0 \varepsilon R_1 R_2}{R_2 - R_1}, \tag{2.7}$$

$$C_{\text{цил}} = \frac{2\pi\varepsilon_0 \varepsilon l}{\ln(R_2/R_1)}. \tag{2.8}$$

З'єднання конденсаторів. В електричних колах окрім поодиноких конденсаторів використовують з'єднання (батареї) декількох певним чином сполучених між собою конденсаторів. Виведемо формули ємності найпростіших паралельного та послідовного з'єднань, виходячи з того, що ємність з'єднання, як і одного конденсатора, визначається формулою (2.5) через сумарний заряд батареї та подану на неї напругу.

Паралельне з'єднання. При паралельному з'єднанні всі конденсатори безпосередньо сполучаються між собою, як показано на рис. 2.9*а*.

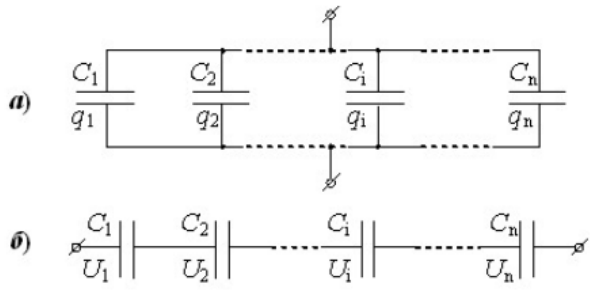


Рис. 2.9

Усі з'єднані обкладки утворюють єдиний провідник і за будь-яких умов мають однаковий потенціал. Тому, якщо подати на батарею напругу U, то напруги U_i всіх конденсаторів будуть такими самими: $U_i = U$. Також є очевидним, що заряд батареї дорівнює сумі зарядів усіх з'єднаних конденсаторів: $q = \sum q_i$. Отже, ємність з'єднання

$$C_{\text{пар}} = \frac{q}{U} = \frac{\sum q_i}{U} = \sum_i \frac{q_i}{U_i} \quad \Rightarrow$$
$$\Rightarrow \quad C_{\text{пар}} = \sum C_i. \tag{2.9}$$

Таким чином,

при паралельному сполученні конденсаторів ємність з'єднання дорівнює сумі ємностей всіх сполучених конденсаторів.

Зокрема, при паралельному з'єднанні n конденсаторів однакової ємності C загальна ємність

$$C_{\text{пар}} = nC. \tag{2.9а}$$

<u>Послідовне з'єднання</u>. В цьому випадку конденсатори утворюють ланцюг, у кожній ланці котрого безпосередньо з'єднуються тільки два конденсатори, як показано на рис.2.9*б*. При такому сполученні питання про ємність з'єднання не є таким очевидним, як при паралельному, тому проаналізуємо сам процес заряджання послідовного ланцюга конденсаторів. У момент приєднання до джерела заряди q і $-q$ подаються тільки на зовнішні пластини першого і останнього конденсаторів C_1 і C_n. Але під дією поля, що виникає, тієї ж миті на всіх з'єднаних між собою парах пластин з'являються індуковані заряди однакової величини q_i і $-q_i$, і всі конденсатори виявляються зарядженими. При цьому напруга на ланцюгу дорівнює сумі напруг на окремих конденсаторах: $U = \sum U_i$, а заряди всіх конденсаторів q_i однакові і дорівнюють заряду всього ланцюга: $q_i = q$. Це легко зрозуміти на прикладі ідеальних плоских конденсаторів. У момент підключення до джерела між крайніми пластинами ланцюга виникає однорідне електричне поле $\vec{E}$. Але всередині з'єднаних пластин поля бути не може, тому індуковані заряди на з'єднаних пластинах мають створювати внутрішнє поле $\vec{E}_i = -\vec{E}$. А для цього величина індукованих зарядів q_i повинна дорівнювати величин поданих на ланцюг сторонніх зарядів q.

Таким чином, ємність послідовного з'єднання конденсаторів

$$C_{\text{пос}} = \frac{q}{U} = \frac{q}{\sum U_i},$$

де

$$U_i = \frac{q_i}{C_i} = \frac{q}{C_i}.$$

Після такої підстановки виходить наступна формула для розрахунку ємності послідовного з'єднання конденсаторів:

$$\frac{1}{C_{\text{пос}}} = \sum \frac{1}{C_i}. \tag{2.10}$$

Отже,

при послідовному з'єднанні конденсаторів додаються не ємності, а їхні обернені величини.

Із загального виразу (2.10) випливають прості формули прямого розрахунку ємності послідовного з'єднання двох конденсаторів з ємностями C_1 і C_2:

$$C_{\text{пос}} = \frac{C_1 C_2}{C_1 + C_2} \tag{2.10а}$$

та n конденсаторів однакової ємності C:

$$C_{\text{пос}} = \frac{C}{n}. \tag{2.10б}$$

Формули (2.9) і (2.10) дозволяють розраховувати й більш складні комбіновані з'єднання, які можна розчленувати на окремі паралельні та послідовні ділянки.

2.3. *Енергія електричного поля*

Енергія та її трансформації є універсальними чинниками, що визначають стан будь-якої фізичної системи і процеси в ній. Згадаймо хоча би про механічну енергію та її роль в механічних процесах. Електростатична система, тобто сукупність заряджених частинок або тіл, має відповідний запас електричної енергії. Її походження неважко збагнути. Сили кулонівської взаємодії завжди намагаються звести докупи і нейтралізувати різнойменні заряди та розкидати на нескінченність однойменні. Тому при створенні заданої конфігурації зарядів треба виконувати роботу по подоланню кулонівських сил, тим самим надаючи системі відповідного запасу електростатичної енергії. Позаяк кулонівські сили є консервативними, ця енергія є різновидом потенціальної енергії і, як усяка енергія, може переходити в інші форми. Наприклад, в кінетичну, якщо зарядам системи дозволити рухатися під дією сил кулонівської взаємодії між ними.

Енергія взаємодії системи точкових зарядів. Розглянемо гіпотетичний процес утворення системи з трьох точкових зарядів шляхом перенесення їх з нескінченності, де потенціал поля $\varphi = 0$, в задані точки 1, 2, 3 (рис. 2.10):

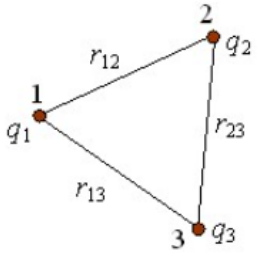


Рис. 2.9

Переміщення першого заряду q_1 в точку 1 відбувається вільно, але наступний заряд q_2 вже рухається в полі заряду q_1 і для його переміщення в точку 2 треба виконати деяку роботу A' проти сил поля. Ця робота дорівнює взятій з протилежним знаком роботі A самого поля і визначає потенціальну енергію W_{12} заряду q_2 в полі заряду q_1. Тож, відповідно до формул ((2.10а) і (2.7), розділ I),

$$W_{12} = k\frac{q_1 q_2}{r_{12}}. \tag{2.11}$$

Але, якщо спочатку розташувати заряд q_2 в точці 2, а потім перенести в точку 1 заряд q_1, то нічого не зміниться крім того, що вираз (2.11) буде трактуватись як потенціальна енергію заряду q_1 в полі заряду q_2. Тому величина W_{12} є "спільним надбанням" обох зарядів і визначає енергію їхньої електричної взаємодії.

При включенні в систему заряду q_3 додасться ще й енергія його взаємодії з першим та другим зарядами, котра виражається аналогічно. Отже, енергія системи з трьох зарядів

$$\begin{aligned} W &= W_{21} + W_{31} + W_{32} = \\ &= k\frac{q_2 q_1}{r_{12}} + k\frac{q_3 q_1}{r_{13}} + k\frac{q_3 q_2}{r_{23}}. \end{aligned}$$

Позаяк у кожній парі заряди є рівноправними партнерами взаємодії, перепишемо цей вираз симетрично, як

$$\begin{aligned} W = \frac{1}{2}\Bigg(& q_1 k\frac{q_2}{r_{12}} + q_2 k\frac{q_1}{r_{12}} + q_2 k\frac{q_3}{r_{23}} + \\ & + q_3 k\frac{q_2}{r_{23}} + q_3 k\frac{q_1}{r_{13}} + q_1 k\frac{q_3}{r13}\Bigg), \end{aligned}$$

або, після перегрупування,

$$W = \frac{1}{2}\left(q_1\left(k\frac{q_2}{r_{12}} + k\frac{q_3}{r_{13}}\right) + \right.$$
$$+ q_2\left(k\frac{q_1}{r_{12}} + k\frac{q_3}{r_{23}}\right) +$$
$$\left. + q_3\left(k\frac{q_1}{r_{13}} + k\frac{q_2}{r_{23}}\right)\right).$$

Згідно з формулою ((2.7а), розділ I), вираз у кожній малій дужці визначає потенціал поля, що створюється в місці розташування відповідного заряду двома іншими.

Зрозуміло, що при додаванні наступних зарядів нічого принципово не зміниться. Тому електростатична енергія взаємодії довільної системи точкових зарядів виражається загальною формулою

$$W = \frac{1}{2}\sum q_i\varphi_i, \tag{2.12}$$

де φ_i — потенціали полів, що створюються в місцях розташування зарядів q_i всіма іншими зарядами системи.

Енергія зарядженого діелектрика та провідника. В протяжних тілах заряд розподіляється в просторі неперервно. Тому для підрахунку електростатичної енергії в зарядженому діелектрику увесь заряд треба подумки розчленувати на елементарні точкові порції $\mathrm{d}q$ і замінити у виразі (2.12) дискретне додавання на інтегрування по всій зарядженій області:

$$W = \frac{1}{2}\int \varphi\,\mathrm{d}q. \tag{2.13}$$

Якщо заряд розподілений по об'єму з густиною $\rho(\vec{r})$, то $\mathrm{d}q = \rho\,\mathrm{d}V$, і

$$W = \frac{1}{2}\int\limits_V \varphi\rho\,\mathrm{d}V. \tag{2.14}$$

При розподілі заряду по поверхні з густиною $\sigma(\vec{r})$ або по тонкому стержню з лінійною густиною $\lambda(\vec{r})$ електростатична енергія виражається, відповідно, як

$$W = \frac{1}{2}\int\limits_S \varphi\sigma\,\mathrm{d}S \tag{2.14а}$$

і

$$W = \frac{1}{2}\int\limits_L \varphi\lambda\,\mathrm{d}l. \tag{2.14б}$$

Провідники в електростатичному стані є еквіпотенціальними. Тому у виразі (2.13) $\varphi = \mathrm{const}$, і електростатична енергія провідника визначається формулою

$$W = \frac{1}{2}q\varphi. \tag{2.15}$$

Стосовно отриманих результатів необхідно зробити важливе зауваження. Хоча вирази (2.12) і (2.13) — (2.15) за загальним змістом видаються ідентичними, між ними є принципова відміна. Це стає зрозумілим, якщо згадати, що точкові заряди є абстракцією. Реально "точкові" заряди є частинками скінченних розмірів, тож заряди в них теж є рознесені в просторі і взаємодіють між собою. З цією взаємодією пов'язана відповідна *власна* електростатична енергія зарядів, не відображена у формулі (2.12). Напроти, формули (2.13) —(2.15) враховують взаємодію всіх зарядів тіла і визначають власну електростатичну енергію. А ось в електростатичну енергію *системи тіл* окрім власної енергії $W_{\text{вл}} = \sum W_i$, яка дорівнює сумі енергій всіх тіл, входить іще й енергія взаємодії $W_{\text{вз}}$ кожного з кожним:

$$W = W_{\text{вл}} + W_{\text{вз}}.$$

Таким чином, електростатична енергія довільної системи не дорівнює сумі енергій окремих її частин (тіл), тобто не є адитивною величиною. Тому в загальному випадку при підрахунках не можна забувати про енергію електростатичної взаємодії $W_{\text{вз}}$, яка визначається *конфігурацією* системи — взаємним розташуванням заряджених тіл. Але для точкових заряджених тіл через їхні дуже малі розміри зміна взаємного розташування зарядів в них практично не впливає на власну електростатичну енергію, тож

$W_{\text{вл}} = \text{const}$. Тому власна енергія системи точкових зарядів ніяк не впливає на процеси в ній і може бути відкинута. Отже, хоча формула (2.12) принципово є неповною, фактично вона дає всю необхідну інформацію про електростатичну енергію системи точкових зарядів.

Енергія зарядженого конденсатора. Ця енергія дорівнює роботі по заряджанню конденсатора. Для її підрахунку будемо вважати, що конденсатор заряджається шляхом перенесення нескінченно малих порцій заряду з однієї обкладки на іншу (як це й відбувається в дійсності). Тоді, якщо при напрузі на конденсаторі u перенести з однієї обкладки на іншу порцію заряду $\mathrm{d}q$, виконавши роботу $\delta A'$, то енергія конденсатора збільшиться на величину виконаної роботи:

$$\mathrm{d}W = \delta A' = u\,\mathrm{d}q.$$

Виразивши за формулою (2.5) перенесений заряд через ємність конденсатора та приріст напруги як $\mathrm{d}q = C\,\mathrm{d}u$, і проінтегрувавши вираз $\mathrm{d}W$ у границях від 0 до кінцевої напруги U, одержимо наступну формулу енергії зарядженого конденсатора:

$$W = \frac{CU^2}{2}, \tag{2.16}$$

або рівнозначні формули

$$W = \frac{q^2}{2C} \tag{2.16а}$$

і

$$W = \frac{1}{2}qU. \tag{2.16б}$$

Енергія електричного поля. Всі отримані вище формули виражають електростатичну енергію через характеристики тіл — заряди, потенціали та ємності. Але її можна подати й інакше. Якщо для плоского конденсатора у формулі (2.16) ємність виразити за формулою (2.6), а напругу — за формулою ((3.6а), розділ I), то вийде

$$W = \frac{\varepsilon_0\varepsilon E^2}{2}V, \tag{2.17}$$

де $Sd = V$ — об'єм області простору між пластинами.

В об'ємі V зосереджене все поле конденсатора, тож формула (2.17) виражає енергію конденсатора вже не через характеристики пластин, а тільки через характеристики електричного поля між ними. Це наводить на думку, що саме в полі, а не на пластинах, зосереджена електростатична енергія конденсатора. Якщо так, то величина

$$w_e = \frac{\varepsilon_0\varepsilon E^2}{2} \tag{2.18}$$

визначає об'ємну густину енергії електричного поля в ізотропному середовищі, тобто енергію, зосереджену в одиниці об'єму поля.

Слід зауважити, що в електростатиці таке припущення неможливо ані довести, ані спростувати. Але в електродинаміці, де досліджуються явища в нестаціонарних (змінних) полях, теорія й експеримент однозначно доводять, що

> ***електричне поле має енергію, яка в ізотропному середовищі розподілена з об'ємною густиною (2.18).***

Об'ємну густину енергії електричного поля можна виразити ще двома рівнозначними формулами, врахувавши зв'язок між векторами $\vec{E}$ і $\vec{D}$, а саме:

$$w_e = \frac{D^2}{2\varepsilon_0\varepsilon} \tag{2.18а}$$

і

$$w_e = \frac{ED}{2}. \tag{2.18б}$$

У неоднорідному полі об'ємна густина характеризує концентрацію енергії в кожній точці і визначається як

$$w = \frac{\mathrm{d}W}{\mathrm{d}V}. \tag{2.19}$$

У цьому випадку енергія поля в заданій області V є інтегралом (сумою) енергій $\mathrm{d}W = w\mathrm{V}$ в усіх елементарних областях $\mathrm{d}V$:

$$W = \int\limits_V w\,\mathrm{d}V.$$

Отже, згідно з формулами (2.18) маємо наступні вирази для обчислення енергії електричного поля в заданому об'ємі:

$$W = \int\limits_V \frac{\varepsilon_0 \varepsilon E^2}{2}\mathrm{d}V; \tag{2.20}$$

$$W = \int\limits_V \frac{D^2}{2\varepsilon_0 \varepsilon}\mathrm{d}V; \tag{2.20а}$$

$$W = \int\limits_V \frac{ED}{2}\mathrm{d}V. \tag{2.20б}$$

На завершення зазначимо, що всі результати, отримані тут для статичних полів, зберігають чинність і для залежних від часу (нестаціонарних) полів.

Контрольні запитання

1. Чому всередині діелектрика електростатичне поле лише послаблюється, а в провіднику повністю зникає?

2. Що таке електростатична індукція? Як розподіляються в провіднику індуковані заряди?

3. У чому полягає причина еквіпотенціальності провідника? Чому потенціали об'єму та поверхні провідника не можуть бути різними?

4. Якою формулою визначається напруженість електричного поля біля поверхні провідника? Поміркуйте, чому напруженість біля поверхні провідника вдвічі більша, ніж біля поверхні зарядженої площини з такою самою густиною заряду (формула (3.2), розділ I)).

5. Від чого залежить поверхнева густина заряду в різних точках поверхні провідника?

6. Які особливості має електричне поле вістер? Чому натертий об чисте волосся гребінець притягає до себе мілкі незаряджені клаптики паперу?

7. Поясніть принцип дії громовідводу. Що й куди він"відводить"?

8. Чому ніякі зовнішні заряди та електричні поля не діють на заряди вміщені всередину провідної оболонки?

9. Знайдіть формулу напруженості електричного поля назовні незарядженої металевої сферичної оболонки, в центрі якої знаходиться точковий заряд q. Чим створюється це поле?

10. Заряд q знаходиться у порожнині в куску металу. Поясніть, що відбуватиметься в порожнині й назовні з напруженістю поля та з розподілом індукованих зарядів на поверхнях провідника при переміщенні заряду q у порожнині.

11. Дайте означення електричної ємності відокремленого провідника та конденсатора.

12. В яких одиницях вимірюється електрична ємність? Чому не можна дати означення одиниці ємності безпосередньо з формули (2.3) для ємності відокремленого провідника?

13. Поясніть, чому ємність відокремленого провідника залежить від зовнішніх тіл і полів, а ємність конденсатора — ні.

14. Запишіть формули для обчислення ємності паралельного та послідовного з'єднань конденсаторів і покажіть, що у випадку послідовного з'єднання загальна ємність є завжди меншою за найменшу із ємностей, що входять до з'єднання.

15. Поясніть походження електростатичної енергії в системи заряджених частинок чи тіл.

16. Поясніть, як буде змінюватись електростатична енергія системи закріплених зарядів, якщо всі, або частину зарядів вивільнити.

ІІІ. ЕЛЕКТРИЧНИЙ СТРУМ

Значна частина електромагнітних явищ, які знаходять широке практичне застосування, пов’язана з електричним струмом — упорядкованим рухом частинок, який супроводжується перенесенням електричного заряду.

У цьому розділі розглянуті такі питання:

1. Загальні закони електричного струму

2. Електричне коло

3. Струм у колі з конденсатором

4. Робота і потужність струму

1. Загальні закони електричного струму

Електричним струмом називається напрямлений рух заряджених частинок — носіїв струму. Носіями струму можуть бути всякі заряджені частинки. В металах і сплавах це електрони провідності — вивільнені при об'єднанні атомів у тверду речовину валентні електрони, які можуть переміщуватися по всьому об'єму зразка. У рідких провідниках (електролітах) носіями струму є іони, а в газах — і те, й інше. Потоки заряджених частинок у вакуумі теж створюють електричний струм.

Кількісні характеристики та загальні закони струму розглядаються далі в наступних питаннях:

1.1. Сила та густина струму

1.2. Рівняння неперервності

1.3. Дрейф носіїв струму. Закон Ома в локальній формі

1.4. Питома потужність струму

1.5. Обмеженість класичної теорії електропровідності

Контрольні запитання

1.1. *Сила та густина струму*

Сутність явища електричного струму полягає в перенесенні в просторі електричного заряду, а його кількісною характеристикою є сила струму, або величина струму, котра визначається кількістю заряду, що переноситься через задану поверхню за одиницю часу. У випадку стаціонарного (постійного) струму сила струму

$$I = \frac{q}{t}. \tag{1.1}$$

Для нестаціонарного (залежного від часу) струму величина (1.1) визначає *середню* силу струму за час t, а його точна (*миттєва*) величина в кожен момент часу визначається відношенням елементарної порції перенесеного заряду $\mathrm{d}q$ до відповідного елементарного проміжку часу $\mathrm{d}t$:

$$I = \frac{\mathrm{d}q}{\mathrm{d}t}. \tag{1.1а}$$

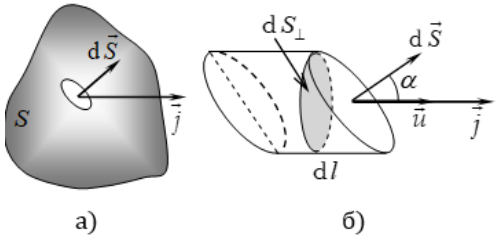
Одиниця сили струму називається ампер (А). Вона визначається через магнітну взаємодію струмів і прийнята як одна з основних одиниць міжнародної системи (SI). Відтак одиниці всіх інших електричних і магнітних величин у кінцевому підсумку визначаються через ампер. Зокрема, одиниця заряду 1 Кл = 1 А · с — то є заряд, який проходить через поперечний переріз провідника за одиницю часу при силі струму 1 А.

Сила струму є скалярною величиною, проте струм як процес має напрям, який визначається напрямом перенесення носіями електричного заряду. При цьому, позаяк існує два знаки заряду, за домовленістю

напрям струму прийнято напрям перенесення позитивного заряду.

Отже в металах, де носіями є електрони, напрям струму є протилежним до напрямку впорядкованого руху носіїв.

У тонких провідниках струм є спрямований уздовж провідника й рівномірно розподілений по перерізу. Але в провідному середовищі маємо справу з об'ємними струмами, котрі в різних точках можуть мати різну величину та напрям. Тому крім інтегральної характеристики — сили струму I використовують іще й локальну величину — вектор густини струму $\vec{j}$, який задає інтенсивність і напрям перенесення заряду в кожній точці. Розглянемо, чим визначається вектор $\vec{j}$ у провідному середовищі з носіями одного сорту. Для цього підрахуємо струм $\mathrm{d}I = \mathrm{d}q/\mathrm{d}t$, який протікає через елементарну ділянку $\mathrm{d}\vec{S}$ якоїсь поверхні S (рис. 1.1*а*).



а) б)

Рис. 1.1

За час $\mathrm{d}t$ крізь указану ділянку пройдуть усі носії, що знаходилися всередині косого циліндрика (рис. 1.1*б*) з основою $\mathrm{d}S$ і твірною $\mathrm{d}l = u\,\mathrm{d}t$, паралельною до швидкості $\vec{u}$ їхнього упорядкованого руху. Заряд цих носіїв

$$\mathrm{d}q = enV\,\mathrm{d}V = enu\,\mathrm{d}t\,\mathrm{d}S\cos\alpha,$$

де e — заряд одного носія, n — їхня концентрація. Отже величина струму через площинку $\mathrm{d}S$ дорівнює

$$\mathrm{d}I = enu\,\mathrm{d}S\cos\alpha = enu\,\mathrm{d}S_\perp, \tag{1.2}$$

де $\mathrm{d}S_\perp = \mathrm{d}S\cos\alpha$ — площа проекції ділянки $\mathrm{d}S$ на площину перпендикулярну до напрямку руху носіїв. Величина

$$j = \frac{\mathrm{d}I}{\mathrm{d}S_\perp} = enu,$$

яка в кожній точці визначає струм, що припадає на одиницю площі поверхні, перпендикулярної до його напрямку, називається густиною струму і вимірюється в $\mathsf{А}/\mathsf{м}^2$. Аби величина j відображала ще й напрям струму, її розглядають як вектор:

$$\vec{j} = en\vec{u}. \tag{1.3}$$

У металах і сплавах носіями струму є електрони, але в інших середовищах у створенні струму можуть брати участь різні частинки-носії, кожна із своєю величиною заряду q_i концентрацію n_i та швидкістю упорядкованого руху $\vec{u}_i$. Тоді

$$\vec{j} = \sum q_i n_i \vec{u}_i. \tag{1.3а}$$

Через вектор густини струму можна згорнуто виразити силу струму крізь елементарну площинку $\mathrm{d}S$ (формула (1.2)) та через довільну поверхню S:

$$\mathrm{d}I = \vec{j}\,\mathrm{d}\vec{S}, \tag{1.4}$$

$$I = \int\limits_S \vec{j}\,\mathrm{d}\vec{S}. \tag{1.4а}$$

Отже, силу струму можна трактувати як потік вектора густини струму через задану поверхню і при необхідності зображувати розподіл струму у просторі лініями вектора $\vec{j}$ (лініями струму).

Якщо струм по площі поперечного перерізу провідника розподіляється рівномірно, то

$$I = jS. \tag{1.4б}$$

1.2. *Рівняння неперервності*

Розглянуті характеристики струму дозволяють подати загальній математичний вираз одного із фундаментальних фізичних законів — закону збереження електричного заряду. Якщо виділити в просторі довільну замкнену поверхню, то згідно з означенням (1.1а) і виразом (1.4а) можна записати:

$$\oint\limits_S \vec{j}\,\mathrm{d}\vec{S} = \frac{\mathrm{d}q_S}{\mathrm{d}t}, \tag{1.5}$$

де $\mathrm{d}q_s/\mathrm{d}t$ — заряд, який проходить за одиницю часу крізь поверхню S назовні. Але позаяк заряд не виникає з нічого і не зникає безслідно, ця похідна в точності дорівнює швидкості зменшення величини заряду в об'ємі, що обмежений даною поверхнею. Отже:

$$\frac{\mathrm{d}q_S}{\mathrm{d}t} = -\frac{\mathrm{d}q_V}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_V \rho\,\mathrm{d}V =$$

$$= -\int\limits_V \frac{\partial\rho}{\partial t}\mathrm{d}V,$$

де ρ — об'ємна густина заряду.

Примітка. Під знаком інтеграла стоїть частинна похідна, тому що об'єм V є обмежений фіксованою (нерухомою) поверхнею, і похідна визначається тільки явною залежністю заряду в об'ємі від часу, а не, скажімо, рухом стінок поверхні.

Підставивши цей вираз у (1.5), отримуємо:

$$\oint\limits_S \vec{j}\,\mathrm{d}\vec{S} = -\int\limits_V \frac{\partial\rho}{\partial t}\mathrm{d}V. \tag{1.6}$$

Рівняння (1.6) називається рівнянням неперервності і є загальним виразом закону збереження електричного заряду. Воно, зокрема, вбирає й закон збереження заряду в електрично ізольованій системі, про який говорилося в електростатиці. Справді, крізь будь-яку замкнену поверхню, що відділяє таку систему від оточення, не проходять заряджені частинки, і в усіх її точках $\vec{j} = 0$. Отже, згідно з (1.6),

$$\int\limits_V \frac{\partial\rho}{\partial t} = 0 \quad \Rightarrow \quad \frac{\partial\rho}{\partial t} = 0 \quad \Rightarrow$$
$$\Rightarrow \quad q(t) = \mathrm{const},$$

тобто повний заряд системи не змінюється з часом (зберігається).

Із рівняння неперервності випливають також деякі загальні висновки щодо умов існування електричного струму. Зокрема, це висновок про те, що постійний струм може існувати лишень у замкнених провідних контурах. Дійсно, при постійному струмі розподіл потенціалів і зарядів у провідному середовищі не змінюється з часом, отже, в рівнянні (1.6) похідна $\partial\rho/\partial t = 0$ в усіх точках. Тому для будь-якої замкненої поверхні

$$\oint\limits_S \vec{j}\,\mathrm{d}\vec{S} = 0.$$

а це означає, що кількість ліній струму (ліній вектора $\vec{j}$), які заходять усередину поверхні та виходять назовні є завжди однаковою. Інакше говорячи, лінії стаціонарного струму є неперервними (замкненими) — вони ніде не починаються і ніде не закінчуються. Це одна з причин, чому рівняння (1.6) називають рівнянням неперервності. З нього випливає той загальновідомий факт, що

постійний струм може протікати тільки по замкненому провідному контуру.

Але з рівняння неперервності випливає й зовсім не самоочевидний висновок про те, що

залежний від часу струм може існувати і в незамкнених контурах,

як от у колах із конденсаторами. При цьому в точках розриву контуру (на пластинах конденсатора) відбувається накопичення зарядів того чи іншого знаку, і в них $(\partial\rho/\partial t) \neq 0$. Тому відповідно до рівняння (1.6), потік $\vec{j}$ крізь малу замкнену поверхню, що охоплює таку точку,

$$\oint\limits_S \vec{j}\,\mathrm{d}\vec{S} > 0 \quad \text{при } \frac{\partial\rho}{\partial t} < 0$$

і

$$\oint\limits_S \vec{j}\,\mathrm{d}\vec{S} < 0 \quad \text{при } \frac{\partial\rho}{\partial t} > 0.$$

Згідно з геометричним змістом потоку ці вирази означають, що лінії змінного струму мають початки і кінці: вони починаються в одних точках і закінчуються в інших. (Чому, попри це, рівняння (1.6) називається рівнянням неперервності, див. розділ V, п. 2.) Отже, рівняння неперервності не заперечує можливості існування електричного струму в незамкненому контурі. Але через обмежену кількість носіїв у провіднику, вони не можуть в точках розриву накопичуватися *весь час*. Тому в незамкненому провіднику довготривалий нестаціонарний струм може існувати лише за умови, що його напрям періодично змінюється на протилежний. Саме такі струми означає термін "змінний струм".

1.3. *Закон Ома в локальній формі*

Дрейф носіїв струму. Розглянемо механізм електричного струму в металевому провіднику на основі уявлень класичної фізики, за якими електрони провідності в металі утворюють "електронний газ" і здійснюють інтенсивний тепловий рух. Його усереднену швидкість v можна оцінити з відомих формул молекулярної фізики:

$$\frac{m\langle v^2\rangle}{2} = \frac{3}{2}kT \quad \Rightarrow \quad v = \sqrt{\frac{3kT}{m}}.$$

Підставивши значення маси електрона $m = 9{,}1 \cdot 10^{-31}$ кг, сталої Больцмана $k = 1{,}38 \cdot 10^{-23}$ Дж/К і прийнявши $T = 293$ К (20° С), отримаємо:

$$v = \sqrt{\frac{3 \cdot 1{,}38 \cdot 10^{-23} \cdot 293}{9{,}1 \cdot 10^{-31}}} \approx 10^5 \text{ м/с}.$$

Отже, електрони здійснюють дуже інтенсивний тепловий рух. Але він є безладним і не спричинює перенесення заряду, тобто електричного струму. Струм створюється впорядкованим рухом носіїв, і саме його швидкість u, відповідно до формули (1.3), визначає густину струму. Задля кращого уявлення про цей рух оцінимо величину u в мідному провіднику при густині струму $j = 10 \text{ А/мм}^2 = 10^7 \text{ А/м}^2$, прийнявши концентрацію носіїв рівною концентрації атомів $n = 8{,}5 \cdot 10^{28} \text{ 1/м}^3$:

$$u = \frac{j}{en} = \frac{10^7}{1{,}6 \cdot 10^{-19} \cdot 8{,}5 \cdot 10^{28}} \approx$$
$$\approx 10^{-3} \text{ м/с} = 1 \text{ мм/с}.$$

Це складає всього 10^{-8} від теплової швидкості, й упорядкований рух окремого електрона на тлі хаотичного теплового руху взагалі не можна розрізнити. Тож величина u визначає не рух окремого електрона, а рух сукупності носіїв струму в цілому. Такий колективний рух називається дрейфом, а швидкість u — дрейфовою швидкістю носіїв. Дрейф носіїв нагадує знесення (дрейф) рою комах “товкунців” слабким вітерцем.

Закон Ома. Мала величина дрейфової швидкості пояснюється тим, що в металі електрони провідності на своєму шляху натикаються на іони кристалічної ґратки, що перешкоджає їхньому впорядкованому рухові вздовж провідника. Між іонами та електронами діють сили притягання. Через це при зустрічі з масивними іонами електрони повністю втрачають швидкість упорядкованого руху, подібно до того, як пластилінова кулька втрачає швидкість при зіткненні зі стіною. Це й перешкоджає дрейфові носіїв. Тому для створення та підтримання струму в провіднику має існувати електричне поле. Це не суперечить твердженню про неможливість в провідниках електричного поля, про що говорилося раніше (розділ II, п. 2), бо в провіднику із струмом заряди-носії не перебувають у рівновазі, й указане поле не є *електростатичним*.

Зауваження. Вільні електрони, звісно, стикаються й між собою. Але ці зіткнення є пружними і, згідно із законом збереження імпульсу, не впливають на сумарний імпульс, тож і на дрейфову швидкість електронів-носіїв струму. Тому зіткнення між електронами не перешкоджає струмові.

Таким чином, через непружні зіткнення з іонами носії вільно рухаються тільки від зіткнення до зіткнення, проходячи в середньому деяку відстань $\langle\lambda\rangle$, яка називається середньою довжиною вільного пробігу. При цьому середній проміжок часу між двома послідовними зіткненнями електрона з іонами $\langle\tau\rangle$, який називається середнім часом вільного пробігу, складає

$$\langle\tau\rangle = \frac{\lambda}{v}, \tag{1.7}$$

де v — швидкість ***теплового*** руху електрона.

Атоми в твердих тілах розміщені дуже щільно, майже впритул, Тому величина $\langle\lambda\rangle$ має бути близькою до відстані між атомами в провіднику. Ця відстань для більшості металів складає $(3 \div 5) \cdot 10^{-10}$ м, а теплова швидкість $v \approx 10^5$ м/с. За таких умов час вільного пробігу виявляється настільки малим, що електрони під дією поля не встигають помітно розігнатися, чим пояснюється мала величина дрейфової швидкості носіїв.

Розглянутий класичний механізм руху електронів провідності в металі дозволяє теоретично встановити один з основних законів електричного струму — закон Ома в *локальній* (інакше *диференціальній*) формі, який визначає зв’язок між струмом та електричним полем у кожній точці металевого провідника. За час між двома послідовними зіткненнями з іонами кожен електрон під дією поля збільшує швидкість руху уздовж провідника від 0 до певної максимальної величини $\vec{u}_m$, яка в середньому дорівнює

$$\langle\vec{u}_m\rangle = \vec{a} \cdot \langle\tau\rangle = \frac{\vec{F}}{m}\langle\tau\rangle = \frac{e\vec{E}}{m}\langle\tau\rangle, \tag{1.8}$$

де $\vec{a}$ — прискорення електрона, e і m — його заряд і маса, $\vec{E}$ — напруженість електричного поля в провіднику і $\langle\tau\rangle$ — середній час вільного пробігу електрона.

Дрейфова швидкість носіїв дорівнює половині величини u_m:

$$\vec{u} = \frac{1}{2}\langle\vec{u}_m\rangle = \frac{e\vec{E}}{2m}\langle\tau\rangle.$$

У такому разі для густини струму відповідно до формули (1.2) маємо:

$$\vec{j} = \frac{e^2 n\langle\tau\rangle}{2m}\vec{E},$$

або, враховуючи вираз (1.7),

$$\vec{j} = \frac{e^2 n\langle\lambda\rangle}{2mv}\vec{E}.$$

Згорнуто ці формули записують у вигляді

$$\vec{j} = \sigma\vec{E} \tag{1.9}$$

або

$$\vec{j} = \frac{\vec{E}}{\rho}. \tag{1.9а}$$

Величина

$$\sigma = \frac{e^2 n\langle\tau\rangle}{2m} = \frac{e^2 n\langle\lambda\rangle}{2mv} \tag{1.10}$$

показує, наскільки добре дана речовина проводить струм і називається питомою електропровідністю. Відповідно, обернена величина

$$\rho = \frac{1}{\sigma} = \frac{2m}{e^2 n\langle\lambda\rangle} \tag{1.10а}$$

називається питомим опором речовини.

Формули (1.10), (1.10а) виражають один з основних законів електричного струму — закон Ома в локальній формі:

> ***густина електричного струму в кожній точці провідника є прямо пропорційною напруженості поля в цій точці.***

Одразу зазначимо, що закон Ома не є загальним. Він виконується тільки в *лінійних провідниках* — таких, в яких електропровідність не залежить від напруженості поля, тож і від напруги на їхніх кінцях. Окрім металів і сплавів, лінійними є однорідні напівпровідники та електроліти (рідкі провідники) — середовища, в яких концентрація носіїв струму n, відтак і електропровідність σ, не залежать від напруженості поля E. А ось гази та контакти різних напівпровідників або напівпровідників із металами є нелінійними провідниками. Тому в газорозрядних і вакуумних лампах, а також у напівпровідникових електронних приладах закон Ома не виконується.

1.4. *Питома потужність струму*

Стикаючись із іонами металу, електрони провідності передають їм імпульс та енергію, набуті в дрейфовому русі. Ця енергія називається енергією електричного струму і далі перетворюється на ті чи інші види. Наприклад, у газорозрядній лампі частина цієї енергії трансформується в енергію збудження атомів газу і по тому — в енергію світлового випромінювання, а в електричному двигуні майже вся енергія струму перетворюється на кінетичну енергію обертання ротора. При цьому в усіх випадках та чи інша частка енергії струму переходить у внутрішню (теплову) енергію провідників. Зокрема, в нерухомих рідких і твердих провідниках із постійним струмом на тепло перетворюється вся енергія струму.

Кількісною характеристикою перетворень енергії струму на тепло в кожній точці провідника є питома потужність струму (Вт/м^3), яка дорівнює енергії, що передається провіднику носіями в одиниці об'єму за одиницю часу

$$w = \frac{\mathrm{d}W}{\mathrm{d}t\,\mathrm{d}V}. \tag{1.11}$$

Величину w можна визначити, прямо підрахувавши енергію $\mathrm{d}W$, що виділяється в об'ємі провідника $\mathrm{d}V$ за час $\mathrm{d}t$, як

$$\mathrm{d}W = \langle K_1 \rangle \mathrm{d}N\,\mathrm{d}Z_1,$$

де $\langle K_1 \rangle$ — середня величина кінетичної енергії дрейфового руху одного електрона на момент зіткнення з іоном металу, $\mathrm{d}N$ — кількість електронів провідності в об'ємі $\mathrm{d}V$, і $\mathrm{d}Z_1$ — кількість зіткнень з іонами, які відбуває один електрон за час $\mathrm{d}t$.

Але зручніше у виразі (6.12) величину $\mathrm{d}W/\mathrm{d}V$ розглядати як роботу, а $\mathrm{d}W/(\mathrm{d}t\cdot\mathrm{d}V)$ — як потужність сил, що діють на всі носії в одиниці об'єму провідника. Тоді, враховуючи відому з механіки формулу $p = Fu$ для потужності сили F, що діє на частинку при швидкості руху u, можна записати:

$$w = np = nuF,$$

де n — кількість носіїв в одиниці об'єму. Відтак, урахувавши вираз (1.3), отримаємо:

$$w = j\frac{F}{e}. \tag{1.12}$$

Формула (1.12) не пов'язана з природою сил, які діють на носії струму, тож є найбільш загальним виразом питомої потужності струму. Але зазвичай $F = eE$, де E — напруженість електричного поля в провіднику із струмом. Тоді

$$w = jE, \tag{1.13}$$

а з урахуванням закону Ома (1.9)

$$w = \rho j^2, \tag{1.13а}$$

або

$$w = \sigma E^2 = \frac{E^2}{\rho}. \tag{1.13б}$$

У випадку постійного струму всі ці формули є рівносильними і виражають закон Джоуля в локальній ("диференціальній") формі:

> ***кількість теплоти, що виділяється за одиницю часу (теплова потужність) в одиниці об'єму однорідного провідника із струмом, в кожній точці є прямо пропорційною квадратові густини струму або квадратові напруженості електричного поля в цій точці.***

1.5. *Обмеженість класичної теорії електропровідності*

Отримані в класичній теорії висновки про лінійну залежність густини струму від напруженості електричного поля (закон Ома) та квадратичну залежність потужності струму від напруженості електричного поля (закон Джоуля) в однорідних лінійних провідниках підтверджуються на досліді. Але низка її кількісних передбачень не справджується, тож класична теорія електропровідності металів є неповною. А саме.

Відомо, що концентрація вільних електронів у міді $n = 8{,}5 \cdot 10^{28}\text{м}^{-3}$, і питомий опір $\rho = 1{,}7 \cdot 10^{-8}$ Ом · м при $t = 20^\circ$ С. Якщо підставити ці величини, отриману раніше величину теплової швидкості електронів $v \approx 10^5$ м/с, та заряд і масу електрона у формулу (1.10а), то для середньої довжини вільного пробігу електрона виходить $\langle\lambda\rangle \approx 5{\cdot}10^{-9}$ м. Позаяк відстань між атомами в міді $d = 3{,}6 \cdot 10^{-10}$ м, то виходить, що електрон без зіткнень проходить у середньому півтора десятка іонів, які трапляються на його шляху. В класичній концепції електронного газу цього не можна збагнути, адже атоми в твердому тілі розміщені дуже щільно, практично впритул. Істотні розбіжності між класичною теорією та дослідом існують і щодо температурної залежності опору металевих провідників $\rho(T)$. Згідно з формулами (1.10а) і (1.7) питомий опір має бути прямо пропорційним величині $\sqrt{T}$, але насправді в широкому температурному інтервалі залежність $\rho(T)$ є лінійною. Крім того класична теорія взагалі не здатна пояснити існування явища надпровідності — повного зникнення електричного опору в деяких металах і сплавах при наднизьких температурах.

Обмеженість класичної теорії електропровідності пояснюється тим, що електрони в металі не можна розглядати як матеріальні точки, підпорядковані законам класичної фізики. Як з'ясувалося, електрони мають складні *квантові* властивості, і їхній рух у твердому тілі радше нагадує не рух кульок, а поширення хвиль. При цьому виявляється, що перешкоду упорядкованому рухові електронів у провіднику, тобто опір, створюють *не самі іони, а дефекти кристалічної структури* — різні порушення строгої періодичності в розташуванні іонів у вузлах кристалічної ґратки. Такими можуть бути відсутність іонів у деяких вузлах ґратки, їх заміщення чужорідними атомами, розташування власних чи чужорідних атомів у міжвузіллях, з рештою, і тепловий рух атомів. В оптиці подібні дефекти середовища є причиною розсіяння світла (розсіяння фотонів). Відповідно, взаємодію носіїв із дефектами структури провідника називають розсіюванням носіїв струму.

Кількість дефектів у провіднику є набагато меншою за кількість власних атомів, а відстань між ними, відповідно, більшою. Це пояснює аномально велику як для класичної теорії, довжину вільного пробігу носіїв струму. Квантова теорія правильно пояснює й інші властивості провідників, зокрема передбачає лінійну залежність опору від температури і пояснює явище надпровідності. В цьому зв'язку те, що класична теорія при принципово хибних засадах дає правильні основні закони струму (1.9) і (1.13) може викликати подив. Але насправді нічого дивного в цьому немає, позаяк залежності (1.9) і (1.13) є наслідком лише того, що при заданій напруженості поля в провіднику встановлюється постійна швидкість дрейфу носіїв. А це пов'язано тільки із самим фактом розсіювання носіїв, а не з його механізмом.

Контрольні запитання

1. У рухомій металевій кульці її вільні електрони, що є носіями струму, теж напрямлено рухаються із швидкістю кульки. Поясніть, чи створює електричний струм цей рух.

2. Як пов'язаний напрям струму із напрямком руху носіїв, які його створюють?

3. По довгій прямій дротині протікає струм. Чи виявить струм у дротини спостерігач, який рухається вздовж неї із швидкістю дрейфу електронів? Якщо так, то якими будуть його величина та напрям?

4. Поясніть, яку властивість електричного заряду відображує рівняння неперервності.

5. Чи можна створити тривалий електричний струм у незамкненому провіднику? Якщо ні, то чому? Якщо так, то за яких умов?

6. Говорять, що струм у металі створюється напрямленим рухом *вільних* електронів. Якщо так, то чому швидкість цього руху є дуже малою?

7. В яких випадках закон Ома не виконується? Які провідники називаються лінійними і що є умовою лінійності провідника?

8. У чому полягає закон Джоуля в локальній формі? Чи завжди вся робота струму переходить у тепло?

9. У чому виявляється неповнота класичної теорії електропровідності металів та якою є її причина?

2. Електричне коло

Розглянуті локальні закони електричного струму дозволяють встановити зв'язки між характеристиками струму та електричного поля не лише в кожній точці провідника, а й на кожній ділянці електричного кола, або ланцюга, тобто з'єднаних між собою провідників, приладів і пристроїв, які утворюють систему провідних контурів для струму. Далі розглядаються найпростіші нерозгалужені кола, котрі включають тільки один контур.

Розглянуті такі питання:

2.1. Сторонні сили. ЕРС

2.2. Узагальнений закон Ома

2.3. Електричний опір

2.4. З'єднання джерел струму

2.5. Розгалужене коло

Контрольні запитання

2.1. *Сторонні сили. ЕРС*

Сторонні сили. Раніше говорилося, що для протікання стаціонарного електричного струму потрібний замкнений провідний контур. Але одного цього замало, бо крім носіїв струму в провіднику є й протилежно заряджені частинки, що своїм електричним полем перешкоджають вільному рухові носіїв по контуру. Тому для підтримування струму в провіднику на носії мають діяти якісь інші сили, що діють проти сил указаного електричного поля і здатні переміщувати носії вздовж провідника. Зрозуміло, що такі сили не можуть створюватися зарядами самого провідника і мають зовнішнє походження. Тому вони називаються сторонніми силами.

Через наявність у провіднику із струмом сторонніх сил у виразах закону Ома для густини струму (1.10) і (1.10а) має стояти сумарна напруженість поля кулонівських та сторонніх сил:

$$\vec{j} = \sigma\left(\vec{E}_{\text{к}} + \vec{E}_{\text{ст}}\right). \qquad (2.1)$$

Теоретично сторонні сили можна створити у різний спосіб. Зазвичай для цього в контур включають спеціальні пристрої — джерела струму. Дія джерел струму ґрунтується теж на різних принципах, але в приладах малої потужності найчастіше використовують хімічні джерела — "батарейки" і акумулятори.

Розглянемо, як діє найпростіше коло, що складається із внутрішньої ділянки (джерела) і зовнішньої ділянки у вигляді провідника, котрий через ключ К приєднаний до полюсів (вивідних контактів) джерела (рис. 2.1).

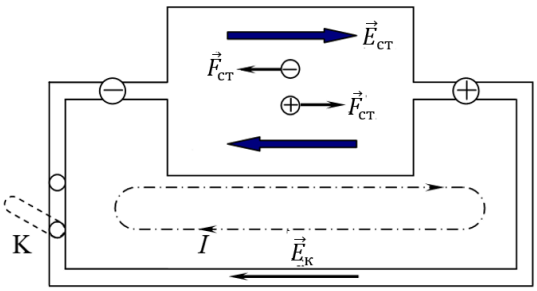


Рис. 2.1

Всередині джерела під дією сторонніх сил певна кількість електронів (носіїв струму) переходить з одного полюса на інший, створюючи між ними відповідну різницю потенціалів і кулонівське електричне поле $\vec{E}_{\text{к}}$ всередині джерела та на зовнішній ділянці. При розімкненому ключі це поле не спричинює струму, бо всередині джерела воно компенсується полем сторонніх сил, а назовні між полюсами для носіїв немає неперервного провідного шляху. Але після замикання ключа надлишкові електрони з негативного полюса під дією кулонівського поля починають рухатися по зовнішній ділянці і з'єднуватися з надлишковими зарядами на позитивному полюсі. Якби в джерелі на електрони не діяли сторонні сили, це призвело би до практично миттєвої нейтралізації полюсів і припинення руху носіїв по колу. Та насправді електрони, що приходять по зовнішній ділянці кола на позитивний полюс, підхоплюються сторонніми силами і, всупереч дії кулонівського поля всередині джерела, знову переміщуються на негативний полюс. Тож, доки діють сторонні сили, в колі зберігається електричне поле і напрямлений рух (дрейф) носіїв, тобто — електричний

струм. Можна сказати, що джерело струму "переганяє" носії струму від одного полюса до іншого, долаючи кулонівське поле всередині джерела, подібно до насоса, котрий переганяє воду від вхідного до вихідного отвору, долаючи опір рухові води в трубопроводі.

Спад напруги та ЕРС. Через розсіювання носіїв у провіднику, вони втрачають енергію впорядкованого руху. Тому для підтримання струму ці втрати мають компенсуватися за рахунок роботи сил, які забезпечують струм. На довільній ділянці кола (рис. 2.2) це можуть бути і кулонівські $\vec{F}_{\text{к}}$, і сторонні $\vec{F}_{\text{ст}}$ сили.

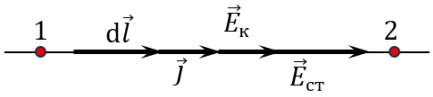


Рис. 2.2

Отже, при елементарному переміщенні $\mathrm{d}\vec{l}$ уздовж провідника над носіями із загальним зарядом q виконується робота

$$\begin{aligned}\delta A = \delta A_{\text{к}} + \delta A_{\text{ст}} = \vec{F}_{\text{к}}\,\mathrm{d}\vec{l} + \vec{F}_{\text{ст}}\,\mathrm{d}\vec{l} = \\ = q\vec{E}_{\text{к}}\,\mathrm{d}\vec{l} + q\vec{E}_{\text{ст}}\,\mathrm{d}\vec{l}.\end{aligned}$$

Для всієї ділянки між точками 1 і 2 маємо:

$$\begin{aligned}A = A_{\text{к}} + A_{\text{ст}} = \\ = q\int\limits_{1\to 2}\vec{E}_{\text{к}}\,\mathrm{d}\vec{l} + q\int\limits_{1\to 2}\vec{E}_{\text{ст}}\,\mathrm{d}\vec{l}.\end{aligned} \tag{2.2}$$

де інтегрування ведеться вздовж провідника.

Робота A є прямо пропорційною до перенесеного носіями заряду q, тож відношення A/q не залежить від q і є характеристикою сумарного силового поля в провіднику. Воно позначається символом U і називається спадом напруги на заданій ділянці кола:

$$U = \frac{A}{q}. \tag{2.3}$$

На елементарній ділянці спад напруги

$$\mathrm{d}U = \frac{\delta A}{q}. \tag{2.3а}$$

Таким чином,

> ***спадом напруги на заданій ділянці кола називається відношення роботи всіх сил, які діють на носії струму, до перенесеного по цій ділянці заряду.***

Відповідно до даного означення і виразу (2.2),

$$U = \int\limits_{1\to 2}\vec{E}_{\text{к}}\,\mathrm{d}\vec{l} + \int\limits_{1\to 2}\vec{E}_{\text{ст}}\,\mathrm{d}\vec{l}.$$

Кулонівське поле є потенціальним (див. розділ I, п. 2), тож перший доданок дорівнює різниці потенціалів на кінцях ділянки:

$$\int\limits_{1\to 2}\vec{E}\,\mathrm{d}\vec{l} = \varphi_1 - \varphi_2. \tag{2.4}$$

Другий доданок характеризує поле сторонніх сил. Він позначається символом $\mathscr{E}$ і називається електрорушійною силою (ЕРС) на даній ділянці кола:

$$\mathscr{E} = \frac{A_{\text{ст}}}{q} = \int\limits_{1\to 2}\vec{E}_{\text{ст}}\,\mathrm{d}\vec{l}. \tag{2.5}$$

Отже,

> ***електрорушійною силою на даній ділянці кола називається відношення роботи сторонніх сил, які діють на носії струму, до перенесеного по цій ділянці заряду.***

Ураховуючи означення (2.3) і (2.5), можна записати наступний загальний вираз для спаду напруги на ділянці кола:

$$U = \varphi_1 - \varphi_2 + \mathscr{E}. \tag{2.6}$$

При цьому на однорідній ділянці, тобто такій, де відсутні сторонні сили, $\mathscr{E}$ = 0, і спад напруги дорівнює різниці потенціалів на кінцях:

$$U = \varphi_1 - \varphi_2. \tag{2.6а}$$

Це означає, що на зовнішній ділянці кола робота по переміщенню зарядів дорівнює

$$A = q(\varphi_1 - \varphi_2)$$

і виконується тільки кулонівським електричним полем. Всередині джерела струму, де $\mathscr{E} \neq 0$, вказана робота виконується обома силовими полями і дорівнює

$$A = qU = q(\varphi_1 - \varphi_2 + \mathscr{E}).$$

А ось при обході замкненого контуру $\varphi_1 = \varphi_2$, і спад напруги у всьому контурі, тобто сума спадів напруги на всіх ділянках,

$$U = \mathscr{E}. \tag{2.6б}$$

Отже, відповідно до (2.3), повна робота, що виконується над носіями струму в замкненому контурі дорівнює роботі тільки сторонніх сил:

$$A = q\mathscr{E}. \tag{2.7}$$

Це пояснюється тим, що через потенціальність кулонівського поля сума його робіт на зовнішній $A'_{\text{к}} = q(\varphi_1 - \varphi_2)$ та на внутрішній $A''_{\text{к}} = q(\varphi_2 - \varphi_1)$ ділянках замкненого контуру дорівнює нулю. Тож одні лише кулонівські сили принципово не можуть забезпечити електричний струм. Відповідно,

> ***електричний струм в колі існує виключно завдяки енергії, що надається носіям джерелом струму.***

Тому роль сторонніх можуть виконувати тільки ті сили, робота і циркуляція поля котрих на даному провідному контурі є відмінною від нуля:

$$\oint \vec{E}_{\text{ст}}\, \mathrm{d}\vec{l} \neq 0.$$

При цьому фізична природа та конкретне походження сторонніх сил не має принципового значення.

Що ж до кулонівського поля, то воно, хоч і не дає енергії для протікання струму, виконує в контурі важливу "диспетчерську" функцію — переносить енергію від джерела до інших ділянок.

Наостанок зробимо ще дві ремарки. По-перше, з означень (2.5) і (2.6) випливає, що ЕРС та спад напруги, як і різниця потенціалів, вимірюються у вольтах (В). І, по-друге, всі згадані величини є алгебраїчними; їхні знаки визначаються напрямками кулонівського поля та поля сторонніх сил відносно напрямку переміщення від початкової до кінцевої точки заданої ділянки кола. Детальніше про це говориться в п. 2.5.

2.2. *Узагальнений закон Ома*

Вище говорилося, що для практики важливо знати зв'язок між струмом і полем не лише в точці, а й на окремих ділянках і в усьому колі в цілому. Цей зв'язок визначається відомим законом Ома, який був установлений дослідним шляхом. Але, виходячи із співвідношень (2.2) і (2.3), указаний закон можна отримати й теоретично.

Розглянемо струм у тонкому провіднику, в якому струм рівномірно розподіляється по площі перерізу, так що густина струму

$$j = \frac{I}{S}, \tag{2.8}$$

і вектори густини струму $\vec{j}$ та полів $\vec{E}_{\text{к}}$, $\vec{E}_{\text{ст}}$ і $\vec{E} = \vec{E}_{\text{к}} + \vec{E}_{\text{ст}}$ спрямовані вздовж провідника.

> *(*Зауважимо, що в теорії термін "тонкий" означає не величину перерізу, а рівномірність розподілу по ньому струму, і на практиці цей критерій задовольняють всі електричні кола.*)*

На малій ділянці $\mathrm{d}l$ провідника елементарна робота δA, яку виконують над носіями *всі сили*, дорівнює

$$\delta A = qE\,\mathrm{d}l,$$

або, відповідно до закону Ома (1.9а) і виразу (2.8),

$$\delta A = q\rho l\frac{\mathrm{d}l}{S}.$$

З іншого боку, згідно з означенням (2.3а),

$$\delta A = q\,\mathrm{d}U,$$

де $\mathrm{d}U$ — спад напруги на ділянці $\mathrm{d}l$. Отже,

$$\mathrm{d}U = \rho l\frac{\mathrm{d}l}{S}.$$

Проінтегрувавши цей вираз по всій довжині провідника з урахуванням того, що для стаціонарного струму $I = \mathrm{const}$ по всій довжині, отримаємо:

$$U = \int\limits_l I\rho\frac{\mathrm{d}l}{S} = I\int\limits_l \rho\frac{\mathrm{d}l}{S}. \tag{2.9}$$

У цьому виразі величина

$$R = \int\limits_l \rho\frac{\mathrm{d}l}{S} \tag{2.10}$$

залежить тільки від розмірів і конфігурації (форми) провідника та властивостей його речовини і називається електричним опором.

Таким чином, струм і спад напруги на довільній ділянці кола пов'язані співвідношеннями

$$U = IR \tag{2.11}$$

або

$$I = \frac{U}{R} = \frac{\varphi_1 - \varphi_2 + \mathscr{E}}{R}, \tag{2.11а}$$

які виражають у**загальнений закон Ома:**

> ***сила струму в довільній ділянці електричного кола дорівнює відношенню спаду напруги на цій ділянці до її опору.***

Закон Ома в такій редакції називається *узагальненим*, оскільки вираз (2.11а) є чинним як для неоднорідної ділянки кола, де на носії діють і кулонівські, і сторонні сили, так і для однорідної ділянки, де діють лише кулонівські сили. В останньому випадку в (2.11а) $\mathscr{E} = 0$, і

$$I = \frac{\varphi_1 - \varphi_2}{R}. \tag{2.12}$$

У підручниках з елементарної фізики цей вираз називають *законом Ома для ділянки кола*:

> ***сила струму в (однорідній) ділянці кола дорівнює відношенню різниці потенціалів на кінцях ділянки до її опору.***

Відповідно, при обході всього замкненого контуру $\varphi_1 = \varphi_2$ і

$$I = \frac{\mathscr{E}}{R_0},$$

або

$$I = \frac{\mathscr{E}}{R + r}, \tag{2.13}$$

де $R_0 = R + r$ — опір всього кола, R — опір зовнішньої ділянки кола, або навантаження, і r — внутрішній опір джерела струму (рис. 2.3).

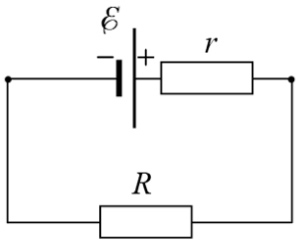


Рис. 2.3

Цю формулу називають законом Ома для повного кола:

> ***сила струму в замкненому колі дорівнює відношенню ЕРС, яка діє в ньому, до повного опору кола.***

Формулу (2.13) можна переписати у вигляді

$$IR + Ir = \mathscr{E}. \tag{2.13а}$$

Величини IR та Ir визначають спади напруги на однорідній зовнішній і неоднорідній внутрішній ділянці кола, відповідно. При цьому згідно з (2.12) $IR = \varphi_1 - \varphi_2$, тож

$$\varphi_1 - \varphi_2 = \mathscr{E} - Ir. \tag{2.14}$$

Отже, різниця потенціалів на полюсах джерела не збігається з ЕРС. Так само не збігаються і спади напруги на цих двох ділянках, попри те, що вони мають спільні крайні точки — полюси джерела.

Зауважимо також, що, тоді як співвідношення (2.11а) виконується для будь-якої ділянки будь-якого кола, формула (2.13) є чинною тільки для найпростішого кола (рис. 2.3), що містить тільки один провідний контур.

2.3. *Електричний опір*

Опір однорідного провідника. Електричний опір спричинюється розсіюванням носіїв у провіднику із струмом і в загальному випадку визначається інтегральною формулою (2.10). Але на практиці зазвичай використовують однорідні провідники ($\rho = \text{const}$) сталого поперечного перерізу ($S = \text{const}$) по всій довжині l. В такому разі

$$R = \frac{\rho}{S}\int\limits_l \mathrm{d}l \quad \Rightarrow \quad R = \rho\frac{l}{S}. \tag{2.15}$$

У деяких випадках замість опору зручніше використовувати обернену величину $\Lambda = (1/R)$, яка називається провідністю і визначається формулою

$$\Lambda = \sigma\frac{S}{l}, \tag{2.15а}$$

де $\sigma = (1/\rho)$ — питома електропровідність речовини провідника.

Електричний опір вимірюється в омах (Ом), а провідність — у сіменсах (См), $1\ \text{См} = 1\ \text{Ом}^{-1}$.

Опір провідника залежить не лише від його довжини та перерізу, а й від температури. Як показує дослід, при нагріванні опір металів зростає і в доволі широкому діапазоні температур виражається формулою:

$$R = R_0(1 + \alpha t), \tag{2.16}$$

де t — температура за шкалою Цельсія, R_0 — опір провідника при температурі 0°С, і α (1/К) — таблична константа, що називається температурним коефіцієнтом опору і показує відносну зміну опору провідника при зміні температури на 1 К.

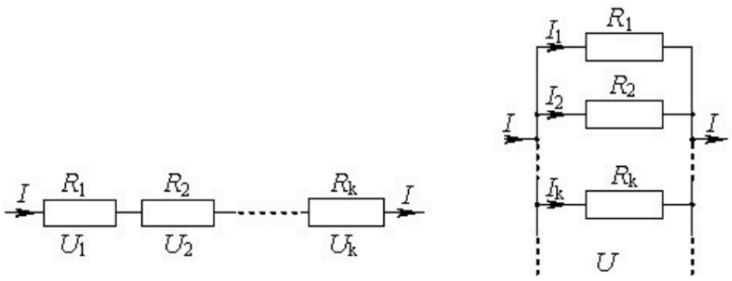
Зростання опору металевих провідників при нагріванні пояснюється зростанням швидкості теплового руху v і, згідно з (1.7), зменшенням часу вільного пробігу носіїв $\langle\tau\rangle$. Через це збільшується питомий опір речовини ρ (1.10а), тож і опір провідника R. В інших провідних середовищах — напівпровідниках і провідниках другого роду (електролітах) — опір теж залежить від температури, але на відміну від металів, при нагріванні зменшується. Це пояснюється тим, що в указаних речовинах при підвищенні температури суттєво зростає концентрація носіїв струму, так що зменшення їхнього часу вільного пробігу перестає бути визначальним.

З'єднання резисторів. Електричний опір є робочим параметром резисторів — найпростіших приладів, які присутні практично в усіх електричних схемах. Резистор по суті є просто зразком провідної речовини із заданим опором R. Для резисторів використовують матеріали з малою електропровідністю, що дозволяє отримувати великі значення опору при малих розмірах резисторів.

В електричних і електронних схемах використовують як поодинокі резистори, так і різні з'єднання (комбінації) резисторів. Але при розрахунках часто буває зручно все з'єднання замінити одним *еквівалентним* резистором із опором, при якому струми в усіх інших ділянках кола лишаються незмінними. Тож опір еквівалентного резистора має бути таким, що напруга і струм у ньому збігаються з напругою та струмом на всьому з'єднанні.

Встановимо формули для розрахунку еквівалентного опору найпростіших з'єднань резисторів — послідовного та паралельного.

Послідовне з'єднання. В такому з'єднанні всі резистори сполучені в ланцюг так, як показано на рис. 2.4*а*.



а) б)

Рис. 2.4

При цьому струм у всіх резисторах однаковий, а напруга на ланцюзі дорівнює сумі напруг на всіх з'єднаних резисторах:

$$I_k = I, \qquad U = \sum_k U_k.$$

Відповідно, опір з'єднання

$$R_{\text{пос}} = \frac{U}{I} = \frac{1}{I}\sum_k U_k = \sum_k \frac{U_k}{I_k},$$

тобто

$$R_{\text{пос}} = \sum_k R_k. \tag{2.17}$$

Отже,

> ***еквівалентний опір послідовного з'єднання дорівнює сумі опорів усіх резисторів, які з'єднані.***

Зокрема, коли послідовно з'єднані n однакових резисторів із опором R_1 кожен, то опір з'єднання

$$R_{\text{пос}} = nR_1. \tag{2.17а}$$

Паралельне з'єднання. Схема такого з'єднання показана на рис. 2.4*б*. Очевидно, що в цьому випадку напруга на всіх резисторах однакова і дорівнює напрузі на всьому з'єднанні, а величина струму, що проходить через з'єднання, згідно із законом збереження заряду, дорівнює сумі струмів у кожному із з'єднаних резисторів:

$$U_k = U, \qquad I = \sum_k I_k.$$

Тому для опору з'єднання маємо:

$$R_{\text{пар}} = \frac{U}{I} = \frac{U}{\sum\limits_k I_k} \quad \Rightarrow$$

$$\Rightarrow \quad \frac{1}{R_{\text{пар}}} = \frac{\sum\limits_k I_k}{U} = \sum_k \frac{I_k}{U_k}.$$

Звідси випливають такі розрахункові формули:

$$\frac{1}{R_{\text{пар}}} = \sum_k \frac{1}{R_k} \tag{2.18}$$

або

$$\Lambda_{\text{пар}} = \sum_k \Lambda_k. \tag{2.18а}$$

Отже,

еквівалентна провідність паралельного з'єднання резисторів дорівнює сумі провідностей всіх з'єднаних резисторів.

Якщо паралельно з'єднані n однакових резисторів опором R_1 кожен, то згідно з формулою (2.18)

$$R_{\text{пар}} = \frac{R_1}{n}. \tag{2.19а}$$

Для двох паралельно з'єднаних двох резисторів R_1 і R_2 також можна записати просту пряму формулу розрахунку еквівалентного опору:

$$R_{\text{пар}} = \frac{R_1 R_2}{R_1 + R_2}. \tag{2.19б}$$

Отримані формули дозволяють розраховувати й опори більш складних комбінованих з'єднань резисторів, які можна"розбити" на окремі послідовні та паралельні ланцюжки, що в свою чергу, з'єднані між собою послідовно та паралельно.

2.4. *З'єднання джерел струму*

Хімічне джерело струму має невелику ЕРС і запас енергії (енергоємність). Тому на практиці крім поодиноких елементів живлення використовують батареї, що складаються з декількох з'єднаних між собою елементів. Параметри батареї — ЕРС $\mathscr{E}$ і внутрішній опір r — так само, як і опір з'єднань резисторів, можна розрахувати на основі закону збереження заряду та закону Ома.

Послідовне з'єднання джерел (рис. 2.5).

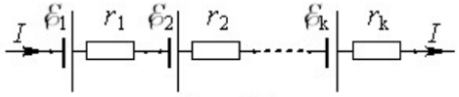


Рис. 2.5

В цьому випадку величина струму в усіх з'єднаних джерелах (E_k, r_k) однакова:

$$I_k = I,$$

а різниці потенціалів на полюсах кожного джерела, згідно з (2.14), дорівнюють

$$\varphi_{1k} - \varphi_{2k} = \mathscr{E}_k - Ir_k.$$

Додавши всі ці вирази, одержимо різницю потенціалів на полюсах батареї:

$$\varphi_1 - \varphi_2 = \sum \mathscr{E}_k - I \sum r_k.$$

З іншого боку, для еквівалентного джерела

$$\varphi_1 - \varphi_2 = \mathscr{E} - Ir.$$

Отже, доходимо висновку, що ЕРС $\mathscr{E}$ і внутрішній опір r батареї *послідовно з'єднаних* елементів визначаються формулами

$$\mathscr{E} = \sum \mathscr{E}_k, \qquad r = \sum r_k. \tag{2.20}$$

Зокрема, при узгодженому з'єднанні n однакових елементів, коли знаки всіх $\mathscr{E}_i$ однакові,

$$\mathscr{E} = n\mathscr{E}_1, \qquad r = nr_1. \tag{2.20а}$$

У такий спосіб отримують джерела живлення з необхідною величиною ЕРС.

Із приводу з'єднання джерел слід зауважити наступне. На практиці при використанні з'єднань (батарей) джерела завжди сполучають узгоджено. Але в навчальних завданнях свідомо використовуються і неузгодженні з'єднання. Тому при розв'язуванні задач треба пам'ятати, що електрорушійна сила є алгебраїчною величиною, і враховувати не тільки її модуль, а й знак.

Паралельне з'єднання джерел (рис. 2.6).

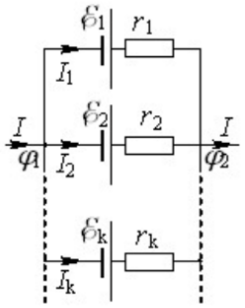


Рис. 2.6

При паралельному з'єднанні різниці потенціалів на полюсах кожного джерела і всієї батареї однакові, а величина струму через батарею відповідно до закону збереження заряду дорівнює сумі струмів у всіх джерелах:

$$\varphi_{1k} - \varphi_{2k} = \varphi_1 - \varphi_2, \qquad I = \sum I_k.$$

Згідно із законом Ома (2.11а), для струмів через окремі джерела можна записати:

$$I_k = \frac{1}{r_k}(\varphi_1 - \varphi_2) + \frac{\mathscr{E}_k}{r_k}.$$

Додавши всі струми I_k, отримаємо струм через батарею:

$$I = (\varphi_1 - \varphi_2)\sum \frac{1}{r_k} + \sum \frac{\mathscr{E}_k}{r_k}.$$

Аналогічно виражається цей струм і через ЕРС $\mathscr{E}$ та внутрішній опір r еквівалентного джерела:

$$I = \frac{1}{r}(\varphi_1 - \varphi_2) + \frac{\mathscr{E}}{r}.$$

Порівнюючи обидва вирази, бачимо, що

при паралельному з'єднанні джерел внутрішній опір r і ЕРС $\mathscr{E}$ батареї визначаються формулами:

$$\frac{1}{r} = \sum \frac{1}{r_k}, \qquad \mathscr{E} = r\sum \frac{\mathscr{E}_k}{r_k}. \tag{2.21}$$

При узгодженому з'єднанні n однакових елементів ($\mathscr{E}_1$, r_1) маємо:

$$\mathscr{E} = \mathscr{E}_1, \qquad r = \frac{r_1}{n}. \tag{2.21а}$$

Тож при паралельному з'єднанні однакових елементів у джерелі живлення його ЕРС не збільшується, але збільшується запас енергії та потужність, яку джерело може віддавати в навантаження.

Таким чином, комбінуючи послідовні та паралельні з'єднання низьковольтних малопотужних елементів, можна отримувати джерела живлення з потрібною величиною ЕРС і великою енергоємністю. Яскравим свідченням цього є акумуляторні батареї для електромобілів, які забезпечують пробіг у сотні кілометрів без дозарядки.

2.5. *Розгалужене коло*

Найпростіше електричне коло, приміром, електрична лампа ввімкнена в мережу, складається з одного провідного контуру, що включає одне джерело струму. Але широко використовуються і більш складні розгалужені кола. Характерною ознакою таких кіл є наявність декількох вузлів, гілок і контурів. Вузлом називається точка, в якій сходиться більше, ніж два провідники. Гілкою (або віткою) називають ділянку кола, що з'єднує два вузли. Відповідно, під контуром розуміють ***сукупність гілок, яка дозволяє носіям рухатися замкненим шляхом, не відгалужуючись***.

Для кожної гілки, звичайно, виконується узагальнений закон Ома (2.11а), але потенціали вузлів визначаються не лише "власним" джерелом, а й джерелами, ввімкненими в усі інші гілки кола. Тому струми в різних частинах одного контуру не однакові, і визначити їх безпосередньо за законом Ома неможливо. Для розгалужених кіл використовують спеціальні методи розрахунку, розроблені в теорії електричних кіл. Із них найпрозорішим є метод розрахунку за допомогою двох правил Кірхґофа, що безпосередньо випливають із основних законів електричного струму.

Перше правило Кірхґофа. Це правило стосується струмів у вузлах розгалуженого кола і є прямим наслідком закону збереження електричного заряду. У випадку постійного струму потенціали вузлів не змінюються з часом, отже, в них не відбувається накопичення зарядів. Це означає, що заряд, який підходить до вузла за одиницю часу, точно дорівнює зарядові, що виходить із нього. Іншими словами, сума струмів, які входить у вузол, дорівнює сумі струмів, які виходять із нього. Цю умову прийнято записувати у вигляді

$$\sum_k I_k = 0, \tag{2.22}$$

розглядаючи величини I_k як алгебраїчні і приписуючи струмам, які виходять з вузла один знак, наприклад додатній, а тим, що входять — протилежний.

Таким чином, перше правило Кірхґофа твердить, що

> ***алгебраїчна сума струмів, які сходяться в довільному вузлі розгалуженого кола, дорівнює нулю.***

Друге правило Кірхґофа. Згідно із узагальненим законом Ома (2.11а), струм у довільній ділянці (гілці) кола із заданим опором визначається різницею потенціалів на її кінцях та електрорушійною силою, що діє на цій ділянці:

$$I_k R_k = \varphi_{1k} - \varphi_{2k} + \mathscr{E}_k.$$

Якщо вибрати в розгалуженому колі якийсь замкнений контур і обходити його в заданому напрямку, то кінцева точка k-ї гілки буде початковою точкою наступної $(k+1)$-ої гілки. Тому при додаванні указаних рівнянь для всіх гілок контуру потенціали скоротяться і вийде:

$$\sum_k I_k R_k = \sum_k \mathscr{E}_k. \tag{2.23}$$

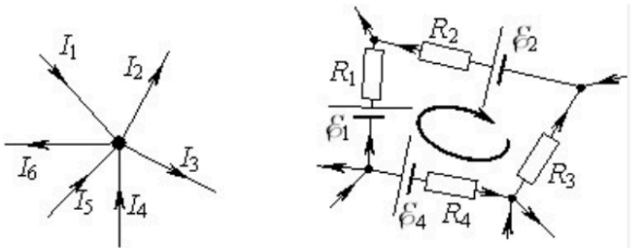
Це рівняння виражає ***друге правило Кірхґофа***, що по суті є узагальненням закону Ома на розгалужене коло. Воно гласить:

> ***алгебраїчна сума добутків сил струму на опори всіх гілок довільного контуру розгалуженого кола дорівнює алгебраїчній сумі ЕРС, які діють у цьому контурі.***

Знаки величин I_k і $\mathscr{E}_k$ є умовними, але мають бути узгодженими. Тому вони визначаються за наступними правилами. Якщо даний струм I_k за напрямом збігається з напрямом обходу контуру, обраним при складанні рівнянь (2.23), то $I_k > 0$. Знак ЕРС збігається зі знаком роботи сторонніх сил у джерелі при переміщенні позитивних носіїв у напрямку обходу контуру. Це означає, що $\mathscr{E}_k > 0$, якщо при обході контуру ми потрапляємо на від'ємний полюс (коротку жирну риску на зображенні джерела на схемі кола). Ці правила ілюструються наступними рівняннями для вузла (рис. 2.7*а*) та контуру (рис. 2.7*б*) якогось розгалуженого кола:

$$I_1 - I_2 - I_3 + I_4 + I_5 - I_6 = 0;$$

$$I_1 R_1 - I_2 R_2 - I_3 R_3 - I_4 R_4 = \\ = \mathscr{E}_1 - \mathscr{E}_2 + \mathscr{E}_4.$$



а) б)

Рис. 2.7

Зауважимо, що величини R_k в рівнянні (2.23) включають і внутрішні опори джерел r_k. Тому для нерозгалуженого кола, котре можна розглядати як контур із двох гілок з опорами R і r та однаковою величиною струму, рівняння (2.23) вироджується в рівняння закону Ома:

$$IR + Ir = \mathscr{E} \quad \Rightarrow \quad I = \frac{\mathscr{E}}{R + r}.$$

У такому контексті закон Ома можна трактувати як окремий випадок другого правила Кірхґофа.

Порядок розрахунку розгалуженого кола. Склавши та розв'язавши систему рівнянь (2.22) і (2.23), можна розрахувати розгалужене коло, тобто визначити величини та напрями всіх струмів у колі через усі наявні в ньому опори R_k та ЕРС $\mathscr{E}_k$. Розрахунок варто проводити в такому порядку.

1. Довільно вибрати й позначити на схемі напрямки струмів I_k у всіх гілках кола;

2. Скласти максимально можливу кількість незалежних рівнянь (2.22) для вузлів, дотримуючись обраного правила знаків.

> Примітки. 1) Максимальна кількість рівнянь для вузлів є не обов'язковою, але доцільною, зважаючи на їхню простоту; 2) Аби рівняння були незалежними, кожен наступний вибраний вузол має включати хоча б одну нову гілку. За такої умови максимальна кількість рівнянь (2.22) дорівнює $N - 1$, де N — повна кількість вузлів кола.

3. Скласти решту необхідних рівнянь за другим правилом Кірхґофа (2.23) для замкнених контурів. При цьому конфігурацію та напрями обходу контурів можна вибирати довільно, але так, аби кожен наступний включав хоча б одну нову гілку.

4. Розв'язати отриману систему рівнянь. При цьому корені I_k системи покажуть не тільки величини, а й напрями всіх струмів. А саме, коли $I_k > 0$, то прийнятий в п. 1 при складанні рівнянь напрям даного струму є правильним. Якщо ж $I_k < 0$, то дійсний напрям даного струму є протилежним до вибраного при складанні рівнянь.

Контрольні запитання

1. Відомо, що на однорідній ділянці кола носії струму рухаються під дією електричного поля, створюваного зарядженими частинками провідника. Тоді для чого потрібні ще й сторонні сили?

2. Що таке сторонні сили? Чи можуть вони мати неелектричну природу?

3. Який фізичний зміст мають різниця потенціалів, електрорушійна сила та спад напруги на заданій ділянці кола?

4. Чи може в нерозгалуженому електричному ланцюзі різниця потенціалів на полюсах джерела струму бути рівною ЕРС?

5. Поясніть природу електричного опору провідників.

6. Поясніть, чому опір провідника зростає при збільшенні його довжини і зменшується при збільшенні площі перерізу.

7. Поясніть, чому опір металів зростає при нагріванні?

8. Покажіть, що при паралельному з'єднанні резисторів опір з'єднання є завжди меншим за найменший з опорів сполучених резисторів.

9. Поясніть, чи може ЕРС послідовного з'єднання елементів живлення бути меншою за суму ЕРС усіх з'єднаних елементів?

10. Поясніть, чи може ЕРС паралельного з'єднання елементів живлення бути: а) меншою і б) більшою за суму ЕРС усіх з'єднаних елементів?

11. Як із однакових елементів з ЕРС 1,5 В можна зібрати батарею з ЕРС 15 В і п'ятикратним запасом енергії?

12. Які характерні ознаки має розгалужене електричне коло? Чи може бути розгалуженим коло, в яке включено всього одне джерело ЕРС?

13. При розрахунку розгалуженого кола напрямки струму в гілках є невідомі. В такому разі, як вони враховуються при складанні рівнянь за правилами Кірхґофа?

14. Як визначаються напрямки струму в гілках розгалуженого кола за результатами розрахунку за правилами Кірхґофа?

15. Чим визначається необхідна кількість рівнянь при розрахунку розгалуженого кола за правилами Кірхґофа?

3. Струм у колі з конденсатором

Стаціонарні (постійні у часі) струми можуть існувати тільки в замкнених провідних контурах. Але з рівняння неперервності (1.6) випливає, що нестаціонарні (залежні від часу) струми можуть існувати і в незамкнених контурах. Про це свідчить широке використання в колах змінного струму конденсаторів — схемних елементів, які створюють розриви провідності у відповідних гілках електричного кола. Нижче на прикладі найпростішого такого кола показана можливість застосування до залежних від часу струмів загальних законів стаціонарного струму, які були отримані вище в пп. 1, 2.

Далі розглядаються такі питання:

3.1. Квазістаціонарні струми

3.2. Струм заряджання конденсатора

3.3. Струм розряджання конденсатора

Контрольні запитання

3.1. *Квазістаціонарні струми*

Те, що *нестаціонарні* (змінні) струми можуть існувати не тільки в замкнених, а й у незамкнених контурах, ставить природнє запитання, чи виконуються для них розглянуті вище закони електричного струму.

Аби отримати відповідь, згадаймо (див. п. 1.3), що час τ і довжина λ вільного пробігу носіїв у провіднику є гранично малими, отже, вони між зіткненнями рухаються в практично постійному полі $\vec{E}$, як і при стаціонарному струмі. Тому в *кожну мить* і *в кожній точці* провідника густина нестаціонарного струму $\vec{j}(t)$ і напруженість поля $\vec{E}(t)$ задовольняють закон Ома в локальній формі (1.9), як і при стаціонарному струмі. Так само лишається чинним і локальний закон Джоуля (1.13).

Але для всього кола чи його ділянки ситуація ускладнюється, позаяк електричне поле поширюється вздовж провідника не миттєво. Через це виникає ефект запізнення: струм у віддалених від джерела точках кола в даний момент часу t визначається напругою, що була створена на вході в якийсь попередній момент $t - \Delta t$. При цьому “час запізнення” Δt визначається швидкістю поширення поля c та відстанню l від джерела до даної точки кола:

$$\Delta t = \frac{l}{c}.$$

Отже, миттєва сила нестаціонарного струму в різних точках кола строго є не однаковою. Тому у виразі (2.9) величину струму I не можна винести з під знаку інтеграла і отримати співвідношення (2.11) і (2.11а). Це означає, що на загал у колах нестаціонарного струму інтегральний закон Ома не виконується. Проте, якщо за час Δt напруга на вході кола не встигає помітно змінитися, то ефект запізнення буде неістотним і закон Ома в кожен момент часу виконуватиметься практично точно для будь-якої ділянки кола. Такі "повільні" струми називаються квазістаціонарними. Тож

> ***для миттєвих значень квазістаціонарних струмів виконуються ті самі основні закони, що й для постійних струмів.***

Умову (критерій) квазістаціонарності можна записати так:

$$\Delta t \ll \tau \qquad \text{або} \qquad \frac{l}{c} \ll \tau, \tag{3.1}$$

де Δt — час запізнення на довжині кола l, c — швидкість поширення поля в провідниках, τ — характерний для даного струму час суттєвої зміни.

Слід зауважити, що позаяк $c \sim 3 \cdot 10^8$ м/с, умова (8.4) не є жорсткою і в теорії електричних кіл має радше теоретичне значення. Наприклад, для промислового струму частота складає $\nu = 50$ Гц, а період $T = 0{,}02$ с. Тому, прийнявши $\tau \sim T$, для граничних розмірів кола L, в якому такий струм є квазістаціонарним, з умови (8.4) отримаємо:

$$L \sim cT = 6 \cdot 10^6 \text{ м}.$$

Аналогічно отримаємо оцінку максимальної частоти, при якій виконується умова квазістаціонарності для різних електронних пристроїв, розміри яких зазвичай не перевищують $l \sim 1$ м:

$$\nu \sim \frac{c}{l} = 300 \text{ МГц},$$

що перекриває весь радіочастотний діапазон, включно з УКХ і FM.

Таким чином, змінні струми в широкому діапазоні частот і розмірів кіл є квазістаціонарними, і для них виконуються ті самі основні закони, що й для постійного струму.

3.2. *Струм зарядки конденсатора*

Як приклад аналізу квазістаціонарних струмів, розглянемо процес заряджання та розряджання конденсатора ємності C що сполучений з резистором R і через ключ може підключатися (чи відключатися) до ідеального ($r = 0$) джерела з ЕРС (рис. 3.1).

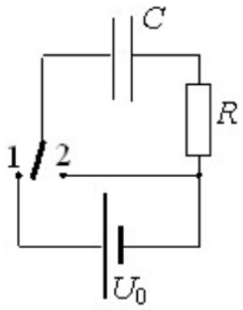


Рис. 3.1

Нехай в момент $t = 0$ ключ переводять у позицію 1, приєднуючи до джерела незаряджений ($q(0) = 0$, $U(0) = 0$) конденсатор C. Після замикання електрони під дією джерела відразу почнуть переходити з однієї обкладки конденсатора на іншу, і через резистор та джерело потече струм, який будемо вважати квазістаціонарним. Тоді для миттєвих значень на основі узагальненого закону Ома (2.11а) можна записати:

$$\begin{gathered} IR = \varphi_- - \varphi_+ + U_0 \quad \Rightarrow \\ \Rightarrow \quad IR + U = U_0. \end{gathered}$$

де $U = \varphi_+ - \varphi_-$ — напруга на обкладках конденсатора. Виражаючи величину U через заряд і ємність конденсатора (ф-ла (2.3), розділ II), отримаємо:

$$IR + \frac{q}{C} = U_0. \tag{3.2}$$

У це рівняння входять дві невідомі величини — струм I у резисторі та заряд конденсатора q. Аби лишити тільки одне, врахуємо, що заряд, який проходить через резистор, дорівнює зміні заряду конденсатора. Отже, миттєва сила струму через резистор $I = (\mathrm{d}q/\mathrm{d}t)$ дорівнює швидкості зміни заряду конденсатора, і рівняння (3.2) набуває вигляду:

$$\begin{gathered} \frac{\mathrm{d}q}{\mathrm{d}t} R + \frac{q}{C} = U_0 \quad \Rightarrow \\ \Rightarrow \quad \frac{\mathrm{d}q}{\mathrm{d}t} = -\frac{q}{\tau} + \frac{U_0}{R}, \end{gathered} \tag{3.2а}$$

де введено позначення

$$\tau = RC. \tag{3.3}$$

Диференціальне рівняння (3.2а) легко розв'язати. Для цього задля зручності введемо нову змінну

$$z = \frac{U_0}{R} - \frac{q}{\tau}, \tag{3.4}$$

і зробимо наступні підстановки в рівняння (3.2а):

$$-\frac{q}{\tau} = z - \frac{U_0}{R}, \qquad \frac{\mathrm{d}q}{\mathrm{d}t} = -\tau \frac{\mathrm{d}z}{\mathrm{d}t}.$$

Відтак рівняння набуває вигляду

$$\tau = \frac{\mathrm{d}z}{\mathrm{d}t} = -z,$$

і його можна переписати як

$$\frac{\mathrm{d}z}{z} = -\frac{\mathrm{d}t}{\tau}. \tag{3.5}$$

Примітка. Таке перетворення обґрунтовується в теорії диференціальних рівнянь і називається відокремленням змінних.

Після цього, інтегруючи праву частину від $t = 0$ до довільного моменту часу t, а ліву — від початкового значення змінної z_0 до значення z в момент t, дістанемо:

$$\begin{gathered}\int\limits_{z_0}^{z} \frac{\mathrm{d}z}{z} = -\frac{1}{\tau}\int\limits_{0}^{t} \mathrm{d}t \quad \Rightarrow \\ \Rightarrow \quad \ln\frac{z}{z_0} = -\frac{t}{\tau} \quad \Rightarrow \quad z = z_0 e^{-\frac{t}{\tau}}.\end{gathered} \tag{3.5а}$$

Зробивши зворотню заміну (3.4) і врахувавши, що $z_0 = (U_0/R)$ і $\tau = RC$, отримаємо остаточний результат:

$$q = q_m(1 - e^{-t/\tau}), \tag{3.6}$$

де величина

$$q_m = CU_0 \tag{3.7}$$

дорівнює максимальному заряду, що може бути накопичений на конденсаторі при заряджанні.

Поділивши q на C, отримаємо залежність від часу напруги на конденсаторі:

$$U = U_0(1 - e^{-t/\tau}), \tag{3.8}$$

де максимальна напруга

$$U_0 = \frac{q_m}{C},$$

природньо, дорівнює ЕРС джерела.

Диференціюючи рівняння (3.6) за часом, знайдемо закон зміни струму заряджання:

$$I = I_0 e^{-t/\tau}, \tag{3.9}$$

де $I_0 = (q_m/\tau) = I(0)$ — струм у резисторі в момент замикання ключа. Якщо сюди підставити вирази (3.3) і (3.7), то вийде

$$I_0 = \frac{U_0}{R}.$$

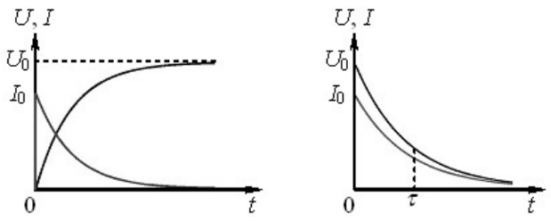
Вираз (3.9) можна отримати й за законом Ома, врахувавши, що напруга на конденсаторі має зустрічну до джерела полярність (“$+$” до “$+$” і “$-$” до “$-$”), отже напруга на опорі

$$U_R = U_0 - U = U_0 e^{-t/\tau},$$

і струм в колі

$$I = \frac{U_R}{R} = \frac{U_0}{R} e^{-t/\tau}.$$

Графіки залежностей $U(t)$ і $I(t)$ показані на рис. 3.2*а*.



а) б)

Рис. 3.2

Впадає в око стрибок струму від 0 до I_0 в момент замикання ключа, немов конденсатора взагалі немає. Але це цілком природньо, адже в першу мить після замикання напруга на конденсаторі $U(0) = 0$, і він ще ніяк не перешкоджає рухові зарядів.

3.3. *Струм розрядки конденсатора*

Якщо після повного заряджання конденсатора перемкнути ключ у позицію 2, то конденсатор почне розряджатися через резистор. При цьому в будь-яку мить напруга (різниця потенціалів) на резисторі та на конденсаторі однакова, тому $IR = q/C$. Сила струму при розряджанні дорівнює швидкості спаду заряду конденсатора: $I = -(\mathrm{d}q/\mathrm{d}t)$, отже,

$$R\frac{\mathrm{d}q}{\mathrm{d}t} = -\frac{q}{C} \quad \Rightarrow \quad \frac{\mathrm{d}q}{\tau}.$$

Відокремивши змінні, як раніше, та проінтегрувавши по тому отримане рівняння, дістанемо:

$$q = q_0 e^{-t/\tau},$$

де $q_0 = CU_0$ — початковий заряд на конденсаторі. Відповідно, напруга на конденсаторі $U = q/C$ і струм розряджання $I = U/R$ змінюються за законом:

$$U = U_0 e^{-t/\tau}, \tag{3.10}$$

$$I = I_0 e^{-t/\tau}, \tag{3.11}$$

де $I_0 = U_0/R$. Графіки залежностей $U(t)$ і $I(t)$ при розряджанні конденсатора показані на рис. 3.2*б*.

Стосовно отриманих тут і в попередньому пункті результатів треба зробити наступні зауваження. Залежності (3.5)—(3.11) є асимптотичними, тому однозначно вказати час повної зарядки чи розрядки конденсатора неможливо. Можна говорити лише про швидкість процесу, мірою якої є параметр τ. Ця величина називається часом релаксації або сталою часу RC-ланцюжка і визначається формулою (3.3).

Час релаксації має прозорий зміст: згідно з формулами (3.10), (3.11), на момент $t = \tau$ напруга на конденсаторі $U(\tau) = U_0/e$ і струм уколі $I(\tau) = I_0/\tau$. Отже,

> ***час релаксації — то є проміжок часу, протягом якого напруга та струм при розряджанні конденсатора зменшуються в*** $e \approx 2{,}7$ ***разів.***

Що ж до повної зарядки або розрядки конденсатора, то теоретично вона взагалі не настає ніколи, бо потребує нескінченно великого часу $t \to \infty$. Але такий висновок є абстрактним. Через хаотичність теплового руху частинок у макроскопічних системах неперервно відбуваються так звані *флуктуації* — неконтрольовані мікроскопічні зміни параметрів системи. У нашому випадку це стосується густини потоку носіїв, а отже сили струму в провідниках і заряду та напруги на конденсаторі. Тому величини q, u, i реально змінюються, лише доки їхня відміна від асимптотичних (граничних) значень не досягне величини флуктуацій. Відтак настає повна розрядка (або зарядка) конденсатора, при якій величини q, u, i будуть здійснювати тільки мікроскопічні флуктуації навколо граничних значень. Причому, такий стан настає достатньо швидко — за час у декілька τ.

При наведених розрахунках уважалося, що струми зарядки і розрядки конденсатора є квазістаціонарними. З'ясуємо, наскільки таке припущення є обґрунтованим. В RC-ланцюжках потужних електричних і електронних кіл використовують конденсатори з ємністю щонайменше в декілька мікрофарад і резистори в десятки й більше омів. За таких умов час релаксації $\tau = RC$, що є мірою суттєвої зміни струму, складає $\geqslant 10^{-5}$ с. Розміри таких ланцюжків можуть досягати $l \approx 10$ см, що відповідає часові запізнення $\Delta t = (l/c) \approx 10^{-10}$ с. Отже, $\Delta t/\tau \approx 10^{-4}$ і критерій квазістаціонарності (3.1) виконується з великим запасом. Тим більше в мікроелектроніці, де розміри схемних елементів і часи запізнення є на порядки менші.

Контрольні запитання

1. Чи є можливим електричний струм у незамкненому контурі? Якщо ні, то чому? Якщо так, то коли?

2. У чому полягає відміна між нестаціонарним і стаціонарним струмом? Нестаціонарним і квазістаціонарним? Сформулюйте і запишіть критерій квазістаціонарності струму.

3. Чи виконуються закони Ома (1.9а) та (2.11а) для нестаціонарного та квазістаціонарного струму? Чому?

4. Чому, коли підключити до джерела конденсатор, напруга на ньому встановлюється не одразу, а поступово?

5. Від чого залежить швидкість заряджання конденсатора? Що таке час релаксації (стала часу) ланцюжка з конденсатором і що показує його числове значення?

6. Як змінюється з часом струм у колі з конденсатором а) при підключенні і б) відключенні від джерела напруги?

7. У скільки разів змінюється струм у колі з конденсатором за проміжок часу 2τ від $t_1 = 0$ до $t_2 = 2\tau$ та від $t_3 = 4\tau$ до $t_4 = 6\tau$?

8. Яка частка (%) енергії ідеального джерела ($r = 0$) йде на зарядку підключеного до нього через резистор R конденсатора? Як пояснити те, що вона не залежить від величини R?

4. Робота і потужність струму

Сили, що забезпечують дрейфовий рух носіїв у провіднику із струмом, виконують над ними роботу, яка називається *роботою електричного струму*. Ця робота виконується за рахунок енергії джерела струму і є мірою її перетворення в інші форми — теплову, механічну, світлову, тощо. Згідно з означенням (2.3), робота постійного струму ($I = \text{const}$) на довільній ділянці кола визначається добутком перенесеного заряду на спад напруги: $A = qU$. При цьому заряд $q = It$, отож

$$A = IUt. \tag{4.1}$$

З урахуванням закону Ома (2.11) роботу постійного струму можна також виразити формулами:

$$A = I^2Rt \tag{4.1а}$$

і

$$A = \frac{U^2}{R}t. \tag{4.1б}$$

У випадку квазістаціонарного струму спад напруги та сила струму змінюються з часом: $u = u(t)$ і $i = i(t)$. Тому за наведеними формулами можна виразити тільки елементарну роботу δA за нескінченно малий проміжок часу $\mathrm{d}t$, а повна робота A за заданий час t визначається інтегруванням:

$$\begin{aligned} A = \int \delta A \quad \Rightarrow \quad A = \int iu\,\mathrm{d}t = \\ = \int i^2R\,\mathrm{d}t = \int \frac{u^2}{R}\,\mathrm{d}t, \end{aligned} \tag{4.2}$$

де i, u — миттєві значення струму та напруги.

Потужність струму P, яка визначає роботу струму за одиницю часу, для постійного струму, відповідно до виразів (4.1), обчислюється за якоюсь із формул:

$$P = IU, \tag{4.3}$$

$$P = I^2R, \tag{4.3а}$$

або

$$P = \frac{U^2}{R}. \tag{4.3б}$$

У випадку квазістаціонарного струму при підстановці в ці формули миттєвих значень струму та спаду напруги $u = u(t)$ і $i = i(t)$ отримаємо миттєву потужність струму $p(t)$ у відповідний момент часу.

Струм у провідниках завжди супроводжується виділенням тепла. При цьому теплова потужність стаціонарного струму, тобто кількість теплоти, що виділяється в провіднику із струмом за одиницю часу, в загальному випадку виражається формулою (4.3а). Відповідно, кількість теплоти Q, що виділяється за час t в провіднику з опором R і силою струму I, визначається формулою

$$Q = I^2Rt, \tag{4.4}$$

яка виражає відомий закон Джоуля.

Якщо провідники не рухаються і в них за рахунок електричної енергії не відбуваються хімічні перетворення, на тепло перетворюється вся енергія струму. В такому разі теплову потужність постійного струму можна обчислювати і за формулами (4.3) і (4.3б), а кількість теплоти — за формулами

$$Q = IUt \tag{4.4а}$$

та

$$Q = \frac{U^2}{R}t. \tag{4.4б}$$

У випадку квазістаціонарного струму $u = u(t)$, $i = i(t)$ відповідні формули мають вигляд:

$$Q = \int i^2 R \, \mathrm{d}t, \tag{4.5}$$

$$Q = \int iu \, \mathrm{d}t, \tag{4.5а}$$

$$Q = \int \frac{u^2}{R} \, \mathrm{d}t. \tag{4.5б}$$

Контрольні запитання

1. Що таке робота електричного струму? Яким є її безпосередній (*на мікроскопічному рівні*) результат?

2. Чи виконують роботу в колі із струмом сили кулонівського електричного поля? Якщо так, то за рахунок якої енергії?

3. В які види, крім теплової, може перетворюватись енергія електричного струму? Наведіть конкретні приклади.

4. За якою із наведених формул (4.4), (4.4а), (4.4б) найдоцільніше обчислювати час закипання води в електричному чайнику з відомим опором нагрівача?

5. Нагрівник електричної плитки складається з трьох однакових секцій, котрі можна використовувати окремо, або як завгодно з'єднавши між собою. За який найменший час можна закип'ятити воду в чайнику, якщо при ввімкненні однієї секції вона закипає за час τ?

6. Електричний двигун працює при напрузі U, опір дроту обмоток його ротора Чи можна за цими даними обчислити кількість теплоти, що виділяється в роторі за заданий час t? Чому?

IV. МАГНІТНЕ ПОЛЕ

Як говорилося на самому початку (Вступ), взаємодія між зарядами, що зумовлена їхнім рухом, називається магнітною і здійснюється через магнітне поле. Магнітні сили відіграють важливу роль у природі та техніці.

Цей розділ присвячено загальним властивостям магнітного поля та його дії на рухомі заряджені частинки й електричні струми. Розглянуто наступні питання:

1. Магнітна взаємодія
2. Потік і циркуляція магнітного поля
3. Магнітне поле в речовині

1. Магнітна взаємодія

Між зарядженими частинками крім електричної спостерігається ще й магнітна взаємодія, яка зумовлена їхнім рухом і здійснюється за допомогою магнітного поля. У цьому можна переконатися на простому досліді. Якщо по двох паралельних дротинах пропустити струми, то вони або притягнуться одна до одної, або відштовхнуться залежно від напрямків струму. Позаяк струм не порушує електричної нейтральності провідників, указана взаємодія зумовлена саме рухом зарядів і здійснюється за польовою схемою: кожен із провідників створює своє магнітне поле, котре діє на інший. При цьому взаємодія припиняється при вимкненні струму в будь-якому з провідників. Це означає, що

магнітне поле створюється рухомими зарядами і діє лише на рухомі заряди.

(Принагідно зауважимо, що магнітні поля породжується також змінними електричними полями (див. розділ V).)

Нижче розглянуті наступні питання:

1.1. Вектор $\vec{B}$

1.2. Магнітне поле точкового заряду

1.3. Магнітне поле струмів

1.4. Дія магнітного поля на провідники із струмом

Контрольні запитання

1.1. *Вектор* $\vec{B}$

Властивості магнітного поля і його кількісні характеристики можна встановити, вивчаючи на досліді магнітні сили, що діють на провідники або контури зі струмом. Але магнітне поле безпосередньо діє не на провідник як такий, а на носії струму, тому логічніше виходити із властивостей сил, які діють у магнітному полі на окремі рухомі заряди. Одразу зауважимо, що, позаяк магнітні сили визначаються не лише величиною заряду, а й величиною та напрямком швидкості його руху, вони мають більш складні властивості, ніж електричні.

Стосовно магнітних сил, які діють на рухомі заряди, дослід показує наступне.

1. У кожній точці магнітного поля серед безлічі напрямів є один такий, при русі зарядженої частинки вздовж котрого магнітне поле на неї не діє. Цей напрям прийнятий за напрям магнітного поля. Інакше говорячи,

при русі зарядженої частинки вздовж напряму магнітного поля магнітна сила на неї не діє.

2. Напрям магнітної сили завжди є перпендикулярним одночасно і до напрямку поля, і до напрямку руху зарядженої частинки, а її модуль F_m є прямо пропорційний величині заряду $|q|$, його швидкості v і синусу кута α між напрямком швидкості та напрямком поля:

$$F_m \sim |q| \cdot v \cdot \sin\alpha.$$

Інакше це можна записати як

$$F_m = B \cdot |q| \cdot v \cdot \sin\alpha, \tag{1.1}$$

де величина

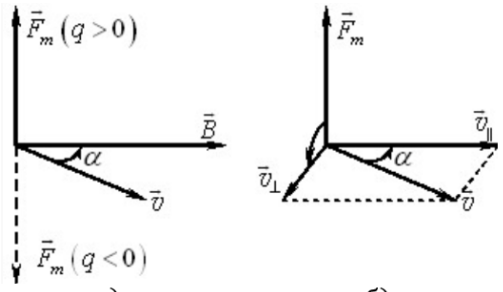
$$B = \frac{F_m}{|q| \cdot v \cdot \sin\alpha} \tag{1.2}$$

не залежить від заряду та характеристик руху частинки і є силовою характеристикою самого поля. Вона називається індукцією магнітного поля або магнітною індукцією. Аби магнітна індукція вказувала також напрям магнітного поля, її розглядають як вектор $\vec{B}$, так що формула (1.1) набуває вигляду:

$$\vec{F}_m = q[\vec{v}\vec{B}], \tag{1.3}$$

де квадратні дужки означають *векторний добуток*. Отже, напрям сили, що діє на рухомий точковий заряд у магнітному полі, визначається правилом правого гвинта («свердлика») з урахуванням знаку заряду (рис. 1.1*а*):

якщо правий гвинт обертати від вектора швидкості заряду до вектора магнітної індукції найкоротшим шляхом, то він буде угвинчуватись у напрямку магнітної сили, що діє на додатній заряд.



а) б)

Рис. 1.1

При зміні знаку заряду напрям магнітної сили змінюється на протилежний.

Варто зауважити, що магнітну силу (1.3) зазвичай називають “силою Лоренца”, що не зовсім точно, бо, строго говорячи, силою Лоренца називається повна сила, що діє на заряд в електромагнітному полі; вона складається з електричної та магнітної компонент:

$$\vec{F} = \vec{F}_e + \vec{F}_m = q\big(\vec{E} + [\vec{v}\vec{B}]\big).$$

Формула (1.3) є однією з основних у теорії магнітного поля. Зокрема, вона дозволяє дати пряме означення вектора $\vec{B}$. Якщо швидкість заряду розкласти на паралельну та перпендикулярну до напрямку поля складові так, що $\vec{v} = \vec{v}_\parallel + \vec{v}_\perp$ (рис. 1.1*б*), то, враховуючи, що $v_\perp = v\sin\alpha$, з виразів (1.1)—(1.3) виходить:

$$\vec{B} = \frac{[\vec{F}\vec{v}_\perp]}{qv_\perp^2}. \tag{1.4}$$

Таким чином,

індукція магнітного поля чисельно дорівнює силі, з якою воно діє на одиничний заряд, що рухається перпендикулярно до напрямку поля з одиничною швидкістю.

Одиниця індукції називається тесла (Тл). Її означення буде дано пізніше, а зараз укажемо, що в полі з індукцією $B = 1$ Тл на заряд $q = 1$ Кл при $v_\perp = 1$ м/с діє сила $F_m = 1$ Н.

1.2. *Магнітне поле точкового заряду*

Формула (1.3) дозволяє обчислити силу, що діє на заряджену частинку в кожній точці магнітного поля. Відтак за допомогою законів механіки можна знайти закон руху частинки в полі. Тому однією з основних задач теорії є встановлення способів розрахунку магнітних полів. Ці способи, як і у випадку електричного поля, ґрунтуються на виразі індукції магнітного поля окремого точкового заряду та принципі суперпозиції.

Із експериментальних даних установлено, що для поля заряду q, що рухається з нерелятивістською швидкістю $\vec{v}$ (рис. 1.2), індукція $\vec{B}$ в будь-якій точці визначається формулою

$$\vec{B} = k\frac{q[\vec{v}\vec{r}]}{r^3}, \qquad B = k\frac{qv_\perp}{r^2}, \tag{1.5}$$

в якій $\vec{r}$ — радіус-вектор даної точки поля *відносно заряду* і $v_\perp = v\sin\vartheta$ — перпендикулярна до $\vec{r}$ складова швидкості заряду.

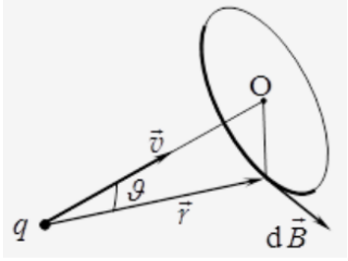


Рис. 1.2

Коефіцієнт пропорційності k залежить від вибраної системи одиниць і в СІ дорівнює

$$k = 10^{-7}\text{Гн/м},$$

де Гн (генрі) — одиниця індуктивності (див. розділ V). Числове значення k є зручним для обчислень, але в літературі при записі формул використовують так звану ***магнітну сталу*** $\mu_0 = 4\pi k$, і формули (1.5) записують у вигляді:

$$\vec{B} = \frac{\mu_0}{4\pi}\frac{q\,[\vec{v}\vec{r}]}{r^3}, \qquad B = \frac{\mu_0}{4\pi}\frac{qv_\perp}{r^2}. \tag{1.5а}$$

Згідно з цією формулою, вектор $\vec{B}$ напрямлений перпендикулярно до площини векторів $\vec{v}$ і $\vec{r}$, тобто по дотичній до кола, що проходить в цій площині через задану точку і має центр O на лінії вектора $\vec{v}$, як показано на рис. 1.2. Зауважимо, що, позаяк величина $\vec{r}$ змінюється з часом, поле точкового заряду є нестаціонарним (залежним від часу): $\vec{B} = \vec{B}(t)$.

Вирази (1.3) і (1.5а) дозволяють визначити силу $\vec{F}_{m1}$, яка діє на заряд q_1, що рухається із швидкістю $\vec{v}_1$ у магнітному полі іншого заряду q_2, який має швидкість $\vec{v}_2$:

$$\vec{F}_{m1} = q_1[\vec{v}_1\vec{B}_2], \quad \text{де} \quad \vec{B}_2 = \frac{\mu_0}{4\pi}\frac{q_2[\vec{v}_2\vec{r}]}{r^3}.$$

Аналогічно можна виразити і силу $\vec{F}_{m2}$, що діє на заряд q_2 з боку поля заряду q_1. Слід зазначити, що позаяк індукції полів і магнітні сили залежать не лише від взаємного розташування, а й від величин та напрямків швидкостей зарядів, магнітна взаємодія між ними є набагато складнішою, ніж електрична. Зокрема, магнітні сили на загал не задовольняють третій закон Ньютона. Виняток становить лише випадок, коли два заряди мають однакову за величиною й напрямом швидкість $v_1 = v_2 = v$ і розташовані один проти одного (рис. 1.3).

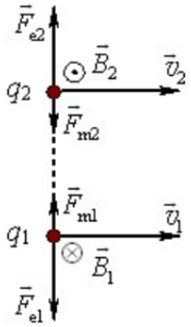


Рис. 1.3

У цьому випадку $\vec{F}_{m1} = -\vec{F}_{m2}$, причому

$$F_{m1} = F_{m2} = \frac{\mu_0}{4\pi}\frac{q_1 q_2 v^2}{r^2}. \tag{1.6}$$

Між зарядами крім магнітних діють ще й електричні сили $\vec{F}_{e2} = -\vec{F}_{e1}$, величина яких визначається законом Кулона:

$$F_{e1} = F_{e2} = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2}. \tag{1.7}$$

У випадку, що розглядається, відношення сил магнітної та електричної взаємодії складає

$$\frac{F_m}{F_e} = \mu_0\epsilon_0 v^2.$$

Величина $\mu_0 = 4\pi \cdot 10^{-7}$ Гн/м, а $\varepsilon_0 = 9 \cdot 10^{-9}/(4\pi)$ Ф/м (див. розділ I, п. 1.2), тож

$$\mu_0\varepsilon_0 = \frac{1}{9 \cdot 10^{16}}\left(\frac{\text{с}}{\text{м}}\right)^2 = \frac{1}{c^2}.$$

де $c = 3 \cdot 10^8$ м/с — швидкість світла у вакуумі. Таким чином

$$\frac{F_m}{F_e} = \frac{v^2}{c^2}. \tag{1.8}$$

1.3. *Магнітне поле струмів*

Із співвідношення (1.8) видно, що при нерелятивістських швидкостях руху магнітна взаємодія між зарядами є несумірно слабшою за електричну. Тому магнітні сили (1.3) є істотними і знаходять практичне застосування тільки за відсутності помітних електричних полів і тільки в потужних магнітних полях, які створюються не поодинокими зарядами, а інтенсивними потоками заряджених частинок, тобто макроскопічними електричними струмами. На практиці магнітні поля створюють, пропускаючи електричний струм по провідниках відповідної конфігурації, як от обмотки ротора електродвигуна або трансформатора чи електромагніта. Властивості магнітного поля струмів і його дія на заряди та інші струми розглядаються в наступних пунктах:

1.3.1. Закон Біо і Савара

1.3.2. Розрахунок магнітного поля за законом Біо і Савара

1.3.1. *Закон Біо і Савара*

Індукцію магнітного поля провідника із струмом можна розрахувати на основі виразу (1.5) для поля точкового заряду та принципу суперпозиції, позаяк кожну елементарну ділянку $\mathrm{d}\vec{l}$ провідника із струмом (рис. 1.4) можна розглядати як рухомий точковий заряд.

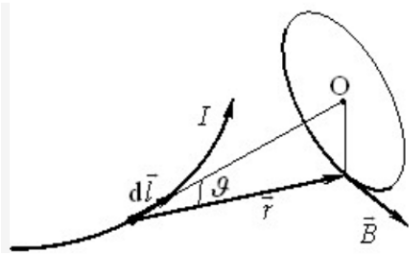


Рис. 1.4

Його величина $\mathrm{d}q$ дорівнює сумарному заряду зосереджених на цій ділянці носіїв струму, а швидкість руху збігається із їхньою дрейфовою швидкістю $\vec{u}$. Отже, в заданій точці простору $\vec{r}$ гранично мала ділянка струму створює елементарне магнітне поле

$$\mathrm{d}\vec{B} = k\frac{\mathrm{d}q[\vec{u}\vec{r}]}{r^3}.$$

Зауважимо, що це поле, на відміну від поля (1.5), є стаціонарним $\mathrm{d}\vec{B} = \mathrm{const}$, бо кожен носій, який виходить із ділянки $\mathrm{d}\vec{l}$ тієї ж миті заміщується носієм, який входить в неї.

Величина $\mathrm{d}q$ у ділянці $\mathrm{d}\vec{l}$ дорівнює

$$\mathrm{d}q = en\,\mathrm{d}V = enS\,\mathrm{d}l,$$

де e, n — заряд і концентрація носіїв, S — площа перерізу провідника. Отже,

$$\mathrm{d}\vec{B} = k\frac{Sen\,\mathrm{d}l\,[\vec{u},\vec{r}]}{r^3} = k\frac{S[en\,\mathrm{d}l\,\vec{u},\vec{r}]}{r^3}.$$

Вектори $\mathrm{d}\vec{l}$ і $\vec{u}$ мають однаковий напрям, тож $\mathrm{d}l\,\vec{u} = u\,\mathrm{d}\vec{l}$, і можна записати:

$$\mathrm{d}\vec{B} = k\frac{enuS[\mathrm{d}\vec{l},\vec{r}]}{r^3}.$$

Згідно з формулами (розділ III, (1.3), (1.4б)), $enuS = I$, отже

$$\mathrm{d}\vec{B} = k\frac{I[\mathrm{d}\vec{l},\vec{r}]}{r^3}.$$

Урахувавши, що $k = \mu_0/4\pi$, отримуємо остаточні вирази:

$$\mathrm{d}\vec{B} = \frac{\mu_0 I[\mathrm{d}\vec{l},\vec{r}]}{4\pi r^3}, \tag{1.9}$$

$$\mathrm{d}B = \frac{\mu_0 I\,\mathrm{d}l\sin\vartheta}{4\pi r^2}. \qquad (1.9\text{a})$$

Вирази (1.9), (1.9а) спочатку були встановлені на основі дослідних фактів і виражають закон Біо і Савара. Згідно з ними вектор $\mathrm{d}\vec{B}$ напрямлений перпендикулярно до площини векторів $\mathrm{d}\vec{l}$ і $\vec{r}$ так само, як поле $\vec{B}$ позитивного точкового заряду, що рухається в напрямку струму. Тобто,

напрям магнітного поля струму збігається із напрямом обертання правого гвинта (свердлика), що укручується в напрямку струму.

Зауважимо також, що для точок на напрямку струму на даній ділянці вектори $\mathrm{d}\vec{l}$ і $\vec{r}$ є колінеарними, і $\mathrm{d}\vec{B} = 0$.

1.3.2. *Розрахунок магнітного поля за законом Біо і Савара*

Закон Біо і Савара разом із принципом суперпозиції дозволяє розрахувати індукцію магнітного поля будь-якого заданого провідника із струмом. Для цього треба знайти вираз (1.9) для довільної нескінченно малої ділянки провідника $\mathrm{d}\vec{l}$ і проінтегрувати його по всій довжині провідника:

$$\vec{B} = \int\limits_L \mathrm{d}\vec{B}. \qquad (1.10)$$

Визначити індукцію у такий спосіб принципово можна завжди. Але розв'язати задачу аналітично і отримати формулу, що визначає поле в будь-якій точці, можна тільки в окремих випадках за сприятливої конфігурації струму. Розглянемо деякі з них:

1. Поле відрізка та нескінченного прямого струму

2. Поле на осі кільцевого струму. Магнітний момент

Поле відрізка прямого струму. Якщо струм I протікає по відрізку прямого провідника (рис. 1.5), то в будь-якій точці P, згідно із законом Біо і Савара (1.9), елементарні поля $\mathrm{d}\vec{B}$ напрямлені перпендикулярно до площини, в якій лежить ця точка та відрізок струму (на рис. 1.5 "до нас"). Тож вектор $\vec{B}$ результуючого поля напрямлений так само, і його модуль (1.10) визначається інтегралом уздовж відрізка

$$B = \frac{\mu_0 I}{4\pi}\int\limits_1^2 \frac{\mathrm{d}l\sin\vartheta}{r^2} \qquad (1.11)$$

від початкової точки 1 до кінцевої точки 2.

Для обчислень позначимо як R відстань від точки P до лінії струму і r — відстань від точки O до довільної точки на відрізку провідника. Тоді величина

$$l = R\,\mathrm{ctg}\,\vartheta$$

є функцією $l(\vartheta)$ кута ϑ (див. рис. 1.5).

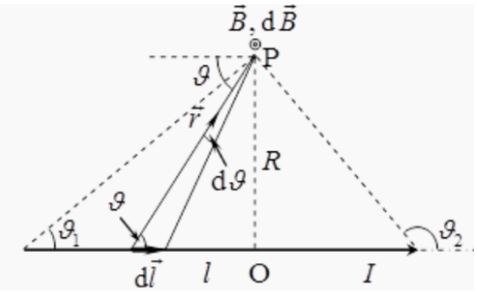


Рис. 1.5

Тому величина $\mathrm{d}l$ під інтегралом (1.11) є модулем приросту цієї функції і за правилами диференціального числення визначається через похідну:

$$\mathrm{d}l = |l'(\vartheta)\mathrm{d}\vartheta| \quad \Rightarrow \quad \mathrm{d}l = R\left(\frac{1}{\sin^2\vartheta}\right)\mathrm{d}\vartheta.$$

Урахуємо також, що $r = (R/\sin\vartheta)$, тоді

$$B = \frac{\mu_0 I}{4\pi R} \int\limits_{\vartheta_1}^{\vartheta_2} \sin\vartheta \, \mathrm{d}\vartheta,$$

і

$$B = \frac{\mu_0 I}{4\pi R} (\cos\vartheta_1 - \cos\vartheta_2), \tag{1.12}$$

де кути ϑ_1 і ϑ_2 визначають довжину відрізка та положення його кінців відносно заданої точки і відраховуються від *напряму струму*; R — відстань від точки до лінії струму.

Поле нескінченного прямого струму. При збільшенні довжини відрізка l кут ϑ_1 буде зменшуватись, а ϑ_2 — зростати так, що при $l \to \infty$ $\vartheta_1 \to 0$ і $\vartheta_2 \to \pi$. Отже, відповідно до виразу (1.12), індукція магнітного поля нескінченного прямого провідника із струмом визначається формулою

$$B = \frac{\mu_0 I}{2\pi R}. \tag{1.13}$$

Напрям вектора $\vec{B}$, як завжди, пов'язаний із напрямом струму правилом правого гвинта.

Поле на осі кільцевого струму. Визначимо тепер індукцію $\vec{B}$ поля на осі колового контуру (кільця) радіуса R із струмом I в залежності від відстані r до центра (рис. 1.6).

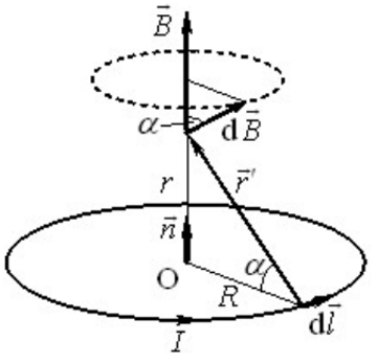


Рис. 1.6

Приступаючи до обчислень за виразом (1.10) одразу врахуємо, що на осі кільця елементарні поля $\mathrm{d}\vec{B}$ від усіх ділянок орієнтовані під однаковим кутом α до осі. Тому перпендикулярні відносно осі складові всіх векторів $\mathrm{d}\vec{B}$ є компенсованими, і результуючий вектор $\vec{B}$ поля всього кільця напрямлений вздовж осі. Отже, для визначення величини B слід інтегрувати тільки осьові складові елементарних полів $\mathrm{d}B_{\parallel} = \mathrm{d}B\cos\alpha$:

$$B = \oint \mathrm{d}B \cos\alpha. \tag{1.14}$$

(Кружальце на символі інтеграла означає, що інтегрування проводиться по замкненому контуру.)

Виражаючи $\mathrm{d}B$ за формулою (1.9а) і враховуючи, що в ній для точок на осі $\vartheta = 90^\circ$ і $r = \mathrm{const}$, отримаємо:

$$B = \frac{\mu_0 I \cos\alpha}{4\pi r'^2} \oint \mathrm{d}l = \frac{\mu_0 I \cos\alpha}{4\pi r'^2} 2\pi R \quad \Rightarrow$$
$$\Rightarrow \quad B = \frac{\mu_0 I R \cos\alpha}{2 r'^2}.$$

Із рис. 1.5 видно, що $\cos\alpha = R/r'$ і $r' = (r^2 + R^2)^{1/2}$, отже,

$$B = \frac{\mu_0 I R^2}{2(r^2 + R^2)^{3/2}}. \tag{1.15}$$

У векторному вигляді

$$\vec{B} = \frac{\mu_0 I R^2 \vec{n}}{2(r^2 + R^2)^{3/2}}, \tag{1.15а}$$

де $\vec{n}$ — орт (одиничний вектор) ***додатньої нормалі*** до площини кільця, тобто нормалі, що пов'язана з напрямом струму правилом правого гвинта.

При $r = 0$ отримуємо формулу індукції магнітного поля в центрі кільцевого струму:

$$\vec{B} = \frac{\mu_0 I \vec{n}}{2R}. \tag{1.16}$$

Напрям магнітного поля на осі кільцевого струму теж визначається правилом правого гвинта, котре в даному випадку звучить так:

якщо правий гвинт обертати в напрямку струму в кільці, то він буде вкручуватись у напрямку магнітного поля, створюваного цим струмом на осі кільця.

Магнітний момент контуру зі струмом. Помноживши у формулі (1.15а) чисельник і знаменник на 2π, отримаємо:

$$\vec{B} = \frac{\mu_0}{4\pi} \frac{2lS\vec{n}}{(r^2 + R^2)^{3/2}},$$

де $S = \pi R^2$ — площа кільця.

Цей вираз показує, що індукція магнітного поля в заданій точці на осі кільця визначається не просто струмом I, а величиною

$$\vec{p} = IS\vec{n} \qquad \text{або} \qquad \vec{p} = I\vec{S}, \tag{1.17}$$

де $\vec{n}$ — орт (одиничний вектор) додатньої нормалі до поверхні контуру, а $\vec{S}$ — «вектор поверхні» контуру. Це поняття можна узагальнити і на неплоскі контури.

Вектор (1.17) називається магнітним моментом контуру і є загальною характеристикою його магнітних властивостей. Вектор $\vec{p}$ визначає магнітне поле кільця також і в інших точках та магнітні поля контурів іншої конфігурації. Дія зовнішнього поля на контур із струмом теж визначається його магнітним моментом. Одиницею магнітного моменту є $1\ \text{А} \cdot \text{м}^2$.

З урахуванням сказаного формула індукції магнітного поля на осі кільцевого струму набуває вигляду:

$$\vec{B} = \frac{\mu_0}{4\pi} \frac{2\vec{p}}{(r^2 + R^2)^{3/2}}. \tag{1.18}$$

Окремий інтерес являють властивості поля на великих відстанях від кільця або при дуже малих його розмірах, коли $r \gg R$. У такому разі, нехтуючи величиною R^2, для поля на осі отримаємо формулу

$$\vec{B} = \frac{\mu_0}{4\pi} \frac{2\vec{p}}{r^3} = k_m \frac{2\vec{p}}{r^3}, \tag{1.19}$$

яка є ідентичною до формули ((4.6), розділ I) для напруженості поля на осі електричного диполя. Так само й інші магнітні властивості малого контуру із струмом і відповідні електричні властивості електричного диполя виражаються ідентичними формулами. Тому малий контур із струмом називають магнітним диполем. Принагідно зауважимо, що через дуже малі розміри магнітного диполя форма контура, що його уособлює, не має значення.

1.4. *Дія магнітного поля на провідники із струмом*

Проходження струму не порушує електричної нейтральності провідників, тому на них не діє зовнішнє електричне поле. Але через дрейф (упорядкований рух) носіїв на провідники зі струмом діє магнітне поле. Далі розглядаються питання:

1.4.1. Сила Ампера

1.4.2. Контур із струмом у магнітному полі

1.4.1. *Сила Ампера*

Закон Ампера. Сила, що діє в магнітному полі на провідник із струмом, називається силою Ампера. Її походження не важко зрозуміти. За наявності струму електрони-носії дрейфують уздовж провідника, і на них діє відповідна усереднена магнітна сила (1.3), що визначається величиною дрейфової швидкості носіїв. Під дією цієї сили електрони на шляху вільного пробігу отримують поперечний імпульс, який відтак передають іонам кристалічної ґратки провідника при зіткненнях. Як наслідок, виникає сила, що діє на провідник як ціле, тобто — сила Ампера. Таким чином, магнітне поле діє на провідник із струмом опосередковано, через носії струму. Тож сила Ампера діє як на рухомі, так і на нерухомі провідники, і її величина дорівнює результуючій магнітній силі, що діє на всі носії струму в провіднику.

Визначимо згідно з описаним механізмом амперову силу $\mathrm{d}\vec{F}$, яка діє на елементарну ділянку $\mathrm{d}\vec{l}$ провідника із струмом I, що вміщений у задане магнітне поле з індукцією $\vec{B}$ (рис. 1.7).

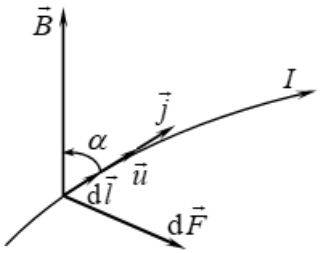


Рис. 1.7

На зосереджені в ділянці носії із сумарним зарядом $\mathrm{d}q$, що рухаються із дрейфовою швидкістю $\vec{u}$, діє магнітна сила

$$\mathrm{d}\vec{F} = \mathrm{d}q[\vec{u}, \vec{B}].$$

Якщо позначити заряд носія, їхню концентрацію та площу перерізу провідника відповідно як e, n і S, то

$$\mathrm{d}q = en\,\mathrm{d}V = enS\,\mathrm{d}l.$$

Тут враховано, що для однаково напрямлених векторів $\mathrm{d}\vec{l}$ і $\vec{u}$ $\mathrm{d}l\,\vec{u} \equiv u\,\mathrm{d}\vec{l}$. Множник перед векторним добутком, відповідно до формул (1.3) і (1.4б), розділ II, дорівнює силі струму в провіднику, отже,

$$\begin{aligned} \mathrm{d}\vec{F} &= I[\mathrm{d}\vec{l}, \vec{B}], \\ \mathrm{d}F &= IB\,\mathrm{d}l\sin\alpha. \end{aligned} \tag{1.20}$$

Формули (1.20) виражають закон Ампера, котрий визначає дію магнітного поля на елементарну ділянку струму.

Напрям сили Ампера визначається правилом правого гвинта (рис.1.7):

якщо правий гвинт обертати від напрямку струму в даній ділянці провідника до напрямку магнітного поля, то він буде укручуватись у напрямку сили Ампера, що діє на дану ділянку.

Зауважимо, що це правило є еквівалентним відомому із школи правилу лівої руки.

Сила Ампера, що діє на весь провідник, складається із сил, які діють на окремі його ділянки, і в загальному випадку визначається інтегруванням виразу (1.20) по всій довжині провідника:

$$\begin{aligned} \vec{F} &= \int \mathrm{d}\vec{F}, \\ \vec{F} &= I\int [\mathrm{d}\vec{l}, \vec{B}]. \end{aligned} \tag{1.21}$$

В окремих випадках формула (1.21) спрощується. До прикладу, якщо довільний провідник знаходиться в однорідному полі $\vec{B} = \mathrm{const}$, то вектор індукції можна винести з-під інтеграла, не змінюючи порядку множників у векторному добутку. Тоді вийде:

$$\begin{aligned} &\vec{F} = I\left[\int \mathrm{d}\vec{l}, \vec{B}\right] \quad \Rightarrow \quad \vec{F} = I[\vec{l}, \vec{B}], \\ &F = IlB\sin\alpha, \end{aligned} \tag{1.22}$$

де $\int \mathrm{d}\vec{l} = \vec{l}$ — вектор проведений від початкової до кінцевої точки провідника (рис. 1.8), l — відстань між цими точками, а для прямолінійного провідника — його довжина.

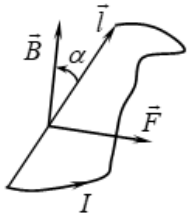


Рис. 1.8

Найпростішою є ситуація, коли маємо прямолінійний провідник розташований перпендикулярно до напрямку однорідного поля. Тоді

$$F = IBl. \tag{1.23}$$

Відмітимо також, що при $\alpha = 0$ або π величина $F = 0$. Отже,

магнітне поле не діє на прямий провідник із струмом, якщо він орієнтований паралельно до напрямку поля.

1.4.2. *Контур із струмом у магнітному полі*

Вирази (1.20) і (1.21) визначають дію магнітного поля на якусь частину замкненого контуру із струмом. Проте зазвичай у полі перебуває весь контур, і при зміні його орієнтації та переміщенні в полі амперові сили виконують роботу. Тому контур із струмом має відповідну енергію у зовнішньому магнітному полі. Далі розглядаються наступні питання:

1) Момент сил, що діє на контур у магнітному полі.

2) Сила, що діє на магнітний диполь у неоднорідному полі.

3) Енергія магнітного диполя у зовнішньому полі.

4) Робота сили Ампера.

Момент сил, що діє на контур. Аби знайти силу, що діє на контур із струмом I в зовнішньому магнітному полі $\vec{B}$, інтегрування у виразі (1.21) треба провести по всьому цьому контуру:

$$\vec{F} = I \oint [d\vec{l}, \vec{B}].$$

При цьому, якщо поле однорідне ($\vec{B} = \text{const}$), то

$$\vec{F} = I \left[\oint \mathrm{d}\vec{l}, \vec{B} \right] = 0, \tag{1.24}$$

бо інтеграл у цьому виразі є сумою безлічі векторів $\mathrm{d}\vec{l}$, які утворюють замкнену лінію.

Таким чином, в однорідному магнітному полі сила Ампера, що діє на будь-який замкнений контур як ціле, дорівнює нулю. Але це не означає, що поле ніяк не діє на контур, бо амперові сили, прикладені до окремих ділянок контуру, не є компенсованими. Аби в цьому переконатися, проаналізуємо дію однорідного поля $\vec{B}$ на жорстку прямокутну рамку із струмом I і сторонами a і b (рис. 1.9*а*).

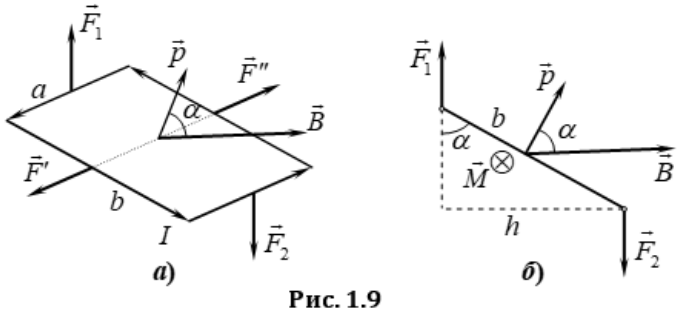


Рис. 1.9

Поле є паралельним, а площина рамки перпендикулярною до площини рисунка (передня сторона рамки і сторона, де струм іде "від нас", наведені жирніше).

При $\vec{B} = \text{const}$ сили (1.22), прикладені до паралельних сторін рамки, мають однакові модулі і протилежні напрямки: $\vec{F}' = -\vec{F}''$ і $\vec{F}_1 = -\vec{F}_2$. При цьому перші дві сили діють по одній лінії в площині рамки і є компенсованими. Але дві інші не компенсовані, бо вони намагаються повернути рамку навколо осі $F' - F''$ і утворюють так звану "пару сил". Обертальна дія пари сил відносно осі $F' - F''$ визначається вектором моменту пари $\vec{M}$, який за правилом правого гвинта напрямлений "від нас" (рис. 1.9б) і має модуль M рівний добутку модуля однієї із сил $F_1 = F_2 = F$ на *плече* пари (відстань h між лініями дії сил):

$$M = Fh.$$

Із рис. 1.9б видно, що $h = b \sin\alpha$, а за формулою (1.23) $F = IaB$, отже,

$$M = IabB \sin\alpha.$$

Величина $Iab = IS = p$ є модулем магнітного моменту рамки, тож

$$M = pB\sin\alpha \quad \Rightarrow \quad \vec{M} = [\vec{p}, \vec{B}]. \tag{1.25}$$

Отриманий вираз показує, що магнітне поле намагається повернути рамку так, аби зорієнтувати її магнітний момент у напрямку поля ($\alpha = 0$), а площину рамки — перпендикулярно до поля. В такому положенні рамки не лише сили F' і F'', а й $\vec{F}_1$ та $\vec{F}_2$ діють уздовж однієї прямої, тож момент $\vec{M} = 0$, і контур перебуває у стані рівноваги. Ця рівновага є стійкою, позаяк при повороті рамки навколо осі, напрям якої не збігається з напрямом поля, виникає момент, що намагається повернути рамку у вихідне положення. При протилежній орієнтації ($\alpha = 180°$) рамка теж буде перебувати в рівновазі, але нестійкій: при щонайменшій зміні орієнтації вона перевернеться у положення $\alpha = 0$.

Можна довести, що в однорідному полі формула (1.25) є правильною для контуру якої завгодно форми та розмірів. Але для довільного контуру в *неоднорідному* полі вона втрачає чинність. Виняток становить магнітний диполь (малий контур із струмом). У цьому випадку поле в усіх точках контуру можна вважати однаковим, тому для магнітного диполя формула (1.25) лишається придатною. Отже,

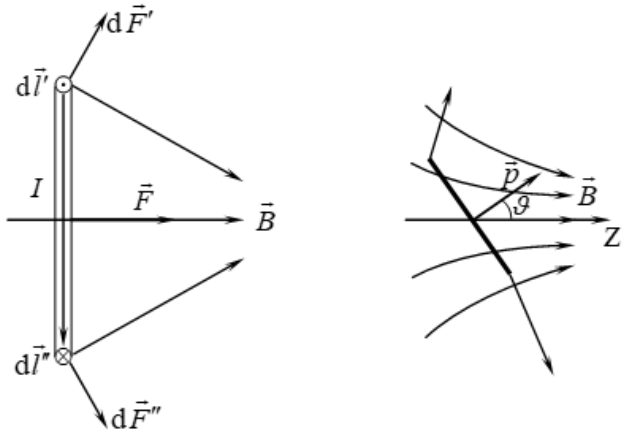
на контур із струмом у зовнішньому магнітному полі діє момент сил, який при малих розмірах контуру визначається виразом (1.25).

Сила, що діє на магнітний диполь у неоднорідному полі. В неоднорідному полі вектор $\vec{B} \neq \text{const}$, тому у виразі (1.22) його не можна винести з-під інтеграла і отримати результат (1.24). А це означає, що

в неоднорідному магнітному поліна на контур із струмом, окрім моменту, діє певна результуюча сила Ампера

$$\vec{F} = I\int [\mathrm{d}\vec{l}, \vec{B}] \neq 0.$$

Це наочно ілюструє рис. 1.10, на якому зображено в профіль круглий контур із струмом (кільце), розміщений в осесиметричному неоднорідному магнітному полі.



а) *б)*

Рис. 1.10

В показаному положенні на всі елементарні ділянки кільця діють однакові за модулем симетричні сили $\mathrm{d}\vec{F}$, які створюють сумарну силу $\vec{F}$, спрямовану вздовж осі Z. При іншій орієнтації контуру з'являється ще й момент сил, який намагається перевести контур у показане положення. Тому у неоднорідному полі вільний магнітний диполь під дією амперових сил буде й обертатися, і рухатися поступально. Так само поводиться електричний диполь у зовнішньому електричному полі (див. розділ I, п. 4). При цьому вирази крутних моментів і результуючих сил в обох випадках теж виявляються ідентичними. Зокрема, сила, що діє на магнітний диполь в осесиметричному неоднорідному полі визначається формулою:

$$F_z = p_z\frac{\partial B}{\partial z}, \tag{1.26}$$

де $p_z = p\cos\alpha$ — проекція магнітного моменту диполя на напрям поля, а $(\partial B/\partial z)$ — градієнт індукції, що є мірою неоднорідності поля.

Енергія магнітного диполя в зовнішньому полі. Вище було показано, що в однорідному зовнішньому полі вільний контур із струмом (магнітний диполь) розташовується у площині перпендикулярній до напрямку поля. Це є положення стійкої рівноваги, тож для зміни орієнтації контуру

на нього має подіяти момент якихось сторонніх сил, аби подолати момент амперових сил, які утримують контур у положенні стійкої рівноваги. Згідно із законами механіки, робота цих сторонніх сил при повороті контуру йде на зміну його енергії:

$$W_2 - W_1 = A_{\text{стор}}. \tag{1.24}$$

На загал енергія контуру складається із механічної енергії та енергії його взаємодії із зовнішнім полем. Але при повільному повороті навколо нерухомої осі механічна енергія контуру не змінюється, і робота, що виконується моментом прикладених сил, йде тільки на зміну магнітної енергії контуру у зовнішньому полі. З механіки відомо, що робота моменту сили при повороті тіла визначається загальним виразом

$$A = \int M \, \mathrm{d}\alpha,$$

де $\mathrm{d}\alpha$ — елементарний кут повороту. Очевидно, що при повільному повороті контуру моменти сторонніх і амперових сил мають однаковий модуль (1.25). Тому при зміні кута α від α_1 до α_2 зміна магнітної енергії контуру в зовнішньому полі складає:

$$W_2 - W_1 = pB \int\limits_{\alpha_1}^{\alpha_2} \sin\alpha \, \mathrm{d}\alpha \quad \Rightarrow$$

$$\Rightarrow \quad W_2 - W_1 = pB\cos\alpha_1 - pB\cos\alpha_2.$$

Цей результат передбачає, що поле однорідне ($B = \text{const}$). Але для малого контуру (магнітного диполя) він зберігає чинність і в неоднорідному полі. Тому на основі отриманого виразу можна зробити висновок, що

магнітний диполь у зовнішньому магнітному полі має енергію, що визначається формулою:

$$W = -\vec{p}\vec{B}, \tag{1.27}$$

або

$$W = -p_z B, \tag{1.27а}$$

де p_z — проекція магнітного моменту диполя на напрям зовнішнього поля.

Робота сили Ампера. Сила Ампера здатна рухати провідник із струмом і виконувати над ним механічну роботу. На цьому принципі, зокрема, працюють електричні двигуни. Але провідник із постійним струмом завжди є частиною якогось провідного контуру. Тому не будемо зосереджуватися на окремому провіднику, а одразу розглянемо роботу сил Ампера при переміщенні в магнітному полі замкненого контуру із струмом. У загальному випадку різні точки контуру здійснюють різні переміщення і на них діють різні за величиною та напрямом амперові сили. Тому спочатку розглянемо роботу $\delta' A = I\vec{F}\,\mathrm{d}\vec{r}$, що виконується при елементарному переміщенні $\mathrm{d}\vec{r}$ однієї елементарної ділянки контуру $\mathrm{d}\vec{l}$ (рис. 1.11).

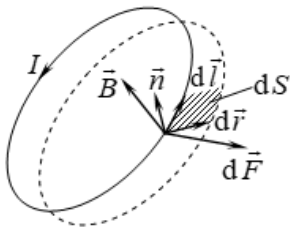


Рис. 1.11

Згідно з виразом (1.20),

$$\delta' A = I[\mathrm{d}\vec{l}, \vec{B}] \cdot \mathrm{d}\vec{r}. \tag{1.28}$$

Отже, ця робота визначається векторно-скалярним (інакше — мішаним) добутком трьох векторів: $\mathrm{d}\vec{l}$, $\vec{B}$ і $\mathrm{d}\vec{r}$. Властивості мішаного добутку трьох векторів $[\vec{a}, \vec{b}] \cdot \vec{c}$ вивчаються у векторній алгебрі. При цьому доводиться, що при так званих циклічних перестановках множників добуток не змінюються, а саме, $[\vec{a}, \vec{b}] \cdot \vec{c} \equiv [\vec{c}, \vec{a}] \cdot \vec{b}$. Тому вираз (1.28) можна переписати як

$$\delta' A = I[\mathrm{d}\vec{r}, \mathrm{d}\vec{l}] \cdot \vec{B}.$$

У цьому виразі вектор $[\mathrm{d}\vec{r}, \mathrm{d}\vec{l}]$ напрямлений по нормалі $\vec{n}$ до площинки $\mathrm{d}S$, а його модуль $\mathrm{d}r\,\mathrm{d}l\sin\alpha = \mathrm{d}S$ дорівнює величині цієї площинки, тож

$$[\mathrm{d}\vec{r}, \mathrm{d}\vec{l}] = \vec{n}\,\mathrm{d}S$$

і

$$\delta' A = I\vec{B}\cdot\vec{n}\,\mathrm{d}S = IB_n\,\mathrm{d}S \quad \Rightarrow$$
$$\Rightarrow \quad \delta\mathsf{A} = I\,\mathrm{d}'\Phi,$$

де $\mathrm{d}'\Phi$ — магнітний потік крізь площадку $\mathrm{d}S$.

Очевидно, що $\mathrm{d}'\Phi$ є частиною приросту потоку $\mathrm{d}\Phi$ крізь усю поверхню контуру при його елементарному переміщенні в полі, а $\delta' A$ — відповідною частиною повної елементарної роботи δA, що виконується при цьому, отже,

$$\delta A = I\,\mathrm{d}\Phi. \tag{1.29}$$

Повна робота при скінченному переміщенні контуру із струмом у магнітному полі дорівнює

$$A = I\cdot\Delta\Phi = I(\Phi_2 - \Phi_1). \tag{1.29а}$$

Таким чином,

при переміщенні контуру із струмом у магнітному полі сили Ампера виконують роботу рівну добутку сили струму в контурі на зміну магнітного потоку крізь поверхню, обмежену цим контуром.

Зауважимо, що коли знак роботи є істотним, при обчисленні потоків слід використовувати додатні нормалі до поверхні контуру, які є пов'язані з напрямом струму правилом правого гвинта.

У зв'язку з отриманим результатом виникає наступне питання. Позаяк амперові сили виконують роботу, вони можуть надавати провідникам із струмом кінетичної, а при наявності, крім магнітного, й інших полів — ще й потенціальної енергії. Але амперові сили є результатом дії на заряди магнітних лоренцових сил (1.3), які є поперечними і ні за яких умов роботи не виконують. То ж звідки береться енергія, за рахунок якої виконують роботу амперові сили? Відповідь полягає в тому, що коли контур із струмом починає рухатися в неоднорідному магнітному полі, то для збереження величини струму на початковому рівні в контурі мають з'явитися додаткові сторонні сили, за рахунок яких і виконується робота (1.29а). Детальніше це питання розглядається в розділі V.

Контрольні запитання

1. Що таке магнітна взаємодія? Чим вона є подібна і чим відмінна від електричної? Що прийнято за напрям магнітного поля?

2. Чи визначає напрям магнітного поля напрям сили, з якою воно діє на частинку із заданим зарядом? Чому?

3. І електричне, і магнітне поля змінюють характеристики руху зарядженої частинки. Чим принципово відрізняються ці зміни?

4. Що називається магнітною індукцією? Дайте формулювання і запишіть математичний вираз вектора $\vec{B}$.

5. Як визначається величина та напрям магнітної сили, що діє на рухомий заряд?

6. Запишіть закон Біо — Савара. Як визначається напрям вектора $\mathrm{d}\vec{B}$ поля елементарної ділянки струму $\mathrm{d}\vec{l}$? Чому дорівнює поле $\mathrm{d}\vec{B}$ на напрямку $\mathrm{d}\vec{l}$?

7. Який напрям має магнітне поле нескінченного прямого струму та поле на осі кільцевого струму?

8. Запишіть математичний вираз закону Ампера. Як визначається напрям сили Ампера?

9. Від чого і як залежить сила Ампера, що діє на довільний провідник із струмом в однорідному магнітному полі?

10. Чому дорівнює сила Ампера, що діє в однорідному магнітному полі на довільний контур із струмом? Чи можна сказати, що в такому випадку дія магнітного поля на струм відсутня?

11. У чому полягає відміна дії однорідного та неоднорідного магнітного поля на контур із струмом?

12. Що таке магнітний момент контура із струмом? Невагому рамку із струмом, яку утримували в однорідному магнітному полі, відпускають. Як поведеться рамка, якщо її магнітний момент складав із напрямком поля кут а) 30° і б) 60°?

Рис. 1.8

2. Потік і циркуляція магнітного поля

У розділі I було показано, що для електричного поля зарядів, окрім напруженості та потенціалу, які характеризують поле в кожній точці (локально), існують й інтегральні величини — потік і циркуляція, — котрі відображують “поведінку” поля на заданій поверхні чи лінії і містять важливу інформацію про його фундаментальні властивості. Нижче подібний аналіз проводиться для магнітного поля струмів у вакуумі. Розглядаються наступні питання:

2.1. Теорема Гаусса для магнітного поля

2.2. Теорема про циркуляцію магнітного поля струмів у вакуумі

2.3. Розрахунок магнітного поля струмів за допомогою теореми про циркуляцію

Контрольні запитання

2.1. *Теорема Гаусса для магнітного поля*

Магнітне поле, як і електричне, задля наочності зображують за допомогою ліній вектора $\vec{B}$ або ліній магнітної індукції. Ці лінії відображають напрям і величину магнітного поля, але не є "силовими" лініями, бо безпосередньо не вказують напрямок магнітних сил.

Дослід свідчить, що лінії магнітного поля є завжди замкненими, тож кожна лінія, що входить всередину якоїсь замкненої поверхні, обов’язково десь виходить назовні. Згідно з геометричною інтерпретацією потоку (розділ I, п. 5), це означає, що потоки на вході та на виході замкненої поверхні завжди мають однакову величину і протилежний знак. Тому

> ***потік вектора індукції магнітного поля крізь довільну замкнену поверхню дорівнює нулю:***

$$\oint\limits_S \vec{B}\,\mathrm{d}\vec{S} = 0. \tag{2.1}$$

У цьому полягає теорема Гаусса для магнітного поля. Вона відображає фундаментальний дослідний факт — відсутність у природі "магнітних зарядів", тобто особливих частинок, які би самі по собі мали здатність створювати магнітне поле подібну до здатності електронів або протони створювати електричне поле. Тому рівняння (2.1) є одним з основних рівнянь класичної електромагнітної теорії.

Наостанок зауважимо, що рівняння (2.1) “завдячує” замкненості ліній магнітного поля, тож є чинним для поля із замкненими лініями, незалежно від його природи. Зокрема, це стосується електричних полів, котрі створюється не зарядами, а нестаціонарними магнітними полями. Картина ліній подібних полів нагадує завихрення, що інколи виникають у потоці води в річці та в повітрі. Тому поля із замкненими лініями називаються *вихровими полями*. Тож

> ***магнітне поле є вихровим силовим полем.***

2.2. *Теорема про циркуляцію магнітного поля струмів у вакуумі*

Розглянемо тепер другу інтегральну характеристику магнітного поля — циркуляцію вектора $\vec{B}$, що за означенням (розділ I, ф-ла (5.2а)) визначається як

$$C = \oint\limits_L \vec{B}\,\mathrm{d}\vec{l}.$$

Циркуляція поля прямого струму. Визначимо спочатку циркуляцію поля одного нескінченного прямого провідника із струмом I по довільному плоскому контуру L, розташованому, як показано на рис. 2.1*а*.

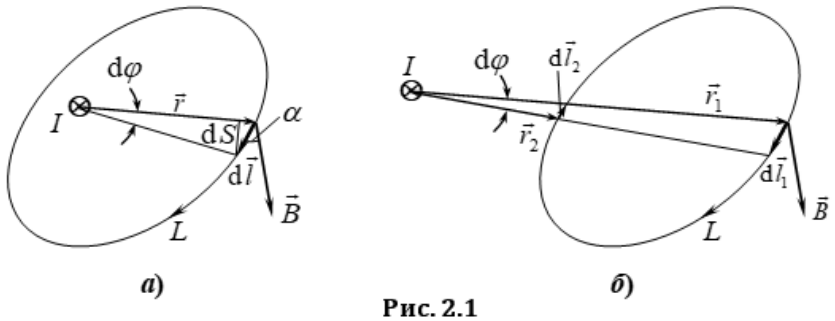


Рис. 2.1

В цьому випадку задача не становить ніяких труднощів. Домовимось обходити контур за годинниковою стрілкою, тоді

$$\oint_L \vec{B}\,\mathrm{d}\vec{l} = \oint_L B\,\mathrm{d}l\cos\alpha, \tag{2.2}$$

а з рис. 2.1*а* зрозуміло, що

$$\mathrm{d}l\cos\alpha = \mathrm{d}s = r\,\mathrm{d}\varphi.$$

Отже,

$$\oint_L \vec{B}\,\mathrm{d}\vec{l} = \oint_L Br\,\mathrm{d}\varphi.$$

Урахувавши, що індукція поля прямого струму визначається формулою (1.13), отримуємо:

$$\oint_L \vec{B}\,\mathrm{d}\vec{l} = \frac{\mu_0 I}{2\pi}\oint_L \mathrm{d}\varphi \quad \Rightarrow$$
$$\Rightarrow \quad \oint_L \vec{B}\,\mathrm{d}\vec{l} = \mu_0 I. \tag{2.3}$$

Для визначення циркуляції у випадку, коли струм не охоплюється контуром, розсічемо його радіальними променями на пари ділянок, як показано на рис. 2.1*б*. Тоді вклад в інтеграл (2.2) від кожної пари складе

$$B_1\,\mathrm{d}l_1\cos\alpha_1 + B_2\,\mathrm{d}l_2\cos\alpha_2 =$$
$$= (B_1 r_1 - B_2 r_2)\mathrm{d}\varphi = 0.$$

Отже,

$$\oint_L \vec{B}\,\mathrm{d}\vec{l} = 0. \tag{2.3а}$$

Теорема про циркуляцію. Отримані результати легко узагальнити на магнітне поле системи паралельних прямих струмів I_i (рис.2.2).

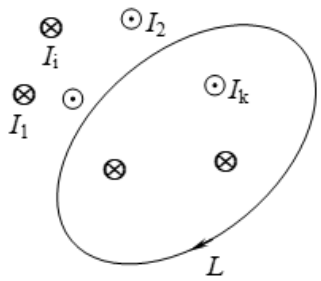


Рис. 2.2

У такому випадку за принципом суперпозиції $\vec{B} = \sum_i \vec{B}_i$ і

$$\oint_L \vec{B}\,\mathrm{d}\vec{l} = \oint_L \sum \vec{B}_i\,\mathrm{d}\vec{l} = \sum \oint_L \vec{B}_i\,\mathrm{d}\vec{l}.$$

Кожен з інтегралів під знаком суми має значення або (2.3), або (2.3а) залежно від розташування струму I_i відносно контуру L. Обидва ці випадки можна об'єднати, приписавши всім струмам, які охоплені контуром L, окремий індекс k. Тоді

$$\oint_L \vec{B}\,\mathrm{d}\vec{l} = \mu_0 \sum_k I_k = \mu_0 I, \tag{2.4}$$

де I — алгебраїчна сума струмів, які охоплені контуром інтегрування. При цьому знак "+" мають струми, напрямки яких пов'язані з обраним при інтегруванні напрямом обходу контуру правилом правого гвинта.

Можна довести, що отриманий результат є загальним. Він зберігає чинність і для провідників довільної форми та для об'ємних струмів, розподілених із заданою густиною $\vec{j}$. Тож у загальному випадку сумарний струм I в рівнянні (2.4) визначається формулою:

$$I = \int_S \vec{j}\,\mathrm{d}\vec{S},$$

де S — довільна поверхня, що спирається на контур інтегрування L. Відтак рівняння (2.4) набуває вигляду

$$\oint_L \vec{B}\,\mathrm{d}\vec{l} = \mu_0 \int_S \vec{j}\,\mathrm{d}\vec{S} \tag{2.4а}$$

і є загальним виразом теореми про циркуляцію магнітного поля струмів у вакуумі:

> ***циркуляція вектора індукції магнітного поля струмів у вакуумі по довільному замкненому контуру дорівнює сумарній силі струму, що охоплений цим контуром, помноженій на*** μ_0**.**

Таким чином, циркуляція магнітного поля по замкненому контуру на загал не дорівнює нулю. Відповідно до критерію потенціальності (розділ I, п. 5.1) це означає, що магнітне поле є непотенціальним полем, тобто, для нього не існує скалярної характеристики подібної до потенціалу φ кулонівського електричного поля.

2.3. *Розрахунок магнітного поля струмів за допомогою теореми про циркуляцію*

Практична значимість теореми про циркуляцію магнітного поля зумовлена тим, що вона спрощує розрахунки магнітних полів високої симетрії, подібно до теореми Гаусса у випадку електричних полів (розділ I, п. 5.2). Розглянемо декілька конкретних прикладів:

1. Поле прямого струму

2. Поле циліндра із струмом

3. Поле циліндричного шару із струмом

4. Поле ідеального соленоїда

5. Поле тороїда

Поле прямого струму. Перш за все продемонструємо переваги використання теореми про циркуляцію на прикладі поля нескінченного прямого провідника із струмом, для якого вже було отримано формулу індукції за допомогою закону Біо і Савара (п. 1.3).

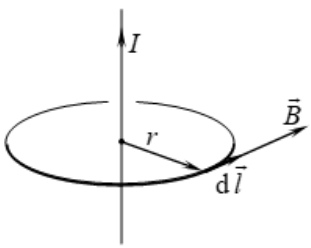


Рис. 2.3

Позаяк індукція такого поля однакова по модулю і напрямлена по дотичній до кола з центром на провіднику, яке є лінією вектора $\vec{B}$ (рис. 2.3), то

$$\oint_L \vec{B}\,\mathrm{d}\vec{l} = B \cdot 2\pi r = \mu_0 I \quad \Rightarrow$$
$$\Rightarrow \quad B = \frac{\mu_0 I}{2\pi r}, \tag{2.5}$$

що збігається з формулою (1.14).

Поле циліндра із струмом. Розрахуємо тепер індукцію магнітного поля струму, що тече вздовж нескінченного суцільного циліндра радіуса R із немагнітної речовини. Будемо вважати, що в усіх точках перерізу циліндра (рис. 2.4) вектор густини струму $\vec{j} = \text{const}$ і напрямлений “до нас”.

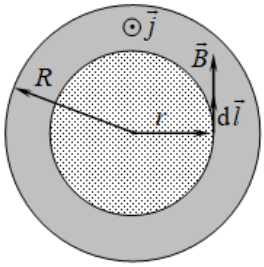


Рис. 2.4

Очевидно, що таке поле, як і в попередньому прикладі, має осьову симетрію, і його індукція теж визначається виразом (2.5). Але тепер величина I залежить від розмірів контуру інтегрування L і має визначатися окремо для області всередині циліндра та поза ним.

Поле всередині циліндра ($r \le R$). У цьому випадку контур L (коло радіуса r на рис.2.4) охоплює частину всього струму, рівну

$$I = j \cdot \pi r^2.$$

Отже, модуль індукції магнітного поля ***всередині циліндра*** дорівнює

$$B = \frac{\mu_0 j r}{2}, \tag{2.6}$$

вектор $\vec{B}$ визначається, як

$$\vec{B} = \frac{\mu_0 [\vec{j}, \vec{r}]}{2}. \tag{2.6а}$$

Поле поза циліндром ($r > R$). За такої умови при будь-якому радіусі контуру L проходить увесь струм

$$I = j \cdot \pi R^2.$$

Тож індукція магнітного поля ***поза циліндром*** визначається формулою

$$B = \frac{\mu_0 j R^2}{2r}, \tag{2.7}$$

або

$$\vec{B} = \frac{\mu_0 R^2 [\vec{j}, \vec{r}]}{2r^2}. \tag{2.7а}$$

На рис. 2.5 показаний графік залежності $B(r)$ поля циліндра.

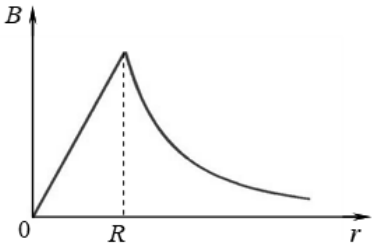


Рис. 2.5

Поле циліндричного шару із струмом. Розглянемо ще й поле, що створюється поздовжнім струмом у циліндричному шарі (циліндрі з коаксіальною порожниною) із радіусами R_1 і R_2 (рис. 2.6), при рівномірному розподілі струму по перерізу із заданою густиною $\vec{j}$.

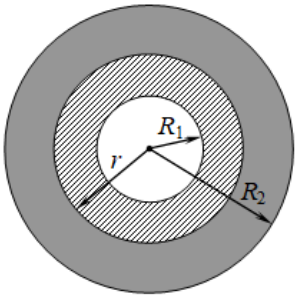


Рис. 2.6

Це поле теж має осьову симетрію, і його розрахунок є аналогічним до попереднього, за винятком того, що він доповнюється розглядом поля в порожнині.

Поле в порожнині ($r < R_1$). При осьовій симетрії поля його індукція скрізь визначається загальним виразом (2.5). Але для будь-якого $r < R_1$ величина $I = 0$, отже,

в порожнині поле відсутнє:

$$\vec{B} = 0.$$

Поле всередині шару ($R_1 \le r \le R_2$). У цьому випадку коло радіуса r охоплює частину всього струму, рівну

$$I = l \cdot \pi(r^2 - R_1^2).$$

Відповідно, індукція поля ***всередині*** циліндричного шару із струмом визначається формулою

$$B = \frac{\mu_0 j \left(r^2 - R_1^2\right)}{2r}, \tag{2.8}$$

або

$$\vec{B} = \frac{\mu_0 [\vec{j}, \vec{r}]}{2} \left(1 - \frac{R_1^2}{r^2}\right). \tag{2.8а}$$

Поле назовні шару ($r > R_2$). Крізь будь-яке коло такого радіуса проходить увесь струм шару. Тому величина I не залежить від r і дорівнює

$$I = j \cdot \pi \left(R_2^2 - R_1^2\right).$$

Отже, індукція поля ***назовні*** циліндричного шару із струмом визначається формулою

$$B = \frac{\mu_0 j \left(R_2^2 - R_1^2\right)}{2r}, \tag{2.9}$$

або

$$\vec{B} = \frac{\mu_0 [\vec{j}, \vec{r}] \left(R_2^2 - R_1^2\right)}{2r^2}. \tag{2.9а}$$

Графік залежності $B(r)$ наведено на рис. 2.7.

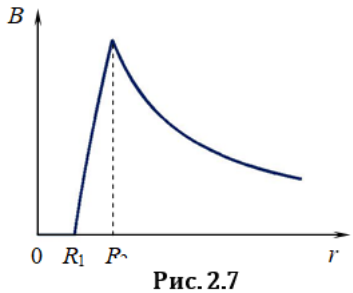
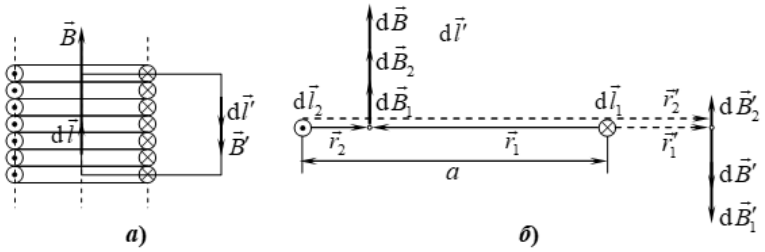


Рис. 2.7

Варто звернути увагу на те, що поле суцільного циліндра і циліндричного шару назовні збігається із полем тонкого провідника розташованого на осі системи, по якому проходить увесь струм. Це ж вірно і для неоднорідного, але симетричного відносно осі розподілу густини струму $\vec{j}(r)$.

Поле ідеального соленоїда. Ідеальним соленоїдом називається нескінченна циліндрична котушка, щільно намотана гранично тонким дротом в один шар (рис. 2.8*а*).

Примітка. Реальний соленоїд — це котушка, довжина котрої є сумірною із діаметром.



а) *б)*

Рис. 2.8

За допомогою закону Біо і Савара можна розрахувати магнітне поле тільки на осі соленоїда. А ось за допомогою теореми про циркуляцію поле ідеального соленоїда можна визначити в усьому просторі. Для цього проаналізуємо симетрію даного поля. Найперше констатуємо, що всі рівновіддалені від осі соленоїда точки є фізично еквівалентними, тож поле в них повинно мати однакову величину. Так само з міркувань симетрії випливає, що поле нескінченного соленоїда має осьовий напрямок, оскільки він є безальтернативним, тобто таким, для якого не існує інших рівноправних напрямів. Всередині соленоїда напрям $\vec{B}$, як і для окремого витка, пов'язаний із струмом правилом правого гвинта, а назовні має бути протилежним. Це випливає із замкненості ліній $\vec{B}$ і видно з рис. 2.8*б*, на якому показані вектори елементарних полів, створюваних парою діаметральних ділянок $d\vec{l}_1$ і $d\vec{l}_2$ одного витка всередині ($d\vec{B}$) і назовні ($d\vec{B}'$). Крім того, поле $\vec{B}'$ має послаблюватися при віддаленні від соленоїда. Такий висновок випливає із закону Біо і Савара. Справді, відповідно до виразу (1.9а) модуль поля dB' двох показаних на рис. 2.8*б* ділянок дорівнює

$$dB' = dB_1' - dB_2' = \frac{\mu_0 I\, dl}{4\pi}\left(\frac{1}{r_1^2} - \frac{1}{r_2^2}\right) = \\ = \frac{\mu_0 I\, dl}{4\pi}\cdot\frac{(r_2 - r_1)(r_2 + r_1)}{r_1^2 r_2^2}.$$

У цьому виразі $r_2 - r_1 = d$ — діаметр витка, а $r_2 + r_1 = 2r$, де r — відстань від осі соленоїда до даної точки. Крім того, за винятком точок поблизу поверхні соленоїда, $r_1 r_2 \approx r^2$. Отже,

$$dB' \approx \frac{\mu_0 I d\, dl}{2\pi}\cdot\frac{1}{r^3}.$$

Зрозуміло, що так само мають спадати з відстанню поля й інших подібних парних ділянок, а отже й поле B' всього соленоїда.

Для визначення поля за теоремою про циркуляцію виберемо прямокутний контур, який має бічну сторону l і розташований у площині діаметрального перерізу соленоїда, як показано на рис. 2.8*а*. При такій орієнтації контуру вектори індукції спрямовані паралельно до бічних і перпендикулярно до поперечних сторін контуру та на кожній бічній стороні мають незмінний модуль. В такому разі

$$\oint_L B_l\, dl = Bl + B'l,$$

і, згідно з теоремою про циркуляцію (2.4),

$$\begin{aligned}(B + B')\, l = \mu_0 N I \quad &\Rightarrow \\ \Rightarrow \quad B + B' = \mu_0 n I,\end{aligned} \tag{2.10}$$

де N — кількість витків соленоїда, що охоплені контуром, а $n = N/l$ — щільність намотки (кількість витків на одиницю довжини соленоїда).

Отриманий вираз означає, що при заданій величині I незалежно від розмірів прямокутника

$$B + B' = \text{const}.$$

Але ця умова може виконуватися тільки при $B' = 0$, бо ця величина не може бути сталою, вона має зменшуватися при віддаленні від соленоїда.

Таким чином,

магнітне поле ідеального соленоїда існує тільки всередині, є однорідним і визначається формулою

$$B = \mu_0 n I. \tag{2.11}$$

Поле тороїда. Тороїдом називається дротяна котушка щільно намотана в один шар по всій поверхні тора (рис. 2.9), який будемо вважати немагнітним.

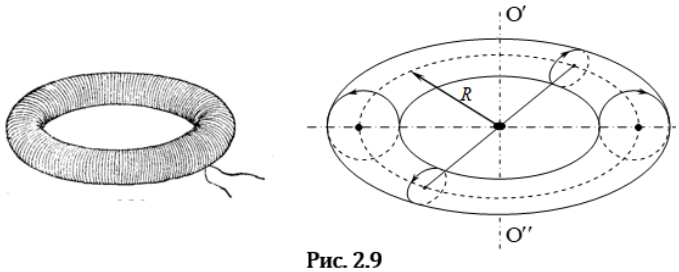


Рис. 2.9

Примітка. *Тор являє собою «математичний бублик», тобто тіло, утворене обертанням кола навколо осі, що розташована в площині кола поза його межами.*

Розрахуємо поле тороїда з радіусами R_1 і R_2, кількістю витків N, по якому тече струм I.

Поле всередині тороїда. З міркувань симетрії очевидно, що всі точки кола з радіусом r, яке проходить всередині тороїда ($R_1 < r < R_2$) і має центр на осі системи $\mathrm{O'O''}$, є фізично еквівалентними. Тому в кожній з них вектор $\vec{B}$ напрямлений перпендикулярно до площини витка і має однаковий модуль. Отже, циркуляція $\vec{B}$ по такому колу

$$\oint\limits_L \vec{B}\, \mathrm{d}\vec{l} = B \cdot 2\pi r,$$

і за теоремою про циркуляцією (10.3)

$$B = \frac{\mu_0 N I}{2\pi r}, \tag{2.12}$$

де NI — повний струм, який охоплюється контуром інтегрування.

Цікаво, що поле всередині тороїда збігається із полем нескінченного прямого провідника із струмом $I_0 = NI$ розташованого на осі $\mathrm{O'O''}$ тороїда.

Формулу (10.11) можна також записати у вигляді

$$B = \mu_0 n I \frac{R_0}{r}, \tag{2.12а}$$

де

$$n = \frac{N}{2\pi R_0}$$

— щільність витків, тобто кількість витків на одиницю довжини середньої лінії тороїда (кола з радіусом $R_0 = (R_1 + R_2)/2$).

Поле поза тороїдом. Якщо коловий контур радіуса r розташований поза тороїдом, то при $r < R_1$ струм узагалі не перетинає поверхню контуру, а при $r > R_2$ струм від кожного витка перетинає її двічі і в протилежних напрямках. Тож в обох випадках

$$\oint\limits_L \vec{B}\,\mathrm{d}\vec{l} = 0 \quad \Rightarrow \quad \vec{B} = 0.$$

Таким чином, як і в соленоїді,

за межами тороїда магнітне поле відсутнє.

Принагідно відмітимо, що з виразу (2.12а) можна отримати формулу індукції магнітного поля ідеального соленоїда, бо його можна формально трактувати як тороїд радіуса $R_0 \to \infty$. У такому разі $(R_0/r) = 1$, і вираз (2.12а) переходить у (2.11).

Наостанку нагадаємо, що в наведених прикладах соленоїд і тороїд розглядалися ідеалізовано, як сукупність щільно складених гранично тонких кілець із струмом. Але насправді кожна з котушок утворює гвинтову лінію. Через це при протіканні струму по витках існує й деяке перенесення заряду, тобто додатковий струм, уздовж осі соленоїда та вздовж середньої лінії тороїда. Тому поза соленоїдом є магнітне поле таке, як поза циліндром із струмом, а поза тороїдом є поле таке, як у кільця радіуса R_0. Але, через те, що котушки зазвичай намотують тонким дротом, крок гвинтової лінії й указані поздовжні струми є дуже малими, тож створювані ними поля є неістотними.

Контрольні запитання

1. Як формулюється та про що свідчить теорема Гаусса для магнітного поля?

2. Яку принципову відміну мають потоки крізь замкнену поверхню у випадку магнітного та кулонівського електричного полів?

3. Як ви гадаєте, чому лінії магнітного поля є завжди замкнені, а лінії кулонівського електричного поля — ні?

4. Яку фундаментальну властивість магнітного поля відображує теорема про циркуляцію вектора $\vec{B}$?

5. Циркуляція електричного поля визначає його роботу при переміщенні одиничного заряду по заданій замкненій траєкторії. Чи можна аналогічно трактувати циркуляцію магнітного поля? Чому?

6. Коли магнітне поле заданого розподілу струму набагато легше розрахувати за допомогою теореми про циркуляцію, ніж за законом Біо — Савара?

3. Магнітне поле в речовині

Речовина змінює характеристики не лише електричного, а й магнітного поля. Загальною (але не єдиною) причиною цього є інтенсивний рух електронів у атомах, який створює в речовині власні мікроскопічні магнітні поля подібно до мікроскопічних електричних полів у діелектриках. При цьому одні речовини (діелектрики) лише послаблюють статичне електричне поле, а інші (провідники) повністю його знищують. Але вплив різних речовин на магнітне поле так радикально не відрізняється, і тому всі вони є магнетиками.

Опис магнітного поля в речовині та загальні властивості різних типів магнетиків розглядаються далі в наступних питаннях:

3.1. Намагнічування речовини

3.2. Вектор $\vec{H}$. Теорема про циркуляцію магнітного поля в речовині

3.3. Магнітні сприйнятливість і проникність

3.4. Умови на межі двох магнетиків

3.5. Магнетики (пара-, діа- і феромагнетики)

Контрольні запитання

3.1. *Намагнічування речовини*

Намагніченість. Замкнений рух електронів в атомах навколо ядер створює мікроскопічні молекулярні струми. Тим самим атоми уподібнюються до мікроскопічних витків із струмом — магнітних диполів, які створюють в околі атомів мікроскопічні магнітні поля. За звичайних умов ці поля в більшості речовин орієнтовані хаотично і є компенсованими. Але при вміщенні магнетика у зовнішнє магнітне поле магнітні моменти і мікрополя атомів частково орієнтуються по зовнішньому полю і перестають бути компенсованими. Як наслідок, у речовині створюється внутрішнє макроскопічне магнітне поле. Цей процес називається намагнічуванням речовини. Він є аналогічним до поляризації діелектриків, при якій відбувається упорядкування електричних моментів молекул і в речовині виникає внутрішнє електричне поле. Подібно до діелектриків уводиться і кількісна характеристика ступеня намагніченості речовини в кожній точці — намагніченість або вектор намагнічування $\vec{J}$:

$$\vec{J} = \frac{1}{\Delta V}\sum \vec{p}_i = \frac{\vec{p}_\Sigma}{\Delta V}, \tag{3.1}$$

де ΔV — *фізично нескінченно малий об'єм* (див. розділ II, п. 1) в околі даної точки речовини, а $\vec{p}_\Sigma = \sum \vec{p}_i$ — сумарний магнітний момент атомів у цьому об'ємі. Помноживши і поділивши вираз (3.1) на кількість N атомів у об'ємі ΔV, намагніченість можна подати через концентрацію частинок речовини $n = (N/\Delta V)$ і середній магнітний момент однієї частинки $\langle \vec{p} \rangle = \frac{1}{N}\sum \vec{p}_i$:

$$\vec{J} = n\langle \vec{p} \rangle. \tag{3.1a}$$

Струми намагнічування. Переважна орієнтація магнітних моментів атомів уздовж напрямку поля при намагнічуванні речовини означає, що електрони в атомах починають рухають узгоджено. Як наслідок, у намагніченій речовині виникають макроскопічні струми намагнічування. Це можна зрозуміти з рис. 3.1, де стрілками схематично показано циркуляцію заряду в атомах поверхневому шару намагніченого вздовж осі циліндричного зразка.

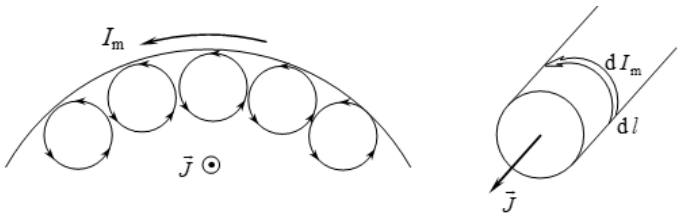


Рис. 3.1

В об'ємі молекулярні струми сусідніх молекул спрямовані взаємно протилежно і є компенсованими, тоді як на поверхні всі молекулярні струми напрямлені однаково. Тому, хоч електрони й не переміщуються від атома до атома, на поверхні відбувається напрямлений рух *електричного заряду*. Таке своєрідне обтікання зарядом поверхні зразка називається поверхневим струмом намагнічування. Розподіл поверхневих струмів намагнічування визначається лінійною густиною струму намагнічування

$$i_m = \frac{\mathrm{d}I_m}{\mathrm{d}l}\ (\text{А/м}), \tag{3.2}$$

тобто, струмом намагнічування, що припадає на одиницю довжини поверхні зразка в напрямку намагніченості (рис. 3.1).

У неоднорідному магнетику повної компенсації молекулярних струмів усередині зразка немає, тому при намагнічуванні виникають не лише поверхневі, а й об'ємні струми намагнічування, котрі, як і звичайні струми, характеризуються поверхневою густиною струму

$$j_m = \frac{\mathrm{d}I_m}{\mathrm{d}S}\ (\text{А/м}^2), \tag{3.3}$$

яка визначає об'ємний струм намагнічування через одиничну площинку, перпендикулярну до напрямку перенесення заряду.

Ще раз наголосимо, що вказане перенесення заряду не пов'язане з рухом електронів (носіїв) по зразку магнетика. Тому для струмів намагнічування немає поняття електричного опору, вони не супроводжуються виділенням джоульового тепла або електрохімічними процесами і можуть існувати в будь-якому середовищі, а не лише в провідниках. Але струми намагнічування і струми провідності є однаковими щодо спроможності створювати магнітне поле. А саме,

> ***магнітне поле струмів намагнічування нічим не відрізняється від магнітного поля струмів провідності такої самої величини та конфігурації.***

Зв'язок між намагніченістю та струмами намагнічування. Уявлення про струми намагнічування дозволяє побудувати послідовну теорію макроскопічного магнітного поля в речовині. Це, зокрема, є зумовлене існуванням безпосереднього зв'язку між струмами намагнічування та намагніченістю. Аби його встановити, виділимо тонкий шар в однорідно намагніченому косому циліндрі з віссю, що складає кут α з вектором $\vec{J}$, і перпендикулярними до нього основами площею S (рис. 3.2).

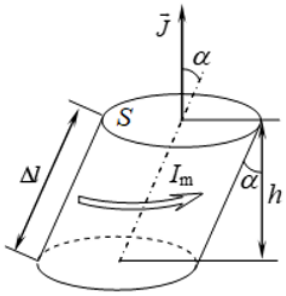


Рис. 3.2

Такий шар можна розглядати як контур із струмом

$$I_m = i_m\,\Delta l,$$

де i_m — лінійна густина поверхневого струму намагнічування, Δl — товщина шару уздовж осі циліндра. Тоді магнітний момент шару має величину $p = I_m S$, а намагніченість, згідно з означенням (3.1),

$$J = \frac{I_m S}{\Delta V} = \frac{i_m \Delta l S}{\Delta V}. \tag{3.3а}$$

Звідси, врахувавши, що $\Delta V = S\Delta l \cos\alpha$, маємо:

$$J\cos\alpha = i_m \qquad \text{або} \qquad J_l = i_m, \tag{3.4}$$

де J_l — осьова складова намагніченості.

Цей результат дозволяє отримати важливий для теорії магнітного поля в речовині вираз циркуляції намагніченості $\oint\limits_L \vec{J}\,\mathrm{d}\vec{l} = \oint\limits_L J_l\,\mathrm{d}l$ по довільному замкненому контуру L (рис. 3.3).

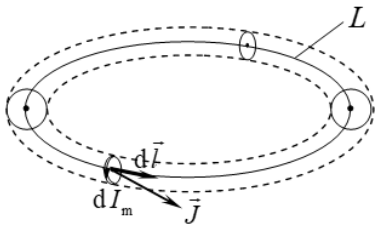


Рис. 3.3

Якщо оточити контур L дуже тонкою циліндричною трубкою, то згідно з (3.4) на кожній ділянці $\mathrm{d}l$ контуру

$$J_l\,\mathrm{d}l = i_m\,\mathrm{d}l = \mathrm{d}I_m,$$

де $\mathrm{d}I_m$ — струм намагнічування, що тече по відповідній ділянці поверхні трубки. Таким чином,

$$\oint\limits_L \vec{J}\,\mathrm{d}\vec{l} = I_m, \tag{3.5}$$

тобто:

> ***циркуляція вектора намагнічування по довільному замкненому контуру дорівнює струмові намагнічування, що охоплюється цим контуром (інакше — перетинає поверхню обмежену цим контуром).***

Цей результат, який можна трактувати як теорему про циркуляцію вектора намагнічування, зберігає чинність і при неоднорідній намагніченості речовини.

3.2. *Вектор $\vec{H}$. Теорема про циркуляцію магнітного поля в речовині*

При намагнічуванні у речовині створюється внутрішнє макроскопічне магнітне поле $\vec{B}'$, яке разом із зовнішнім полем $\vec{B}_0$ визначає результуюче поле в магнетику

$$\vec{B} = \vec{B}_0 + \vec{B}'.$$

Зовнішнє поле створюється макроскопічним рухом (дрейфом) електронів у провідниках, а внутрішнє — мікроскопічними рухами електронів навколо ядер в атомах магнетика. Але, як говорилося, ніякої різниці у загальних властивостях між цими полями немає. Тому так само, як поле $\vec{B}_0$ визначається величиною та конфігурацією струмів провідності, поле $\vec{B}'$ визначається величиною та конфігурацією струмів намагнічування. Отже, на основі теореми про циркуляцію магнітного поля струмів у вакуумі (2.4), для поля в речовині можна записати

$$\oint\limits_L \vec{B}\,\mathrm{d}\vec{l} = \mu_0(I + I_m), \tag{3.6}$$

де I та I_m — сумарна величина струмів провідності та струмів намагнічування, що охоплюються контуром інтегрування.

Записаний вираз формально є подібним до рівняння (2.4). Проте трактувати його як теорему про циркуляцію магнітного поля в речовині по суті не можна, бо він визначає циркуляцію поля $\vec{B}$ через величину I_m, яка сама залежить від $\vec{B}$. Тому для опису магнітного поля в речовині, як і в теорії електричного поля в діелектриках, уводиться відповідний допоміжний вектор. Зважаючи на співвідношення (3.5), перетворимо вираз (3.6) наступним чином:

$$\begin{aligned} &\oint\limits_L \frac{\vec{B}}{\mu_0}\,\mathrm{d}\vec{l} = I + \oint\limits_L \vec{J}\,\mathrm{d}\vec{l} \quad \Rightarrow \\ &\Rightarrow \quad \oint\limits_L \left(\frac{\vec{B}}{\mu_0} - \vec{J}\right)\mathrm{d}\vec{l} = I \end{aligned} \tag{3.7}$$

і введемо величину

$$\vec{H} = \frac{\vec{B}}{\mu_0} - \vec{J}, \tag{3.8}$$

яку у фізиці зазвичай називають “вектором $\vec{H}$”, рідше — “напруженістю” магнітного поля, терміном, який є поширеним в електротехніці. Відтак рівняння (3.7) набуває вигляду

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = I \tag{3.9}$$

і виражає теорему про циркуляцію магнітного поля струмів у речовині:

> ***циркуляція вектора $\vec{H}$ по довільному замкненому контуру дорівнює сумарному струмові провідності I, котрий проходить крізь довільну поверхню обмежену цим контуром.***

Величина I залежить від розподілу струмів провідності у просторі і в загальному випадку визначається інтегральною формулою ((1.4а), розділ I). Тому розгорнуто рівняння (3.9) записується, як

$$\oint \vec{H}\,\mathrm{d}\vec{l} = \int\limits_S \vec{j}\,\mathrm{d}\vec{S}. \tag{3.9а}$$

Слід сказати, що рівняння (3.9) і (3.9а) як окремий випадок вбирають і доведену раніше теорему про циркуляцію магнітного поля у вакуумі. У цьому легко переконатися, взявши до уваги, що у вакуумі $\vec{J} = 0$, тож

$$\vec{H} = \frac{\vec{B}}{\mu_0}.$$

Отже, рівняння (3.9а) відображує загальні властивості магнітного поля більш повно, ніж рівнянням (2.4а) для циркуляції вектора $\vec{B}$. У той же час $\vec{H}$ є допоміжним, а не основним (“польовим”) вектором, бо, згідно з (3.8), у кожній точці він визначається не лише станом поля, а й станом речовини. Але це не зменшує значимість вектора $\vec{H}$, бо він дозволяє одним рівнянням (3.9а) виразити важливі властивості магнітного поля як у вакуумі, так і в будь-якому середовищі.

3.3. *Магнітні сприйнятливість і проникність*

Циркуляція вектора $\vec{H}$ не залежить від струмів намагнічування, тож і від речовини магнетика. Але цього, звичайно, не можна сказати про сам вектор $\vec{H}$, бо згідно з означенням (3.8) він через вектор $\vec{J}$ є пов’язаний із струмами намагнічування. Проте в деяких випадках не тільки циркуляція, а й сам вектор $\vec{H}$ виявляється прямо не залежним від струмів намагнічування, що дозволяє його визначити тільки через струми провідності. З іншого боку, у фізиці матеріалів установлені функціональні зв’язки $\vec{J}(\vec{H})$ між намагніченістю $\vec{J}$ та вектором $\vec{H}$ у різних типах магнетиків. За такої умови означення (3.8) перетворюється на функціональне рівняння, з якого можна знайти індукцію $\vec{B}$ магнітного поля, створеного в магнетику заданим розподілом струмів провідності:

$$\begin{aligned} \vec{H} = \frac{\vec{B}}{\mu_0} - \vec{J}(\vec{H}) \quad &\Rightarrow \\ \Rightarrow \quad \vec{B} = \mu_0\big(\vec{H} &+ \vec{J}(\vec{H})\big). \end{aligned} \tag{3.10}$$

Такий розрахунок включає два етапи:

— визначення допоміжного вектора $\vec{H}(\vec{r})$ через розподіл сторонніх струмів провідності $\vec{j}\,(\vec{r})$,

— визначення вектора індукції поля $\vec{B}(\vec{r})$ за допомогою рівняння (3.10).

У загальному випадку розрахунок є складним через складну залежність $\vec{J}(\vec{H})$. Але майже всі технологічні магнітні матеріали є ізотропними, тобто їхні властивості не залежать від напрямку поля. А в таких речовинах зв’язок між векторами $\vec{J}$ і $\vec{H}$ є простим і має вигляд:

$$\vec{J} = \chi\vec{H}, \tag{3.11}$$

де безрозмірне число χ (читається “*хі*”) називається магнітною сприйнятливістю і показує наскільки добре дана речовина намагнічується у заданому магнітному полі. Тож підставивши співвідношення (3.11) у рівняння (3.10), отримаємо:

$$\vec{B} = \mu_0(1+\chi)\vec{H}. \tag{3.12}$$

Величину $1+\chi$ в цьому виразі позначають одним символом μ і називають магнітною проникністю речовини. Отже, за означенням

$$\mu = 1+\chi. \tag{3.13}$$

Слід сказати, що для переважної більшості речовин $\chi \approx 0$ і $\mu \approx 1$, тобто вони майже не намагнічуються і не впливають на магнітне поле. Але деякі, як до прикладу залізо, є дуже сильними магнетиками, і для них величина χ сягає 10^5 і навіть більше.

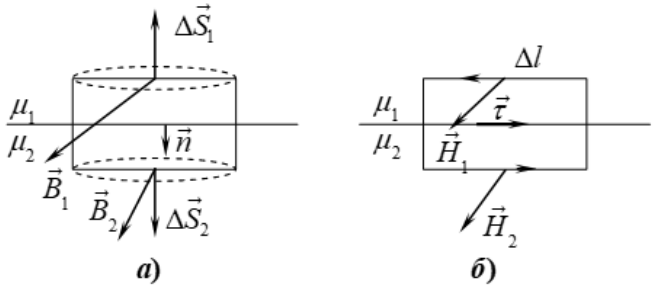
Із виразів (3.12) і (3.13) випливає важлива формула, що напряму пов’язує вектори $\vec{B}$ і $\vec{H}$ в ізотропному магнетику:

$$\vec{B} = \mu_0\mu\vec{H}. \tag{3.14}$$

3.4. *Умови на межі двох магнетиків*

При переході через межу поділу намагнічених речовин характеристики магнітного поля закономірно змінюються. При цьому співвідношення між векторами $\vec{B}_1$ і $\vec{B}_2$ та $\vec{H}_1$ і $\vec{H}_2$ визначаються умовами на межі, котрі випливають із теорем про потік та про циркуляцію магнітного поля.

Спочатку встановимо зв’язок між нормальними складовими векторів у двох ізотропних магнетиках із проникностями μ_1 і μ_2 за відсутності на межі поділу струмів провідності. Для цього розглянемо потік вектора $\vec{B}$ крізь маленький замкнений циліндр із висотою h і паралельними до межі поділу основами ΔS (рис. 3.4*а*).



а) *б)*

Рис. 3.4

Він складається з потоків через основи циліндра $\Delta\Phi_1$, $\Delta\Phi_2$ та бічну поверхню $\Delta\Phi'$ і за теоремою Гаусса (2.1) дорівнює нулю:

$$\Delta\Phi_1 + \Delta\Phi_2 + \Delta\Phi' = 0.$$

Циліндр малий, тому, спроектувавши вектори на спільну нормаль $\vec{n}$ до межі поділу, потоки через основи можна наближено виразити, як $\Delta\Phi_1 \approx -B_{1n}\Delta S$ і $\Delta\Phi_2 \approx B_{2n}\Delta S$. Тоді

$$-B_{1n}\Delta S + B_{2n}\Delta S + \Delta\Phi' \approx 0.$$

Неточність цього виразу зумовлена можливою відміною значень B_n у різних точках площинок ΔS. Тому для отримання точної рівності треба перейти до границі, “стягнувши” циліндр у якусь точку на межі поділу. Для цього спочатку будемо подумки зменшувати висоту циліндра, поступово наближаючи основи до межі поділу. Тоді у границі $h \to 0$ зникне потік крізь бічну поверхню:

$$\begin{gathered} -B_{1n}\Delta S + B_{2n}\Delta S \approx 0 \quad \Rightarrow \\ \Rightarrow \quad B_{1n}\Delta S \approx B_{2n}\Delta S. \end{gathered}$$

При подальшому стяганні основ точність отриманої рівності буде підвищуватись, і в границі $\Delta S \to 0$ отримаємо:

$$B_{1n} = B_{2n}. \tag{3.15а}$$

Відтак за допомогою співвідношення (3.14) встановлюємо зв’язок між нормальними складовими для векторів $\vec{H}$:

$$\mu_1 H_{1n} = \mu_2 H_{2n} \quad \Rightarrow$$
$$\Rightarrow \quad \frac{H_{1n}}{H_{2n}} = \frac{\mu_2}{\mu_1}. \tag{3.15б}$$

Аналогічно за допомогою теореми про циркуляцію вектора $\vec{H}$ (3.9) і формули (3.14) встановимо межові умови для тангенціальних складових векторів. Для цього розглянемо циркуляцію C вектора $\vec{H}$ уздовж малого прямокутника, що охоплює деяку точку на межі поділу магнетиків і орієнтований так, як показано на рис. 3.4*б*. Ураховуючи малість контуру, інтеграл у рівнянні (2.4) наближено виразимо через вклади окремих сторін:

$$-H_{1\tau}\Delta l + H_{2\tau}\Delta l + \Delta C_{\text{п}} \approx 0,$$

де $\Delta C_{\text{п}}$ — спільний вклад від поперечних сторін, а проекції векторів поля на напрям переміщення вздовж поздовжніх сторін записані через спільний орт τ дотичної до межі поділу.

Далі, стягаючи прямокутник у вибрану точку на межі поділу магнетиків, отримаємо

$$H_{1\tau} = H_{2\tau} \tag{3.16а}$$

і, за допомогою формули (3.14),

$$\frac{B_{1\tau}}{\mu_1} = \frac{B_{2\tau}}{\mu_2}. \tag{3.16б}$$

Зведемо всі отримані результати, які складають умови на межі двох магнетиків за відсутності струмів провідності:

$$\begin{aligned} &B_{1n} = B_{2n}, && \frac{B_{1\tau}}{B_{2\tau}} = \frac{\mu_1}{\mu_2}; \\ &\frac{H_{1n}}{H_{2n}} = \frac{\mu_2}{\mu_1}, && H_{1\tau} = H_{2\tau}. \end{aligned} \tag{3.17}$$

Умови (3.17) визначають зв'язок як між модулями, так і між напрямками векторів $\vec{B}$ і $\vec{H}$ на межі ізотропних магнетиків. Зокрема, з них випливає, що напрямки цих векторів задовольняють наступний "закон заломлення" (рис. 3.5):

$$\frac{\text{tg}\alpha_1}{\text{tg}\alpha_2} = \frac{\mu_1}{\mu_2}.$$

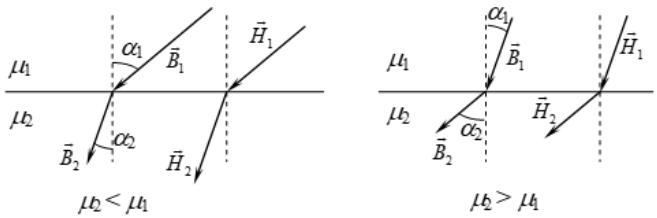


Рис. 3.5

Наостанку зауважимо наступне. Умови на межі (3.17) і зв'язок між векторами магнітного поля (3.14) в ізотропних магнетиках є такий самий, як і в діелектриках ((2.18), (2.14), розділ II). Тож формули для характеристик магнітного поля в ізотропних середовищах, можна отримати з відповідних формул для поля у вакуумі заміною μ_0 на $\mu_0\mu$. Зокрема, для поля ідеального соленоїда з ізотропним осердям це дає

$$B = \mu_0 \mu n I \tag{3.18}$$

та

$$H = nI. \tag{3.19}$$

3.5. *Магнетики (пара-, діа- і феромагнетики)*

Усі речовини є магнетиками, тобто вони так, чи інакше змінюють характеристики магнітного поля. При цьому за ступенем та характером впливу на магнітне поле виділяють декілька типів магнетиків, серед яких найпоширенішими є парамагнетики, діамагнетики та феромагнетики.

Пара- і діамагнетики. До цих двох типів магнетиків належить переважна більшість речовин. Попри різний механізм намагнічування пара- та діамагнетики єднає те, що вони дуже слабко впливають на магнітне поле, їхня сприйнятливість $\chi \approx 0$, а проникність $\mu \approx 1$. При цьому парамагнетики трохи підсилюють, а діамагнетики трохи послаблюють магнітне поле.

Парамагнетиками є речовини, в яких власні магнітні моменти атомів зазвичай орієнтовані хаотично. При внесенні в зовнішнє магнітне поле напрямки магнітних моментів атомів парамагнетика впорядковуються, що зумовлює намагнічування у напрямку поля. Але цьому процесу перешкоджає тепловий рух атомів, отож ступінь упорядкованості моментів і магнітна сприйнятливість парамагнетиків є дуже слабкими і зменшуються при підвищенні температури.

До ***діамагнетиків*** належать речовини, в атомах яких магнітні моменти електронів є компенсованими. Тому атоми діамагнетика не мають власного магнітного моменту і, здавалося б, взагалі не повинні намагнічуватись. Але зовнішнє поле діє не на атом як ціле, а на кожен окремий електрон. При цьому рух електронів навколо ядра змінюється так, що у кожного електрона й атома в цілому з'являється індукований магнітний момент, *напрямлений протилежно* до зовнішнього поля. Тому в діамагнетиках вектори $\vec{J} = \chi\vec{H}$ і $\vec{H}$ є антипаралельними, тож $\chi < 0$ і $\mu < 1$.

Ще однією особливістю діамагнетиків є незалежність магнітної сприйнятливості від температури. Це пояснюється тим, що тепловий рух атомів не спроможний впливати на рух електронів навколо ядра. Слід також зауважити, що зовнішнє магнітне поле створює індуковані моменти електронів у будь-яких атомах. Тому діамагнітна намагніченість є притаманною всім магнетикам. Але це практично не впливає на властивості інших видів магнетиків, оскільки діамагнітна намагніченість є значно слабшою, ніж парамагнітна.

Феромагнетики. До цього класу магнетиків належать атоми і деякі сполуки елементів групи заліза (Fe, Co, Ni) та деякі рідкоземельні елементи, а також сплави на їхній основі. Феромагнетики є сильними магнетиками ($\mu \gg 1$), що зумовлює їхнє практичне застосування в якості технологічних магнітних матеріалів.

Відміна феромагнетиків від слабких магнетиків зумовлена їхньою внутрішньою будовою. З причин, які пояснюються в квантовій фізиці, у феромагнетиках атоми об'єднуються в домени — окремі малі, але макроскопічні (розміром $(1 \div 10)$ мкм і навіть більше) області, в яких магнітні моменти атомів є співнапрямленими. Тож домени є спонтанно (самі по собі) намагнічені до теоретичної границі. Наявність доменної структури саме й спричинює особливі магнітні властивості феромагнетиків.

Основна крива намагнічування. У розмагніченому феромагнетику магнітні моменти доменів орієнтовані хаотично, і в ньому, як і в слабких магнетиках, немає власного макроскопічного магнітного поля. Але, на відміну від них, намагнічування феромагнетика при вміщенні у зовнішнє поле зумовлюється упорядкуванням магнітних моментів не окремих атомів, а цілих доменів, і має складний характер. Його можна дослідити, вмістивши залізний сердечник всередину соленоїда із струмом. Оскільки поле соленоїда напрямлене уздовж осі, модулі напруженості в магнетику H та повітрі H_0 збігаються з їхніми проекціями на поверхню сердечника, тож за умовами (3.17) $H = H_0$. Відтак, узявши до уваги формули (2.10) і (3.14), отримуємо

$$H = nI.$$

Отже, величина H у сердечнику однозначно визначається силою струму в котушці. Тому, поступово збільшуючи величину струму, можна експериментально визначити залежність намагніченості J та індукції B від напруженості H. Вигляд графіка $J(H)$, який називається основною кривою намагнічування феромагнетика, наведений на рис. 3.6*а*.

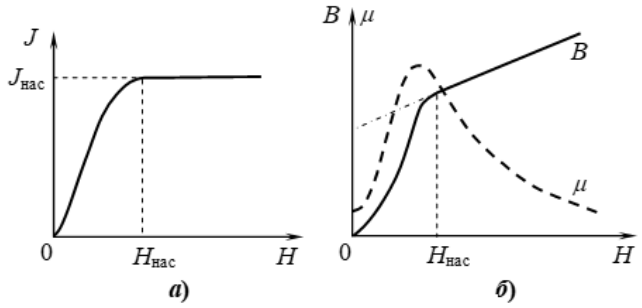


Рис. 3.6

Ця крива наочно відображає особливості намагнічування феромагнетиків — нелінійність і ефект насичення: вже при порівняно невеликій величині $H_{нас}$ в сердечнику намагніченість досягає максимального значення $J_{нас}$ і далі не змінюється. Такий характер намагнічування пояснюється складними процесами впорядкування магнітних моментів доменів, яке завершується практично повною їхньою орієнтацією в напрямку поля. Тому намагніченість $J_{нас}$ є дуже великою і близькою до теоретичного

максимуму. Величина індукції $B(H)$, яка визначається виразом (3.10), при намагнічуванні теж спочатку зростає подібно до $J(H)$, рис. 3.6*б*, але на відміну індукція після насичення намагніченості продовжує потроху (значно повільніше, ніж зображено) зростати за лінійним законом:

$$B(H) = B_m + \mu_0 H, \qquad B_m = \mu_0 J_{\text{нас}}.$$

Феромагнетики зазвичай мають полікристалічну структуру, тобто складаються з маленьких неупорядкованих кристалічних зерен — кристалітів. Тому макроскопічно феромагнітна речовина є ізотропним магнетиком, і формула (3.14) зберігає чинність. Але через нелінійність намагнічування проникність феромагнетика $\mu \neq \text{const}$ і є не табличною характеристикою, а залежною від поля величиною

$$\mu(H) = \frac{B(H)}{\mu_0 H}.$$

Із кривої намагнічування випливає, що $\mu(h)$ спочатку зростає, потім досягає максимуму поблизу насичення, і далі асимптотично спадає до одиниці, як показано на рис. 3.6*б* штриховою кривою.

Гістерезис. Ще однією особливістю феромагнетиків є часткова необоротність процесу намагнічування, яку ілюструє рис. 3.7.

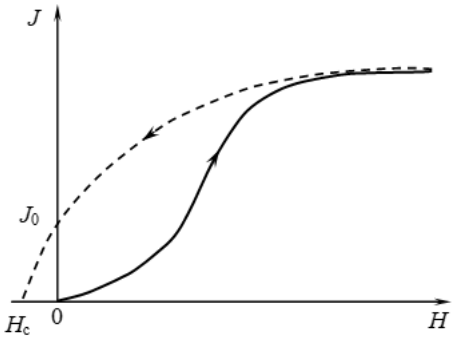
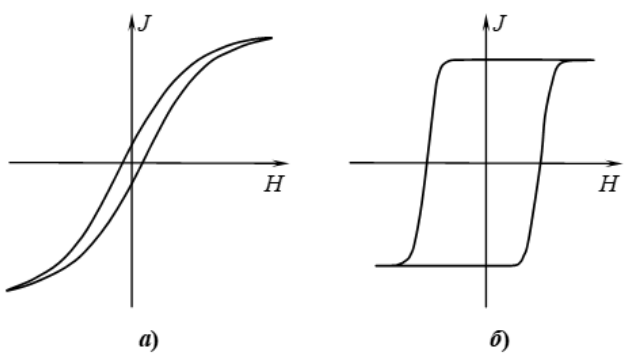


Рис. 3.7

Якщо феромагнітний сердечник намагнітити по основній кривій до насичення (точка А), а потім почати зменшувати струм у котушці, то зміна J не буде поспівати за зміною H, і феромагнетик буде розмагнічуватися не по шляху намагнічування А — 0, а по деякій кривій А — J_0. Таке відставання намагніченості називається *магнітним гістерезисом* ("гістерезис" грецькою саме й означає відставання або спізнення). Через гістерезис при зменшенні напруженості H до 0 феромагнетик повністю не розмагнічується, в ньому зберігається деяка залишкова намагніченість J_0. Тому для повного розмагнічування феромагнетика по котушці треба пропустити струм зворотнього напрямку, котрий створить необхідну для розмагнічування напруженість H_c, яка називається коерцитивною силою.

Залежно від величини коерцитивної сили магнітні матеріали поділяють на магнітном'які та магнітнотверді. Перші мають малу величину H_c, легко перемагнічуються змінним зовнішнім полем і за його відсутності є ненамагніченими. Із таких матеріалів виготовляють сердечники трансформаторів, електромагнітів та інших пристроїв. Напроти, магнітнотверді матеріали мають велику коерцитивну силу і тому зберігають намагніченість навіть у сильних розмагнічуючих полях. Вони, зокрема, використовуються для виготовлення постійних магнітів.

Унаслідок гістерезису при перемагнічуванні феромагнетика змінним струмом залежність $B(H)$ зображується на графіку замкненою кривою, що називається *петлею гістерезису* і виглядає, як на рис. 3.8*а* для магнітом'яких і на рис. 3.8*б* для магнітнотвердих матеріалів.



а) *б*)

Рис. 3.8

При цьому намагнічений магнітнотвердий феромагнетик фактично може перебувати тільки в якомусь із двох станів насичення, перехід між якими відбувається "стрибком", тобто при дуже малій зміні напруженості в околі H_c. Це зумовлює використання таких матеріалів у різноманітних логічних електронних схемах та накопичувачах інформації.

Доменна структура і спроможність до феромагнітного намагнічування є не у всіх кристалічних речовин і не при всіх температурах. У більшості феромагнетиків вона зменшується при нагріванні і зовсім зникає при характерній для кожного феромагнетика температурі, що називається точкою Кюрі. Це пояснюється тим, що в точці Кюрі змінюється взаємне розташування атомів (тип кристалічної ґратки) феромагнетика. Відтак існування доменів стає неможливим і феромагнетик перетворюється на звичайний парамагнетик. Але при охолодженні нижче точки Кюрі феромагнітні властивості речовини повністю відновлюються.

Контрольні запитання

1. Якою є загальна причина впливу речовини на магнітне поле?

2. Чим вплив речовини на магнітне поле відрізняється від впливу на електричне поле?

3. Що таке намагнічування та намагніченість речовини?

4. Що таке струми намагнічування? Чим вони відрізняються від “звичайних” струмів (струмів провідності)?

5. Чи придатні рівняння (2.1) і (2.4) для магнітного поля в речовині? Чому?

6. Запишіть і сформулюйте теорему про циркуляцію магнітного поля в речовині. Чи придатна вона для поля у вакуумі? Чому?

7. Чи є напруженість основним вектором магнітного поля? Чому?

8. Чи завжди виконується співвідношення (3.14)?

9. Що є спільного у впливі на магнітне поле між а) пара- та діамагнетиками і б) пара- та феромагнетиками?

10. У чому полягає та чим пояснюється відміна у впливі на магнітне поле між а) пара- та діамагнетиками і б) пара- та феромагнетиками?

V. ЕЛЕКТРОМАГНЕТИЗМ

У розглянутих досі явищах електричне і магнітне поля виявляють себе як незалежні поля. Але в дійсності вони є двома проявами єдиного електромагнітного поля. Це проявляється в нестаціонарних електромагнітних процесах, які спостерігаються в залежних від часу полях. Далі розглядаються:

1. Явище електромагнітної індукції

2. Рівняння Максвелла

1. Явище електромагнітної індукції

Органічний зв'язок між електричним і магнітним полями найбільш яскраво проявляється у їхній здатності до взаємних перетворень. Зокрема — у здатності нестаціонарного магнітного поля породжувати електричне поле. В цьому полягає глибинний зміст явища електромагнітної індукції (ЕМІ), яке розглядається далі в наступних питаннях:

1.1. Основний закон ЕМІ

1.2. Самоіндукція

1.3. Струм при замиканні та розмиканні кола з індуктивністю

1.4. Енергія магнітного поля

1.5. Природа ЕМІ. Вихрове електричне поле

Контрольні запитання

1.1. *Основний закон ЕМІ*

Явище ЕМІ. Закон Фарадея. Раніше (розділ IV, п. 1.4) було показано, що при зміні магнітного потоку крізь поверхню контуру зі струмом при малому переміщення в магнітному полі виконується робота

$$\delta A = I\,\mathrm{d}\Phi.$$

Але магнітні сили, що діють на носії струму в контурі, є поперечними і як такі роботи не виконують. Тоді за рахунок чого вона виконується?

Аби це з'ясувати, розглянемо такий дослід. Нехай у магнітному полі $\vec{B}$ закріплено дві паралельні металеві рейки, кінці яких з одного боку замкнені на джерело струму із заданою ЕРС $\mathscr{E}$, рис. 1.1.

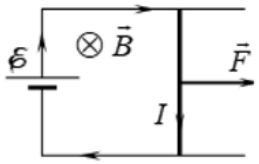


Рис. 1.1

По рейках може ковзати без тертя перетинка із струмом опором R, яку спочатку утримують у фіксованому положенні.

Якщо задля простоти знехтувати опором джерела і рейок, то сила струму в перетинці дорівнює

$$I_0 = \frac{\mathscr{E}}{R}.$$

При цьому вся робота джерела йде на нагрівання провідників і за час $\mathrm{d}t$ складає

$$\delta A = \mathscr{E} I_0\,\mathrm{d}t = I_0^2 R\,\mathrm{d}t.$$

Але після звільнення перетинка почне рухатися під дією сили Ампера, і енергія джерела буде витрачатися вже не тільки на тепло, а й на виконання роботи по переміщенню перетинки в магнітному полі:

$$\mathscr{E} I\,\mathrm{d}t = I^2 R\,\mathrm{d}t + I\,\mathrm{d}\Phi.$$

У такому разі струм у колі зміниться і стане рівним

$$I = \frac{\mathscr{E} - \mathrm{d}\Phi/\mathrm{d}t}{R}.$$

Отже, виходить, що під час руху перетинки в контурі крім ЕРС джерела $\mathscr{E}$ діє ще одна ЕРС, яка дорівнює

$$\mathscr{E}_i = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}. \tag{1.1}$$

Відповідно, сила струму

$$I = \frac{\mathscr{E} + \mathscr{E}_i}{R}.$$

Отриманий результат є одним із проявів відкритого М. Фарадеєм явища електромагнітної індукції (ЕМІ), котре полягає в тому, що

при зміні магнітного потоку крізь поверхню, обмежену провідним контуром, в ньому виникає електричний струм.

Такий струм називається індукційним струмом і створюється електрорушійною силою (ЕРС) індукції (1.1). Отже,

$$I = \frac{\mathscr{E}}{R} = -\frac{1}{R} \cdot \frac{\mathrm{d}\Phi}{\mathrm{d}t}. \tag{1.2}$$

Важливою особливістю явища ЕМІ є його універсальність. Як показали досліди Фарадея, виникнення ЕРС індукції не залежить від причини та способу зміни потоку, причому величина $\mathscr{E}_i$ в усіх випадках визначається виразом (1.1). Тому його називають ***основним законом електромагнітної індукції, або законом Фарадея***.

Потік крізь контур можна змінювати багатьма конкретними способами, але є тільки дві різні причини його зміни. Це або рух контура чи його частини у магнітному полі, або зміна із часом (нестаціонарність) самого магнітного поля. Істотно й те, що ЕРС індукції визначається швидкістю зміни потоку, а не його величиною. Тому за відповідних умов у слабкому полі ЕРС $\mathscr{E}_i$ може виявитися великою, а в сильному — незначною.

Зазначимо також, що в СІ на основі закону (1.1) встановлюється одиниця магнітного потоку вебер (Вб, $1\ \text{Вб} = 1\ \text{В} \cdot \text{с}$):

1 Вб ***— то є магнітний потік, при рівномірній зміні котрого до нуля за час*** 1 с ***у зв'язаному з ним контурі індукується ЕРС*** 1 В**.**

Відтак установлюються одиниці інших магнітних величин. Зокрема, одиниця індукції магнітного поля — тесла (1 Тл) має наступне означення:

1 Тл ***дорівнює індукції такого однорідного магнітного поля, в якому крізь перпендикулярну до нього плоску поверхню площею*** $1\ \text{м}^2$ ***створюється магнітний потік*** 1 Вб**.**

Отже, $1\ \text{Тл} = 1\ \text{Вб}/\text{м}^2$.

Правило Ленца. Досліджуючи електромагнітну індукцію, Фарадей виявив, що напрям індукційного струму залежить від того, зростає потік крізь контур із часом, чи зменшується. Пізніше Ленц дослідним шляхом установив загальне правило, що носить його ім'я і дозволяє визначити напрям індукційного струму в кожному випадку. А саме,

індукційний струм у контурі завжди має такий напрям, при якому перешкоджає причині, що його породжує.

Конкретно це означає, що магнітне поле $\vec{B}_i$, яке створюється індукційним струмом і завжди пов'язане з ним правилом правого гвинта, має бути напрямлене так, аби власним потоком послаблювати зміну потоку поля, яка викликає індукційний струм. Приміром, якщо нерухоме дротяне кільце знаходиться в однорідному нестаціонарному магнітному полі $\vec{B}(t)$, що послаблюється з часом ($\mathrm{d}B/\mathrm{d}t < 0$), то потік крізь кільце спадає. Тоді поле індукційного струму $\vec{B}_i$ має підживлювати поле $\vec{B}$ й індукційний струм буде циркулювати за правим гвинтом відносно напряму $\vec{B}$, як показано на рис. 1.2*а*.

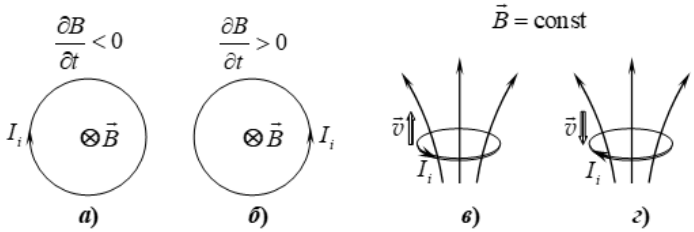


Рис. 1.2

Із тих самих міркувань при ($\mathrm{d}B/\mathrm{d}t > 0$) напрям індукційного струму є протилежним (рис. 1.2*б*). Так можна визначати напрям індукційного струму і в будь-якому іншому випадку. Але зручніше робити це безпосередньо за виразом (1.2), використовуючи наступне правило знаків (див. рис. 1.2*а,б*):

якщо за формулою (1.2) виходить $I_i > 0$***, то індукційний струм пов'язаний із напрямом магнітного поля, що його породжує, правилом правого гвинта. При*** $I_i < 0$ ***напрям струму протилежний.***

До прикладу, коли кільце рухається в постійному неоднорідному полі показаному на рис. 1.2*в*,*г*, то при русі в напрямку поля $\mathrm{d}\Phi/\mathrm{d}t < 0$, а в протилежному напрямку $\mathrm{d}\Phi/\mathrm{d}t > 0$. Тому, у випадку *в*) індукційний струм $I_i > 0$ і тече за годинниковою стрілкою, якщо дивитися вздовж напрямку поля, а у випадку *г*) $I_i < 0$, і струм протікає в протилежному напрямку.

На завершення скажемо, що правило Ленца диктується законом збереження енергії і є одним із проявів загального принципу Ле Шательє — Брауна, за яким усякий зовнішній вплив, що намагається вивести систему із стійкої рівноваги, продукує в ній компенсаційні процеси, спрямовані на нівелювання цього впливу.

1.2. *Самоіндукція*

ЕРС у контурі може індукуватись як зовнішнім магнітним полем, так і власним полем, створюваним струмом у самому контурі. Останній випадок називається самоіндукцією.

Власний потік та індуктивність контуру. Магнітне поле струму, що протікає по замкненому контуру, створює відповідний потік крізь поверхню, обмежену цим контуром. В силу неперервності ліній магнітного поля, цей потік не залежить від площі та форми такої поверхні і називається власним потоком або потоком самоіндукції контуру.

За відсутності феромагнетиків величина індукції магнітного поля провідника із струмом у будь-якій точці простору є прямо пропорційною величині струму: $B \sim I$. Це стосується й потоку поля крізь будь-яку задану поверхню, тож для власного потоку Φ_c контуру із струмом I можна записати:

$$\Phi_\mathrm{c} = LI, \tag{1.3}$$

де величина

$$L = \frac{\Phi_\mathrm{c}}{I} \tag{1.4}$$

називається індуктивністю або коефіцієнтом самоіндукції контуру.

> ***Зауваження.*** *Означення (1.4), строго говорячи, є неоднозначним, бо невідомо, як саме має проходити всередині провідника контур поверхні, через яку визначається величина* Φ_c*. Але, позаяк на практиці контури виготовляють із тонкого дроту, указана обставина є не істотною.*

Одиницею індуктивності є генрі (Гн) — індуктивність контуру, власний потік якого складає 1 Вб при силі струму 1 А:

$$1\ \text{Гн} = 1\ \text{Вб}/\text{А}.$$

Величина L не залежить від сили струму і визначається тільки конфігурацією контуру та магнітними властивостями середовища, в якому він знаходиться. У цьому легко переконатися на прикладі довгого соленоїда, в якому при струмі I створюється магнітна поле з індукцією, що визначається формулою (3.18), розділ IV:

$$B = \mu_0 \mu n I,$$

де n — кількість витків на одиницю довжини соленоїда.

В такому разі потік крізь один виток складає

$$\Phi_1 = \mu_0 \mu n I S,$$

а власний потік усього соленоїда, котрий ще називається ***потокозчепленням***, дорівнює

$$\begin{aligned}\Phi_\mathrm{c} &= N \mu_0 \mu n I S = \mu_0 \mu n^2 I S l = \\ &= \mu_0 \mu n^2 I V,\end{aligned}$$

де N — повна кількість витків соленоїда, S — площа його перерізу, l — довжина і V — об'єм. Відтак за формулою (1.4) індуктивність соленоїда

$$L = \mu_0 \mu n^2 V. \tag{1.5}$$

ЕРС і струм самоіндукції. При зміні з часом власного магнітного потоку контуру в ньому виникає ЕРС самоіндукції $\mathscr{E}_{\text{с}}$, яка, природньо, визначається законом Фарадея (1.1):

$$\mathscr{E}_{\text{с}} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t} = -\frac{\mathrm{d}(LI)}{\mathrm{d}t}.$$

Якщо контур є жорстким і знаходиться в діа- чи парамагнітному середовищі, то $\mu = \text{const}$, $L = \text{const}$, і величина $E_{\text{с}}$ визначається тільки швидкістю зміни струму в контурі:

$$\mathscr{E}_{\text{с}} = -L\frac{\mathrm{d}I}{\mathrm{d}t}. \tag{1.6}$$

Але при деформації не жорсткого контуру чи зміні струму в котушці з феромагнітним осердям індуктивність теж змінюється з часом, і вираз $\mathscr{E}_{\text{с}}$ істотно ускладнюється:

$$\mathscr{E}_{\text{с}} = -L\frac{\mathrm{d}I}{\mathrm{t}} - I\frac{\mathrm{d}L}{\mathrm{d}t}.$$

За таких умов чисто аналітичні розрахунки ланцюгів із індуктивностями стають неможливими, і використовуються інші методи аналізу.

ЕРС самоіндукції створює в контурі додатковий струм самоіндукції $I_{\text{с}}$, напрям якого є підпорядкований правилу Ленца і встановлюється дуже просто:

> ***напрям струму самоіндукції збігається з напрямом основного струму при його зменшенні і є протилежним до основного струму при його зростанні.***

1.3. *Струм при замиканні та розмиканні кола з індуктивністю*

При замиканні та розмиканні кола, в якому є тільки резистори, струм установлюється і зникає практично миттєво. Але в колах із котушками індуктивності при комутаціях виникають індукційні струми, котрі за правилом Ленца уповільнюють зміну основних струмів. Тому при замиканні чи розмиканні кола з індуктивністю струм змінюється поступово.

Струм при замиканні кола. Розглянемо процес установлення струму в простому колі, що складається з котушки індуктивності $L = \text{const}$, резистора R, джерела з ЕРС $\mathscr{E}$ та ключа К (рис. 1.3).

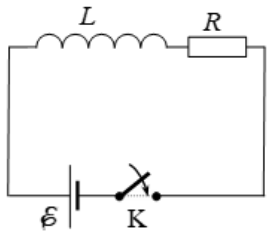


Рис. 1.3

Будемо вважати, що опір дроту котушки включено в R, а опір джерела дорівнює нулю. Також приймемо, що струм є квазістаціонарним (див. розділ III, п. 3.1).

Після замикання в момент $t = 0$ ключа К у котушці виникне ЕРС самоіндукції $\mathscr{E}_{\text{с}}$, так що струм у колі дорівнюватиме

$$I = \frac{\mathscr{E} + \mathscr{E}_{\text{с}}}{R}.$$

Підставивши сюди вираз (1.6), отримаємо диференціальне рівняння

$$\begin{aligned} IR &= \mathscr{E} - L\frac{\mathrm{d}I}{\mathrm{d}t} \quad \Rightarrow \\ \Rightarrow \quad \frac{\mathrm{d}I}{\mathrm{d}t} &= \frac{\mathscr{E} - IR}{L}, \end{aligned} \tag{1.7}$$

з якого можна знайти закон зміни струму $I = I(t)$.

Рівняння (1.7) легко інтегрується методом відокремлення змінних. Але задля зручності спочатку введемо позначення :

$$V = \mathscr{E} - IR. \tag{1.7а}$$

Тоді

$$\frac{\mathrm{d}I}{\mathrm{d}t} = -\frac{1}{R}\frac{\mathrm{d}V}{\mathrm{d}t}$$

і рівняння (1.7) набуває вигляду

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -\frac{V}{\tau} \quad \Rightarrow \quad \frac{\mathrm{d}V}{V} = -\frac{\mathrm{d}t}{\tau},$$
$$\tau = \frac{L}{R}.$$

Проінтегрувавши це рівняння від $t = 0$ до довільного моменту t і від V_0 до V, отримуємо:

$$\int\limits_{V_0}^{V} \frac{\mathrm{d}V}{V} = -\int\limits_{0}^{\tau} \frac{\mathrm{d}t}{\tau} \quad \Rightarrow$$
$$\Rightarrow \quad \ln\frac{V}{V_0} = -\frac{t}{\tau} \quad \Rightarrow \quad V = V_0 e^{-\frac{t}{\tau}}.$$

У момент замикання ключа $t = 0$ струм у колі відсутній: $I = 0$. Отже, відповідно до (1.7а) маємо $V_0 = \mathscr{E}$.

Відтак, зробивши зворотню заміну $V = \mathscr{E} - IR$, дістанемо:

$$I(t) = I_m\left(1 - \mathrm{e}^{-t/\tau}\right),$$
$$I_m = \frac{\mathscr{E}}{R}. \tag{1.8}$$

Отриманий результат має прозорий зміст. А саме: струм у колі складається із двох частин

$$I = I_m + I',$$

де $I_m = \mathscr{E}/R$ — струм, який створюється джерелом, а $I' = -I_m e^{-t/\tau}$ — зустрічний струм самоіндукції, що виникає в котушці.

З часом зустрічний струм самоіндукції поступово зменшується, а повний струм I зростає і наближається до усталеного значення I_m. Наскільки швидко це буде відбуватися, залежить від величини

$$\tau = \frac{L}{R}, \tag{1.9}$$

яка називається часом установлення струму (сталою часу) RL-ланцюжка.

Зокрема, через час τ після замикання ключа струм у колі досягає величини $I = I_m(1 - e^{-1}) = 0{,}63I_m$. При цьому чим більша величина τ, тим повільніше наростає струм (рис. 1.4).

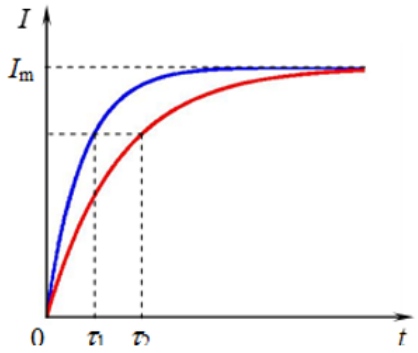


Рис. 1.4

Струм при розмиканні кола. Розглянемо тепер поведінку струму в котушці, яка має індуктивність L і опір r, після розмикання ключа К у колі рис. 1.5. Як і раніше, опір джерела дорівнює нулю.

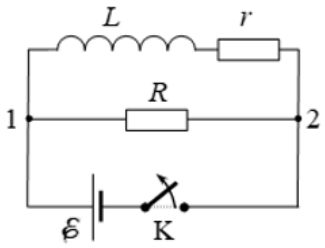


Рис. 1.5

Після розмикання ключа струм через котушку та резистор визначається тільки ЕРС самоіндукції і дорівнює

$$I = \frac{\mathscr{E}_\text{c}}{R + r} = \frac{\mathscr{E}_\text{c}}{R_0},$$

де $R_0 = R + r$ — повний опір кола.

Цей струм називається екстраструмом розмикання і відповідно до (1.6) задовольняє диференціальне рівняння

$$I = -\frac{L}{R_0}\frac{\mathrm{d}I}{\mathrm{d}t} \quad \Rightarrow \quad \frac{\mathrm{d}I}{\mathrm{d}t} = -\frac{I}{\tau},$$
$$\tau = \frac{L}{R_0},$$

що, як і раніше, розв'язується методом відокремлення змінних:

$$\frac{\mathrm{d}I}{I} = -\frac{\mathrm{d}t}{\tau} \quad \Rightarrow \quad \int\limits_{I_0}^{I}\frac{\mathrm{d}I}{I} = -\int\limits_{0}^{t}\frac{\mathrm{d}t}{\tau} \quad \Rightarrow$$
$$\Rightarrow \quad \ln\frac{I}{I_0} = -\frac{t}{\tau}.$$

Отже,

$$I = I_0 \mathrm{e}^{-t/\tau}, \tag{1.10}$$

де величина $I_0 = I(0)$ — струм, який створювався в котушці джерелом $\mathscr{E}$ до розмикання ключа і дорівнював

$$I_0 = \frac{\mathscr{E}}{r}.$$

Таким чином, після відключення джерела в котушці виникає струм, що експоненціально спадає з часом як показано на рис. 1.6.

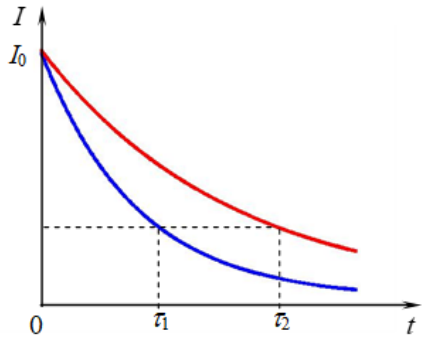


Рис. 1.6

Цей струм тече і через резистор R, тож між точками 1 і 2 створюється напруга

$$U = IR = I_0 R\mathrm{e}^{-t/\tau} = \frac{R}{r}\mathscr{E}\mathrm{e}^{-t/\tau}.$$

Позаяк опір котушки зазвичай малий $r \ll R$, у момент розмикання ключа К створюється стрибок напруги $U \gg \mathscr{E}$. Це може викликати електричний пробій ізоляції, тому в потужних колах з індуктивностями передбачають спеціальні системи захисту від екстраструмів розмикання.

1.4. *Енергія магнітного поля*

При встановленні струму в контурі з індуктивністю джерело виконує роботу по подоланню ЕРС самоіндукції. Тому його енергія витрачається не тільки на джоульове тепло, а й частково переходить в іншу форму, що називається магнітною енергією струму W_m. Про це свідчить те, що після відключення джерела у колі з індуктивністю продовжує текти струм (1.10) і виділятися тепло. Очевидно, що це відбувається за рахунок магнітної енергії струму, адже ніяких інших джерел енергії в колі немає. З іншого боку, за умови квазістаціонарності струму немає й інших каналів перетворення магнітної енергії.

(***Примітка.*** *У випадку нестаціонарних, тобто швидкозмінних, струмів частина цієї енергії йде на електромагнітне випромінювання).*

Тому $W_m = Q$, де Q — кількість тепла, що виділяється при повному згасанні струму в контурі після відключення зовнішнього джерела. Це дозволяє розрахувати магнітну енергію W_m через силу струму I в контурі.

Якщо в колі на рис. 1.5 на момент розмиканні ключа К t = 0 струм у котушці індуктивності дорівнює I, то, згідно з (1.10), надалі він буде зменшуватися за законом $i = Ie^{-t/\tau}$. При цьому магнітна енергія котушки буде перетворюватися в тепло на повному опорі кола R_0=R+r, згідно із законом Джоуля (розділ III, формула (4.2)). Формально цей процес триватиме необмежений час, тож кількість тепла, що виділиться за час до повного згасання струму, визначається, як:

$$Q = R_0 I^2 \int\limits_0^\infty e^{-2t/\tau}\,\mathrm{d}t = \frac{\tau R_0 I^2}{2}.$$

(***Довідка***. Інтеграли з нескінченними границями називаються невласними інтегралами і обчислюється через граничний перехід:

$$\int\limits_0^\infty f(x)\mathrm{d}x = \lim_{a\to\infty}\int\limits_0^a f(x)\mathrm{d}x.$$

Але коли границя перехід є очевидною, із символом ∞ поводяться, як із звичайним числом.

Звідси, враховуючи, що $W_m = Q$ і $\tau = L/R_0$, отримуємо наступну формулу магнітної енергії струму:

$$W_m = \frac{LI^2}{2}. \tag{1.11}$$

Вираз (1.11) показує, що крім величини струму магнітна енергія залежить від параметрів контуру. Наприклад, для соленоїда, зважаючи на формулу (1.5),

$$W_m = \frac{\mu_0\mu n^2 I^2}{2}V. \tag{1.11а}$$

Але для соленоїда $nI = H$ — напруженість поля, тож

$$W_m = \frac{\mu_0\mu H^2}{2}V. \tag{1.11б}$$

Отже, магнітну енергію соленоїда можна виразити не через конфігурацію контуру та величину струму ньому, а через характеристики магнітного поля в соленоїді та об'єм простору який воно займає. В об'ємі соленоїда зосереджене все його поле, тому отриманий результат дає підстави думати, що магнітна енергія зосереджена не в самій котушці із струмом, а в її магнітному полі. Інакше говорячи, магнітне поле має енергію. Такий висновок повністю підтверджується і теорією, і дослідом.

Магнітне поле ідеального соленоїда є однорідним, і його енергія розподілена у просторі рівномірно з об'ємною густиною

$$w_m = \frac{W}{V}\ (\text{Дж}/\text{м}^3),$$

що, згідно з виразом (1.11б), визначається формулою

$$w_m = \frac{\mu_0\mu H^2}{2}. \tag{1.12}$$

Урахувавши співвідношення $B = \mu_0\mu H$ (див. розділ IV), можна записати й такі формули:

$$w_m = \frac{B^2}{2\mu_0\mu} \tag{1.12а}$$

та

$$w_m = \frac{BH}{2}. \tag{1.12б}$$

Таким чином, енергія заданого об'єму однорідного магнітного поля виражається, як

$$W_m = wV.$$

Формули (1.12), (1.12а) і (1.12б) зберігають чинність і для неоднорідного поля, але в такому разі об'ємна густина енергії не скрізь однакова і в кожній точці визначається як

$$w_m = \frac{\mathrm{d}W}{\mathrm{d}V}.$$

Відповідно, енергія в заданому об'ємі V неоднорідного поля знаходиться інтегруванням енергій

$$\mathrm{d}W_m = w\,\mathrm{d}V,$$

зосереджених у всіх елементарних ділянках цього об'єму:

$$W_m = \int\limits_V w\,\mathrm{d}V. \tag{1.13}$$

1.5. *Природа ЕМІ. Вихрове електричне поле*

На початку розділу відзначалося, що є дві різні причини зміни магнітного потоку і виникнення ЕРС індукції та індукційного струму — рух провідників та зміна з часом магнітного поля, в якому знаходиться провідний контур, — але в обох випадках виконується один і той самий закон (1.1). Така універсальність є унікальною рисою явища ЕМІ, бо сторонні сили, що забезпечують рух електронів уздовж контуру, в кожному з указаних випадків мають різне походження.

Індукція в рухомих провідниках. Розглянемо виникнення індукційного струму в провідному контурі утвореному нерухомим металевим каркасом із рухомою дротяною перетинкою, який вміщений у перпендикулярне до його площини однорідне магнітне поле (рис.1.7*а*).

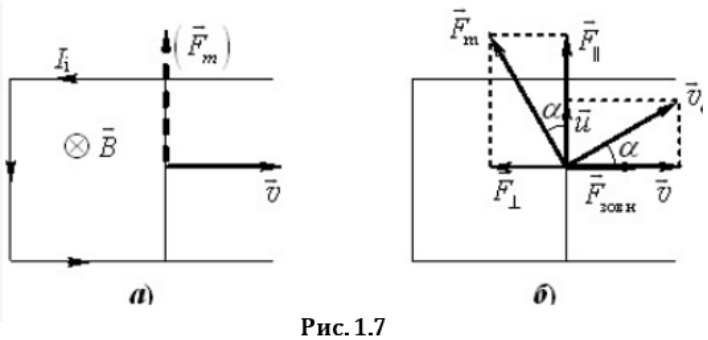


Рис. 1.7

Якщо перетинку привести в рух, як показано, то внаслідок збільшення потоку крізь поверхню контуру, по ньому, відповідно до (1.2), потече індукційний струм I_0, спрямований проти годинникової стрілки. Природа сторонньої сили, яка його створює, на перший погляд є очевидною. Адже з початком руху перетинки вздовж неї на носії струму починає діяти магнітна сила Лоренца $\vec{F}_m$ (розділ IV, ф-ла (1.3)). Тож вона і є причиною переміщення (дрейфу) носіїв по перетинці та виникнення струму. Але в такому поясненні є протиріччя. При протіканні струму сторонні сили мають виконувати над носіями роботу по подоланню електричного опору, тоді як магнітна сила є поперечною і не може виконувати роботу. Пояснення цього "парадоксу" полягає в тому, що насправді швидкість носіїв *відносно магнітного поля* дорівнює не швидкості перетинки $\vec{v}$, а величині $\vec{v}_e = \vec{v} + \vec{u}$, де $\vec{u}$ — дрейфова швидкість, тобто швидкість руху носіїв відносно перетинки. Тому магнітна сила, що діє на носії,

$$\vec{F}_m = e\left[\vec{v}_e, \vec{B}\right], \tag{1.14}$$

напрямлена не вздовж перетинки, а під кутом, як показано на рис.1.7*б* (задля зручності заряд носіїв будемо вважати позитивним). Тому вона має дві складові:

$$\vec{F}_m = \vec{F}_{\parallel} + \vec{F}_{\perp}.$$

При цьому лише паралельна до перетинки складова

$$\vec{F}_{\parallel} = \vec{F}_{\perp} - \vec{F}_m.$$

безпосередньо рухає носії по перетинці. А інша складова $\vec{F}_{\perp}$ гальмує перетинку (це є наочним проявом правила Ленца), і для підтримки руху до неї має бути прикладена якась зовнішня сила

$$\vec{F}_{\text{зовн}} = -\vec{F}_{\perp}.$$

Таким чином, за *сторонню силу*, що створює струм і визначає ЕРС у контурі, править рівнодійна

$$\vec{F}_{\text{стор}} = \vec{F}_m + \vec{F}_{\text{зовн}} = \overrightarrow{F_{\parallel}}.$$

При цьому, позаяк магнітна сила роботи не виконує, робота сторонньої сили збігається з роботою зовнішньої сили:

$$A_{\text{стор}} = A_{\text{зовн}} = A_{\parallel}$$

і виконується за рахунок якогось зовнішнього джерела енергії.

Робота сторонніх сил над носіями струму визначає ЕРС, отже, констатуємо, що

> ***ЕРС індукції в рухомих провідниках зумовлена сумісною дією на носії струму магнітного поля та зовнішніх сил, які підтримують рух. При цьому сили магнітного поля забезпечують дрейф носіїв, а зовнішні сили надають необхідну для цього енергію.***

Описаний ***електронний механізм*** ЕМІ повністю узгоджується із законом Фарадея, в чому можна переконатися прямим розрахунком ЕРС індукції на основі означення ЕРС (див. розділ III, п. 2.1):

$$\mathscr{E} = \frac{A_{\text{стор}}}{q},$$

де $A_{\text{стор}}$ — робота сторонньої сили при переміщенні заряду q по всій довжині перетинки l. Вона дорівнює

$$A_{\text{стор}} = F_{\parallel} l = F_m \cos\alpha \cdot l.$$

З урахуванням виразу (1.14)

$$A_{\text{стор}} = qBv_e \cos\alpha \cdot l = qBvl,$$

де v — швидкість руху перетинки. Відтак для ЕРС індукції в рухомій перетинці отримуємо:

$$\mathscr{E}_i = Blv. \tag{1.15}$$

Якщо тепер узяти до уваги, що $\mathrm{d}v = \mathrm{d}x/\mathrm{d}t$ і $l\,\mathrm{d}x = \mathrm{d}S$ — зміна площі контуру за час $\mathrm{d}t$, то для ЕРС індукції виходить

$$\mathscr{E}_i = \frac{B\,\mathrm{d}S}{\mathrm{d}t} = \frac{\mathrm{d}\Phi}{\mathrm{d}t},$$

у відповідності із законом Фарадея.

Наостанку вкажемо, що отримані результати зберігають чинність і при довільній орієнтації та напрямку руху прямого провідника в магнітному полі. При цьому в будь-який момент часу величина ЕРС індукції визначається, як

$$\mathscr{E}_i = B_{\perp} l v_{\perp}, \tag{1.15а}$$

де $v_{\perp}$ — перпендикулярна до провідника складова його швидкості, а $B_{\perp}$ - складова вектора $\vec{B}$, перпендикулярна до площини руху провідника в цей момент.

Вихрове електричне поле. Коли контур, який знаходиться в стаціонарному магнітному полі $\vec{B} = \mathrm{const}$, не рухається, то потік крізь нього не змінюється з часом, і електромагнітна індукція не спостерігається. Але щойно магнітне поле починає змінюватись, у контурі з'являється ЕРС та індукційний струм у повній згоді з основним законом ЕМІ (1.1). Пояснити цей факт на основі розглянутого електронного механізму неможливо, бо через хаотичність теплового руху на носії в нерухомому контурі ніяка магнітна

сила не діє: $\vec{F}_m = 0$. Отже, стороння сила, що створює індукційний струм, має чисто електричне походження. А це означає, що в нерухомому контурі вміщеному у змінне магнітне поле з'являється електричне поле $\vec{E}_B$, яке й створює сторонню силу

$$\vec{F} = q\vec{E}_B, \tag{1.16}$$

відповідальну за індукційний струм. Такого висновку дійшов Максвелл, який сформулював наступне твердження:

всяке нестаціонарне магнітне поле генерує у просторі відповідне вихрове електричне поле.

На відміну від кулонівського, вихрове поле не створюється зарядами і має замкнені силові лінії. Тому воно не є потенціальним і здатне переміщувати носії струму по замкненому контуру, тобто створювати струм. Робота такого поля віднесена до одиниці перенесеного по контуру заряду — то і є ЕРС індукції. Це дозволяє на основі закону ЕМІ (1.1) і означення ЕРС встановити зв'язок між вихровим електричним полем $\vec{E}_B$ та нестаціонарним магнітним полем $\vec{B} = \vec{B}(t)$, яке його продукує. Згідно з означенням ЕРС,

$$\mathscr{E}_i = \frac{1}{q}\oint_L q\vec{E}_B\,\mathrm{d}\vec{l} \quad \Rightarrow \quad \mathscr{E}_i = \oint_L \vec{E}_B\,\mathrm{d}\vec{l}\,.$$

А за основним законом ЕМІ (1.1) та означенням потоку (розділ I, п. 5.2.1) ЕРС індукції можна подати як

$$\mathscr{E}_i = -\frac{\mathrm{d}}{\mathrm{d}t} = \int_S \vec{B}\,\mathrm{d}\vec{S} = -\int_S \frac{\partial\vec{B}}{\partial t}\,\mathrm{d}\vec{S}.$$

(*Під інтегралом поставлено знак частинної похідної, позаяк при нерухомому контурі зміна потоку крізь поверхню S зумовлюється тільки явною залежністю індукції поля від часу.*)

Таким чином,

$$\oint_L \vec{E}_B\,\mathrm{d}\vec{l} = -\int_S \frac{\partial\vec{B}}{\partial t}\,\mathrm{d}\vec{S}. \tag{1.17}$$

Це встановлене Максвеллом рівняння означає, що

циркуляція напруженості вихрового електричного поля по довільному контуру дорівнює взятому з протилежним знаком потокові швидкості зміни індукції магнітного поля через довільну поверхню обмежену цим контуром.

При цьому напрям обходу контуру пов'язаний із векторами $\mathrm{d}\vec{S}$ правилом правого гвинта, що випливає з правила Ленца. Максвелл також довів, що вектори $\vec{E}_B$ і $\vec{B}(t)$ є завжди взаємно перпендикулярні, тож якщо поле $\vec{B}(t)$ однорідне, то лінії вектора $\vec{E}_B$ лежать у площинах, перпендикулярних до $\vec{B}(t)$ і напрямлені, як на рис. 1.8.

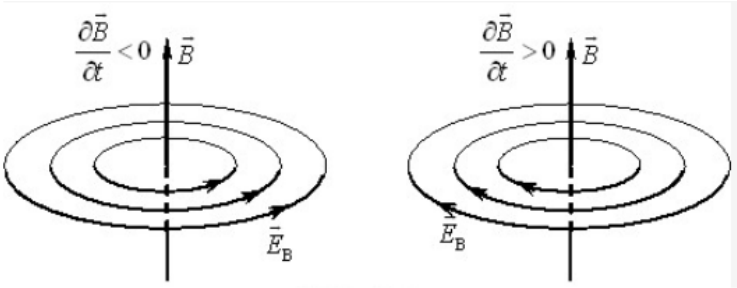


Рис. 1.8

Фарадей трактував електромагнітну індукцію як виникнення струму у замкненому провіднику, вміщеному в магнітне поле. Але за Максвеллом сутність ЕМІ полягає не у виникненні струму в провідниках, а в створенні змінним магнітним полем електричного поля у просторі. А дротяний контур є лишень індикатором цього електричного поля, тому індукційні ефекти мають спостерігатись і в діелектриках та вакуумі. Такий висновок повністю підтверджується дослідом. Ілюстрацією може слугувати бетатрон — один із типів прискорювачів елементарних частинок, у якому електронні пучки у вакуумній камері, вміщеній у змінне магнітне поле, під дією індукованого вихрового електричного поля розганяються майже до швидкості світла.

Таким чином, за Максвеллом співвідношення (1.17) слід трактувати не як вираз ЕРС індукції в дротяному витку, а як рівняння циркуляції вихрового електричного поля по вибраній у просторі замкненій лінії. Але в природі існує ще й кулонівське електричне поле $\vec{E}_K$, створюване зарядами. Тому в загальному випадку електричне поле є суперпозицією полів обох видів, і його напруженість

$$\vec{E} = \vec{E}_K + \vec{E}_B.$$

При цьому кулонівське поле $\vec{E}_K$ є потенціальним і не дає вкладу в циркуляцію (див. розділ I, п. 5.1). Тож будь-яке електричне поле $\vec{E}$ задовольняє рівняння

$$\oint_L \vec{E}\,\mathrm{d}\vec{l} = -\int_S \frac{\partial \vec{B}}{\partial t}\,\mathrm{d}\vec{S}, \qquad (1.17a)$$

яке є загальним виразом теореми про циркуляцію електричного поля.

Відносність механізмів ЕМІ. Істотна відміна розглянутих умов виникнення індукційного струму в дротяному контурі може навести на думку, що індукційні ефекти пов'язані з двома окремими фізичними явищами. Але насправді це не так, бо однозначно розділити їх неможливо. Це наочно ілюструє наступний позірний дослід. Нехай один спостерігач знаходиться на візку 1 біля нерухомого штабового магніту, другий — на візку 2 із закріпленим дротяним витком, а третій — ще на одному візку 3, як схематично показано на рис. 1.9.

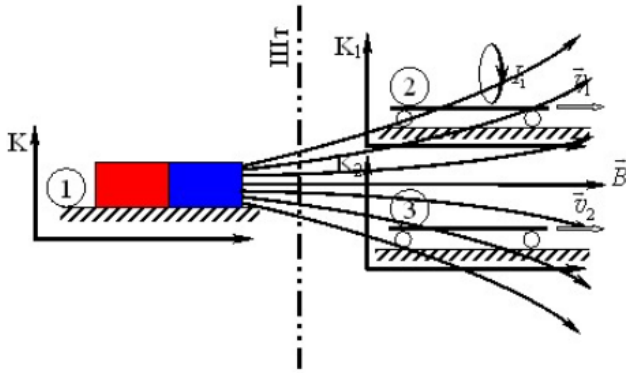


Рис. 1.9

При цьому візки відділені від магніту шторкою Шт із тонованого скла, так що перший спостерігач бачить другого та третього, а вони його — ні. Іншими словами, спостерігачі знаходяться в трьох різних системах відліку K, K_1 і K_2.

Уявімо, що кожен із спостерігачів може реєструвати струм у витку і вимірювати індукцію магнітного поля в місці свого розташування. Тоді, якщо візки 2 і 3 почнуть рухатись із різними швидкостями, всі троє спостерігачів зафіксують появу у витку індукційного струму. Проте вони розійдуться в поясненні причини його виникнення. Перший спостерігач пояснить це дією постійного магнітного поля на носії струму внаслідок руху витка. Другий буде спостерігати струм у нерухомому витку, і пов'яже його з вихровим електричним полем, оскільки виявить, що магнітне поле змінюється з часом. *(Ця зміна зумовлена зменшенням величини B у напрямку руху візка з витком.)* Третій спостерігач пояснить струм обома причинами, бо для нього й магнітне поле змінюється, й виток рухається. При цьому аргументи кожного спостерігача є неспростовними. Отже, відповідь на запитання, чому виникла ЕМІ і яка її частка зумовлена магнітним, а яка електричним полем, в кожному випадку залежить від умов спостереження, тобто від системи відліку. Це промовисто ілюструє тезу, висловлену ще у Вступі:

> ***носієм електромагнітної взаємодії є єдине електромагнітне поле, поділ якого на електричне та магнітне поля є відносним, залежним від системи відліку.***

Іншими словами, електричне та магнітне поля є не різними полями, а різними проявами одного поля, котрі в різній мірі виявляють себе в різних системах відліку. Відповідно, існують встановлені Лоренцом формули перетворень електричного та магнітного полів від однієї до іншої системи відліку, на яких не будемо тут зупинятись. Укажемо лише, що в окремих випадках можна знайти таку систему відліку, в якій існує тільки одне з полів (у розглянутому прикладі це система K_1, де є лише магнітне поле), а в інших випадках, як, наприклад, для електромагнітних хвиль, таку систему знайти неможливо.

Контрольні запитання

1. У чому полягає відміна між термінами "магнітна індукція" та "електромагнітна індукція"?

2. Що є умовою виникнення в контурі ЕРС індукції та чим визначається її величина?

3. За якої умови в контурі, що рухається в магнітному полі, не буде індукуватись ЕРС? Наведіть приклади.

4. У чому полягає правило Ленца?

5. Поясніть, як визначається напрям індукційного струму в контурі через знак ЕРС індукції за формулою (1.1)?

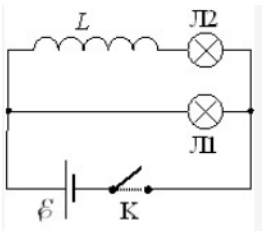
6. Доведіть, що при будь-якій зміні магнітного потоку крізь дротяний виток унаслідок його руху, магнітне поле буде гальмувати цей рух.

7. Дротяна рамка може обертатися в магнітному полі навколо осі, що лежить у площині рамки і перпендикулярна до напрямку поля. Поясніть, чому рамку легше обертати, коли вона розімкнена.

8. Поясніть, як зміниться індуктивність котушки, щільно намотаної на циліндричне осердя, при її розтяганні в довжину.

9. Кругле дротяне кільце, приєднане через резистор до джерела струму, починають розтягати по діаметру в лінію. Як це вплине на величину струму в резисторі? Чому?

10. Паралельний ланцюжок складається з електричної лампочки та такої самої лампочки послідовно приєднаної до котушки з великою індуктивністю L:



Опишіть та поясніть поведінку лампочок при замиканні та розмиканні ключа К.

11. Поясніть походження сторонніх сил, які є відповідальні за індукційний струм у рухомому та в нерухомому контурі.

2. Рівняння Максвелла

Явище електромагнітної індукції свідчить про те, що електричне поле породжується не лише зарядами, а й магнітними полями. Але Максвелл також установив, що й магнітне поле створюється не тільки рухомими зарядами (струмами), а й електричними полями. Відтак стало можливим відобразити загальні властивості електромагнітного поля в чотирьох рівняннях, які складають фундамент усієї класичної електродинаміки.

Далі розглядаються наступні питання:

2.1. Струм зміщення

2.2. Рівняння Максвелла в інтегральній формі

2.3. Диференціальні характеристики поля

2.4. Рівняння Максвелла в локальній формі

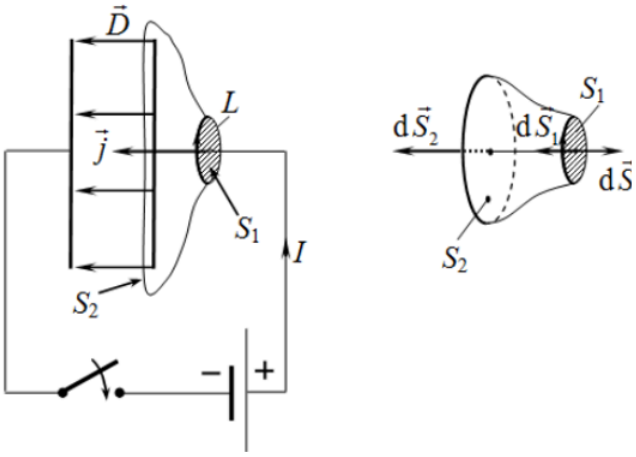
Контрольні запитання

2.1. *Струм зміщення*

До думки про здатність електричного поля створювати магнітне поле можна дійти, до прикладу, аналізуючи процес заряджання конденсатора. При підключенні конденсатора до джерела у просторі виникне магнітне поле, створене струмом заряджання. При цьому, якщо охопити якимось контуром L провідник, що підходить до обкладки конденсатора (рис. 2.1*а*), то, відповідно до теореми про циркуляцію магнітного поля ((3.9), розділ IV), можна записати:

$$\oint_L \vec{H}\,\mathrm{d}\vec{l} = \int_{S_1} \vec{j}\,\mathrm{d}\vec{S}_1, \tag{2.1}$$

де S_1 — поверхня, що спирається на контур L, $\vec{j}$ — густина струму на цій поверхні.



а) б)

Рис. 2.1

Інтеграл у правій частині (2.1) визначає сумарний струм, який охоплюється контуром L, тож не залежить від розмірів та форми поверхні S_1. Отже, можна взяти якусь іншу поверхню, скажімо у формі "глечика", що охоплює одну обкладку конденсатора, як схематично показано на рис. 2.1. Ніякі струми цю поверхню не перетинають, і на ній скрізь $\vec{j} = 0$. Тому рівняння (2.1) мало би виглядати як

$$\oint_L \vec{H}\,\mathrm{d}\vec{l} = 0. \tag{2.1а}$$

Але цього не може бути бо йдеться про циркуляцію того самого поля $\vec{H}$ по тому самому контуру L. Тож логічно припустити, що циркуляцію нестаціонарного магнітного поля, крім струму провідності, визначає ще якийсь фактор. Його вірогідну природу можна відшукати на основі закону збереження електричного заряду, який для нестаціонарних процесів виражається рівнянням неперервності (розділ III, п. 1.2):

$$\oint_S \vec{j}\,\mathrm{d}\vec{S} = -\int_V \frac{\partial \rho}{\partial t}\mathrm{d}V, \tag{2.2}$$

яке засвідчує, що зміна величини заряду в об'ємі V за будь-який проміжок часу в точності дорівнює кількості заряду, що проходить за той самий час крізь замкнену поверхню S, що обмежує цей об'єм.

Для вибраної замкненої поверхні (рис. 2.1*а*) підінтегральна функція $\vec{j}$ в рівнянні (2.2) відмінна від 0 тільки на "кришці" S_1 (рис. 2.1*б*), отже, його можна подати як

$$\int\limits_{S_1} \vec{j}\,\mathrm{d}\vec{S}_1' = -\int\limits_V \frac{\partial\rho}{\partial t}\mathrm{d}V. \tag{2.2а}$$

Урахуємо також, що при обчисленні потоку крізь замкнену поверхню вектори елементарних площинок спрямовують по зовнішніх нормалях, а при обчисленні циркуляції за рівнянням (2.1) узгоджують із напрямком обходу контуру L за правилом правого гвинта. В такому разі $\mathrm{d}\vec{S}_1' = -\mathrm{d}\vec{S}_1$, і попередній вираз набуває вигляду:

$$\int\limits_{S_1} \vec{j}\,\mathrm{d}\vec{S}_1 = \int\limits_V \frac{\partial\rho}{\partial t}\,\mathrm{d}V. \tag{2.2б}$$

Інтеграл в лівій частині визначає струм, який проходить через поверхню S_1, тож

$$\int\limits_{S_1} \vec{j}\,\mathrm{d}\vec{S}_1 = \frac{\mathrm{d}q}{\mathrm{d}t},$$

де $\mathrm{d}q$ — заряд, який заходить всередину "глечика" і осідає на пластині конденсатора за час $\mathrm{d}t$, а похідна визначає швидкість зміни заряду пластини q. Цей заряд створює електричне поле, в якому за теоремою Гаусса ((2.9а), розділ II) потік вектора $\vec{D}$ крізь будь-яку замкнену поверхню

$$\oint\limits_S \vec{D}\,\mathrm{d}\vec{S} = q,$$

де q — вільний заряд всередині поверхні S, який в нашому випадку зосереджений на пластині конденсатора. Тому поле є лише між пластинами, і потік $\vec{D}$ створюється тільки через дно "глечика" S_2, так що

$$q = \int\limits_{S_2} \vec{D}\,\mathrm{d}\vec{S}_2.$$

Відповідно, швидкість зміни заряду

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \int\limits_{S_2} \frac{\partial\vec{D}}{\partial t}\mathrm{d}\vec{S}_2.$$

Порівнюючи цей вираз із (2.2б), доходимо висновку, що

$$\int\limits_{S_1} \vec{j}\,\mathrm{d}\vec{S}_1 = \int\limits_{S_2} \frac{\partial\vec{D}}{\partial t}\mathrm{d}\vec{S}_2. \tag{2.3}$$

Це й вирішує "парадокс" рівняння (2.1а): якщо його записати у вигляді

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = \int\limits_{S_2} \frac{\partial\vec{D}}{\partial t}\mathrm{d}\vec{S}_2, \tag{2.4}$$

то результат обчислення циркуляції $\vec{H}$ за виразами (2.1) і (2.4) буде однаковий, незалежно від вибору поверхні. При цьому в одному випадку поверхню перетинає тільки струм, а в іншому — тільки електричне поле. Тому вирази (2.1) і (2.4) можна об'єднати в одне загальне рівняння

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = \int\limits_S \left(\vec{j} + \frac{\partial\vec{D}}{\partial t}\right)\mathrm{d}\vec{S}, \tag{2.5}$$

що є чинним при будь-якій формі поверхні S, обмеженої контуром L.

Треба сказати, що цей результат сам по собі має лише формальний зміст. Він отриманий тільки для окремого випадку — магнітного поля струму заряджання конденсатора — і пояснюється тим, що між струмом заряджання та електричним полем конденсатора є прямий зв'язок. По закінченню заряджання рівняння (2.5) втрачає зміст, бо струм припиняється ($\vec{j} = 0$), електричне поле у конденсаторі стає постійним ($\partial\vec{D}/\partial t = 0$), і магнітне поле зникає. Але Максвелл дійшов висновку, що *рівняння* (2.5) *є чинним завжди*, не втрачаючи змісту і при $\vec{j} = 0$. Якщо так, то при $(\partial\vec{D}/\partial t) \neq 0$

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = \int\limits_S \frac{\partial\vec{D}}{\partial t}\,\mathrm{d}\vec{S} \neq 0 \quad \Rightarrow$$
$$\Rightarrow \quad \vec{H} \neq 0. \tag{2.5а}$$

Тим самим Максвелл постулював, що

магнітне поле створюється не лише струмами провідності (рухом заряджених частинок), а й змінними електричними полями.

Слід зауважити, що розглянуті міркування і викладки не доводять із необхідністю такий умовивід, а тільки вказують на його вірогідність. Але його істинність, як і інші фундаментальні відкриття Максвела, незаперечно підтверджується дослідом і всім розвитком класичної електродинаміки.

Похідна $\partial\vec{D}/\partial t$, яка фігурує в рівняннях (2.5), (2.5а), як і густина струму провідності $\vec{j}$, має розмірність [А/м^2] і по аналогії називається "густиною струму зміщення" позначається таким самим символом:

$$\vec{j}_{\text{зм}} = \frac{\partial\vec{D}}{\partial t}. \tag{2.6}$$

Але цю термінологію і символіку не слід розуміти буквально, бо

струм зміщення є мірою швидкості зміни електричного поля

і в загальному випадку не пов'язаний із рухом зарядів-носіїв струму.

Уведення струму зміщення дозволяє зручно записувати і формулювати закон (2.5):

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = \int\limits_S (\vec{j} + \vec{j}_{\text{зм}})\mathrm{d}\vec{S}, \tag{2.7}$$

або

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = \int\limits_S \vec{j}_{\text{повн}}\,\mathrm{d}\vec{S}, \tag{2.7а}$$

де величина

$$\vec{j}_{\text{повн}} = \vec{j} + \vec{j}_{\text{зм}} = \vec{j} + \frac{\partial\vec{D}}{\partial t} \tag{2.8}$$

називається густиною повного струму. Відповідно, рівняння (2.7а) інколи називають законом повного струму.

Поняття струму зміщення дозволяє також інакше подати співвідношення (2.2б), що випливає з рівняння неперервності (2.2):

$$\oint\limits_S \vec{j}\,\mathrm{d}\vec{S} = -\oint\limits_S \vec{j}_{\text{зм}}\,\mathrm{d}\vec{S} \quad \Rightarrow$$
$$\Rightarrow \quad \oint\limits_S (\vec{j} + \vec{j}_{\text{зм}})\mathrm{d}\vec{S} = 0 \quad \Rightarrow$$
$$\Rightarrow \quad \oint\limits_S \vec{j}_{\text{повн}}\mathrm{d}\vec{S} = 0.$$

Це означає, що лінії повного струму є завжди неперервні — в точках розриву провідності кожна лінія струму провідності (лінія вектора $\vec{j}$), продовжується лінією струму зміщення (лінією вектора $\partial\vec{D}/\partial t$) і навпаки. (У розглянутому прикладі це відбувається на обкладках конденсатора.) Саме з цієї причини рівняння (2.6), розділ III, й називається рівнянням неперервності.

2.2. *Рівняння Максвелла в інтегральній формі*

Відкриття вихрового електричного поля та струму зміщення як джерела магнітного поля лягло в основу створеної Максвеллом класичної електродинаміки — теорії, що трактує макроскопічні електромагнітні явища при нерелятивістських ($v \ll c$) швидкостях руху зарядів.

Фундаментальні рівняння. Найбільш загальні властивості електромагнітного поля відображаються системою фундаментальних рівнянь електромагнітного поля, або рівнянь Максвелла. Вони є найбільш загальними виразами розглянутих у попередніх розділах теорем про циркуляції та потоки електричного і магнітного полів і мають вигляд:

$$\begin{aligned} &\oint\limits_L \vec{E}\,\mathrm{d}\vec{l} = -\int\limits_S \frac{\partial \vec{B}}{\partial t}\,\mathrm{d}\vec{S};\\ &\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = \int\limits_S \left(\vec{j} + \frac{\partial \vec{D}}{\partial t}\right)\mathrm{d}\vec{S};\\ &\oint\limits_S \vec{D}\,\mathrm{d}\vec{S} = \int\limits_V \rho\,\mathrm{d}V;\\ &\oint\limits_S \vec{B}\,\mathrm{d}\vec{S} = 0. \end{aligned} \tag{2.9}$$

Зауважимо, що ця система не включає рівняння неперервності, яке виражає закон збереження електричного заряду. Але його неявно відображає неперервність ліній повного струму, про що говорилося вище.

У випадку стаціонарних полів, для яких $(\partial \vec{D}/\partial t) = (\partial \vec{B}/\partial t) = 0$, система (2.9) розпадається на дві пари незалежних рівнянь електростатики та магнітостатики, котрі визначають найголовніші властивості стаціонарних полів:

$$\begin{aligned} &\oint\limits_L \vec{E}\,\mathrm{d}\vec{l} = 0;\\ &\oint\limits_S \vec{D}\,\mathrm{d}\vec{S} = \int\limits_V \rho\,\mathrm{d}V, \end{aligned}$$

і

$$\begin{aligned} &\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = \int\limits_S \vec{j}\,\mathrm{d}\vec{S};\\ &\oint\limits_S \vec{B}\,\mathrm{d}\vec{S} = 0. \end{aligned}$$

Перша пара відображає той факт, що статичні електричні поля створюються зарядженими частинками і є потенціальними, а друга — що статичні магнітні поля створюються тільки струмами провідності (рухом зарядів) і є непотенціальними (вихровими).

У випадку нестатичних (змінних у часі) полів із загальної системи фундаментальних рівнянь випливає, що джерелом якогось одного поля є не лише заряди або струми, а й інше поле. Тож нестаціонарні електричне і магнітне поля є органічно взаємопов'язані, й одне не може існувати без іншого. Змінне електричне поле породжує магнітне, котре в свою чергу породжує електричне, а воно — знов магнітне, і т.д. При цьому, як показав Максвелл, із рівнянь (2.9) випливає, що цей процес має хвильовий характер і поширюється у просторі зі швидкістю світла. Тим самим Максвелл теоретично відкрив існування електромагнітних хвиль та електромагнітну природу світла, що потім було підтверджено експериментально.

Матеріальні рівняння. Рівняння Максвелла виконуються за будь-яких умов, на те вони й фундаментальні. Але в математичному сенсі система рівнянь (2.9) є неповною. Вона містить два векторні рівняння для циркуляцій і два скалярні для потоків. При записі в проекціях це дає 8 рівнянь, тоді як навіть при заданих функціях ρ і $\vec{j}$ для визначення чотирьох векторів полів треба мати 12 рівнянь.

Причина такої недостатньості фундаментальних рівнянь полягає в тому, що вони не містять електричних і магнітних характеристик середовища, котрі саме й визначають перебіг електромагнітних процесів у кожному конкретному випадку. Тому фундаментальні рівняння мають бути доповнені матеріальними рівняннями — співвідношеннями, що відображають зв'язки між різними електричними та магнітними характеристиками поля в речовині ("матеріалі"). Ці співвідношення істотно відрізняються для різних класів речовин, тому для кожного є свої матеріальні рівняння. Для ізотропних лінійних речовин, які є найбільш поширені і вживані на практиці, матеріальні рівняння мають знайомий з попередніх розділів вигляд:

$$\vec{D} = \varepsilon_0 \varepsilon \vec{E};$$
$$\vec{B} = \mu_0 \mu \vec{H};$$
$$\vec{j} = \sigma \vec{E},$$

де ε, μ — діелектрична та магнітна проникності, а σ — питома електропровідність речовини.

Фундаментальні рівняння разом із матеріальними співвідношеннями утворюють повну систему рівнянь електромагнітного поля в нерухомому середовищі, котра в кожному випадку принципово дозволяє дослідити будь-яке електромагнітне явище.

2.3 *Диференціальні характеристики поля*

Розглянуті інтегральні рівняння (2.9) встановлюють зв'язок між джерелами і характеристиками електромагнітного поля через потоки й циркуляції, тож відображують його властивості не в кожній точці (локально), а лише інтегрально — на заданій лінії або поверхні. Через це вони не дають принципової відповіді на деякі важливі запитання. До прикладу, нульова величина циркуляції напруженості заданого електричного поля по обраному контуру L *не може бути доказом* його потенціальності, бо не виключає існування в цьому полі інших контурів, для котрих

$$\oint\limits_L \vec{E}\,\mathrm{d}\vec{l} \neq 0.$$

Так само, коли сума струмів, які охоплює даний замкнений контур L, дорівнює нулю, то за відсутності змінного електричного поля з другого рівняння (2.9) випливає, що

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = 0.$$

Та це зовсім не означає, що магнітне поле є потенціальним.

Але такі питання, у принципі, не виникають при використанні рівнянь Максвелла в *диференціальній формі*, котрі встановлюють зв'язок між джерелами та характеристиками поля локально, тобто в кожній точці. Ці рівняння оперують величинами, що визначають не саме поле, а його поведінку в кожній точці. Їхній зміст , розглянуто в питаннях:

2.3.1. Дивергенція поля

2.3.2. Ротор поля

2.3.1. *Дивергенція поля*

Ця величина принципово вирішує задачу визначення розподілу зарядів і струмів за заданими характеристиками створюваного ними поля, що є важливим і в теорії, і на практиці. Як це робиться, розглянемо на прикладі електричного поля.

Третє рівняння Максвелла (2.9) дозволяє знайти сумарний електричний заряд у будь-якій області простору за відомим електричним зміщенням $\vec{D}(\vec{r})$. Але так, у принципі, можна визначити не лише величину заряду в заданій області, а й його просторовий розподіл. Для цього будемо діяти за наступною логічною схемою.

Розчленуємо подумки все поле на маленькі комірки і підрахуємо потік вектора $\vec{D}$ крізь кожну. Тоді там, де він дорівнює нулю, заряду немає, а в комірках із ненульовим потоком заряд є, причому його величина дорівнює величині потоку. Тому, зменшуючи розміри комірок, у такий спосіб можна врешті решт отримати достатньо точну інформацію про заряд, зосереджений у кожній точці поля.

Запропонована позірна схема може здатися практично нездійсненою через гранично велику необхідну кількість указаних комірок. Але насправді вона досить просто реалізується методами вищої математики.

Визначимо потік вектора $\vec{D}$ крізь якусь одну комірку у формі невеликого паралелепіпеда з ребрами Δx, Δy, Δz , паралельними до осей координат, як показано на рис. 2.2.

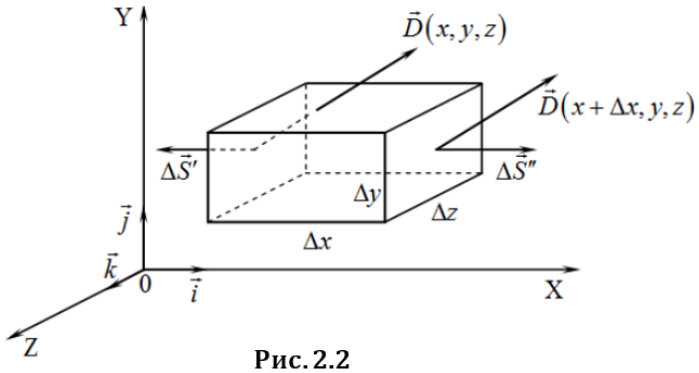


Рис. 2.2

Для цього подамо його як суму потоків крізь пари паралельних граней:

$$\oint_S \vec{D}\,\mathrm{d}\vec{S} = \Phi_1 + \Phi_2 + \Phi_3.$$

При малих розмірах комірок величину $\vec{D}$ в усіх точках заданої грані можна наближено вважати однаковою. Тоді потік крізь одну грань

$$\Delta\Phi \approx \vec{D}\,\Delta\vec{S},$$

де ΔS — площа грані.

Відповідно, для потоку крізь пару граней з координатами x і $x + \Delta x$ матимемо:

$$\Phi_1 \approx \vec{D}(x,y,z)\Delta\vec{S}_1' + \vec{D}(x+\Delta x,y,z)\Delta\vec{S}_1''.$$

Вектори площинок

$$\Delta\vec{S}_1' = -\vec{i}\,\Delta y\,\Delta z \qquad \text{і} \qquad \Delta\vec{S}_1'' = \vec{i}\,\Delta y\,\Delta z,$$

отож,

$$\Phi_1 \approx \big(-D_x(x,y,z) + D_x(x+\Delta x,y,z)\big)\Delta y\,\Delta z = \\ = \Delta D_x\,\Delta y\,\Delta z.$$

Величина ΔD_x визначається відомим з математики виразом

$$\Delta D_x \approx \frac{\partial D_x}{\partial x}\Delta x.$$

Примітка. Частинна похідна стоїть тому, що зміна вектора зміщення зумовлена зміною тільки координати x.

Відтак для потоку крізь дану пару граней маємо

$$\Phi_1 \approx \frac{\partial D_x}{\partial x}\Delta x\,\Delta y\,\Delta z \quad \Rightarrow \\ \Rightarrow \quad \Phi_1 \approx \frac{\partial D_x}{\partial x}\Delta V,$$

де ΔV — об'єм комірки.

Очевидно, що аналогічно виражаються потоки і крізь дві інші пари граней:

$$\Phi_2 \approx \frac{\partial D_y}{\partial y}\Delta V, \qquad \Phi_3 \approx \frac{\partial D_z}{\partial z}\Delta V.$$

Отже, потік крізь усю поверхню комірки наближено виражається формулою:

$$\oint\limits_S \vec{D}\,\mathrm{d}\vec{S} \approx \\ \approx \left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}\right)\Delta V. \tag{2.11}$$

Вираз у дужках часто стрічається у різних задачах математики та фізики. Тому він має спеціальну назву “дивергенція” поля і позначення “div”. Отже,

$$\operatorname{div}\vec{D} = \left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}\right). \tag{2.12}$$

Зауважимо, що залежно від симетрії поля можуть виявитися зручнішими сферичні або циліндричні координати, в яких формули дивергенції при збереженні загальної структури відрізняються від виразу (2.12). Тому він є не означенням, а тільки формулою для обчислення дивергенції в *декартових координатах*. А інваріантне означення, що висвітлює фізичний зміст дивергенції, випливає з виразу (2.11) при переході до границі $\Delta V \to 0$ шляхом стягання комірки в точку:

$$\oint\limits_S \vec{D}\,\mathrm{d}\vec{S} \approx \operatorname{div}\vec{D}\cdot\Delta V \quad \Rightarrow \\ \Rightarrow \quad \operatorname{div}\vec{D} = \lim_{\Delta V\to 0}\frac{\oint \vec{D}\,\mathrm{d}\vec{S}}{\Delta V}. \tag{2.13}$$

Таким чином,

> ***дивергенцією векторного поля в заданій точці називається границя відношення потоку поля крізь замкнену поверхню до обмеженого нею об’єму при стяганні поверхні в задану точку.***

Ураховуючи формулу (2.12) і теорему Гаусса для вектора $\vec{D}$, вираз (2.11) можна подати, як

$$\operatorname{div}\vec{D}\cdot\Delta V \approx q = \langle\rho\rangle\Delta V,$$

де $\langle\rho\rangle$ — середня густина заряду в об’ємі ΔV. Очевидно, що при $\Delta V \to 0$ ця рівність стане точною. Таким чином, у кожній точці поля

$$\operatorname{div}\vec{D} = \rho. \tag{2.14}$$

Позаяк цей результат випливає з інтегральної теореми Гаусса, рівняння (2.14) називається теоремою Гаусса в диференціальній формі. Воно встановлює локальний (у кожній точці) зв’язок між полем і його джерелами, й вирішує проблему визначення розподілу заряду $\rho(x, y, z)$, яким створюється задане поле. Для цього достатньо провести в (2.14) необхідне диференціювання заданої функції $\vec{D}(x, y, z)$.

Якщо в рівнянні ((4.9), розділ II) зробити заміну (2.14), то отримаємо вираз математичної *теореми Остроградського — Гаусса*

$$\oint\limits_S \vec{D}\,\mathrm{d}\vec{S} = \int\limits_V \operatorname{div}\vec{D}\cdot\mathrm{d}V. \tag{2.15}$$

А саме,

> ***потік векторного поля крізь довільну замкнену поверхню дорівнює інтегралу від дивергенції цього поля по об’єму обмеженому даною поверхнею.***

Формула (2.15) є чинною для будь-якого векторного поля, зокрема, для магнітного. Тож

$$\oint\limits_S \vec{B}\,\mathrm{d}\vec{S} = \int\limits_V \operatorname{div}\vec{B}\cdot\mathrm{d}V.$$

Але, згідно з інтегральною теоремою Гаусса (10.1), ліва частина записаного рівняння дорівнює нулю при будь-яких розмірах і формі замкненої поверхні S. Тож і права частина дорівнює нулю при будь-якій величині V, а це можливо тільки за умови, що

> ***в будь-якій точці магнітного поля дивергенція вектора $\vec{B}$ дорівнює нулю:***

$$\operatorname{div} \vec{B} = 0. \tag{2.16}$$

Це твердження виражає **теорему Гаусса для магнітного поля в диференціальній (*локальній*) формі.**

2.3.2. *Ротор поля*

Теорему про циркуляцію векторного поля теж можна виразити в диференціальній формі. Покажемо це на прикладі магнітного поля струмів провідності.

Нехай магнітне поле $\vec{H}(x, y, z)$ створюється заданим розподілом струмів провідності $\vec{J}(x, y, z)$.

Примітка. Позначення густини струму збігається з позначенням одного з ортів декартової системи координат. Тому, аби запобігти плутанині, далі до кінця п. 2.3 густина струму позначається як $\vec{J}$.

Виберемо невеликий прямокутний контур із сторонами Δx і Δy, розташований у площині XOY, як показано на рис. 2.3,

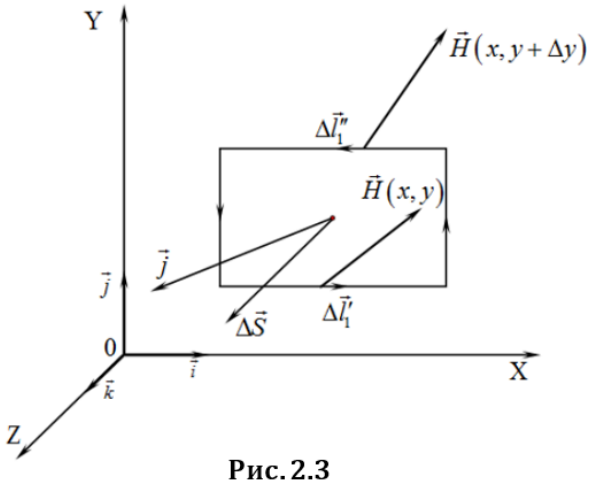


Рис. 2.3

і запишемо для нього наближений вираз теореми про циркуляцію вектора $\vec{H}$, взявши до уваги, що орт нормалі $\vec{n}$ збігається з ортом $\vec{k}$ координатної системи:

$$\begin{aligned} \oint_L \vec{H}\, \mathrm{d}\vec{l} \approx \vec{J}\, \Delta\vec{S} = J_n\, \Delta S = \\ = J_z\, \Delta x\, \Delta y. \end{aligned} \tag{2.17}$$

Циркуляція по вказаному контуру складається із внесків від усіх сторін прямокутника. Ці внески не важко визначити. Розглянемо для прикладу внесок C_1 від горизонтальних сторін, обходячи контур проти годинникової стрілки (у додатному напрямі відносно вектора $\vec{J}$). При малих розмірах прямокутника можна нехтувати зміною поля на переміщенні вздовж кожної із сторін і наближено записати:

$$\begin{aligned} C_1 \approx \vec{H}(x, y)\Delta\vec{l}_1' + \vec{H}(x, y+\Delta y)\Delta\vec{l}_1'' = \\ = \big(\vec{H}(x, y)\vec{i} - \vec{H}(x, y+\Delta y)\vec{i}\big)\Delta x = \\ = -(\Delta H_x)_y \Delta x. \end{aligned}$$

Тут ураховано, що $\Delta\vec{l}_1' = \vec{i}\Delta x$, $\Delta l_1'' = -\vec{i}\Delta x$ і, що скалярний добуток вектора на орт є проекцією цього вектора на відповідну вісь. У такому разі величина $(\Delta H_x)_y$ є частинним приростом величини H_x при зміні координати y при переході до верхньої сторони контуру. За правилами диференціального числення

$$(\Delta H_x)_y \approx \frac{\partial H_x}{\partial y}\Delta y,$$

отже, розглянутий внесок у циркуляцію складає

$$C_1 \approx -\frac{\partial H_x}{\partial y}\Delta x\, \Delta y.$$

Аналогічно можна показати, що вклад від іншої пари сторін

$$C_2 \approx \left(\vec{H}(x+\Delta x, y)\vec{j} - \vec{H}(x,y)\vec{j}\right)\Delta y = \\ = (\Delta H_y)_x \Delta y \approx \frac{\partial H_y}{\partial x}\Delta x\,\Delta y.$$

Відповідно, циркуляція по всьому контуру $C = C_1 + C_2$ дорівнює

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} \approx \\ \approx \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\Delta y\,\Delta x = \\ = \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\Delta S. \tag{2.17a}$$

Порівнюючи цей вираз із (2.17), бачимо, що

$$\left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\Delta S \approx J_z\,\Delta S.$$

Отримана рівність є наближеною, тому що в ній не врахована зміна проекцій поля H_x і H_y при переміщенні вздовж сторін прямокутника в процесі інтегрування, а також не визначене точне положення точки, в якій береться величина $\vec{J}$. Але в границі $\Delta S \to 0$ при стяганні прямокутника в задану точку рівність стає точною:

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = J_z.$$

Вибравши в околі тієї ж точки ще два контури, паралельні іншим координатним площинам, аналогічно отримаємо:

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = J_x,$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = J_y.$$

Вектор $\vec{J}$ через проекції та орти зображується, як

$$\vec{J} = J_x \cdot \vec{i} + J_y \cdot \vec{j} + J_z \cdot \vec{k}.$$

Тож маємо наступний вираз, який визначає вектор густини струму $\vec{J}$ в будь-якій точці заданого поля $\vec{H}(x,y,z)$:

$$\vec{J} = \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right)\vec{i} + \\ + \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\right)\vec{j} + \\ + \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\vec{k}.$$

Подібний громіздкий набір операцій доводиться виконувати в багатьох задачах математики і фізики. Тому права частина даного виразу має спеціальне позначення “rot” і назву “ротор” поля. Отже,

$$\operatorname{rot}\vec{H} = \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right)\vec{i} + \\ + \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\right)\vec{j} + \\ + \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\vec{k}. \tag{2.18}$$

Відтак зв’язок між векторами $\vec{H}$ і $\vec{J}$ згорнуто записується у вигляді:

$$\operatorname{rot}\vec{H} = \vec{J}. \tag{2.19}$$

За цим рівнянням для визначення розподілу струмів $\vec{J}(x, y, z)$, які створюють задане магнітне поле, треба лише виконати необхідні диференціювання заданої функції $\vec{H}(x, y, z)$. За допомогою рівняння (2.19), у принципі, можна визначити й поле, створене заданим розподілом струмів. Але така математична задача є досить складною і має точні розв'язки тільки в найпростіших випадках.

Зауважимо, що рівняння (2.19) є чинним і в інших системах координат, але формула ротора в кожній з них має свій вигляд. Тому вираз (2.18) є не означенням, а тільки формулою для обчислення ротора в декартових координатах. Аби дати незалежне від координатної системи означення ротора, звернемо увагу на те, що в наведених викладках орти $(\vec{i}, \vec{j}, \vec{k})$ є й одиничними векторами нормалей $\vec{n}$ до поверхонь розглянутих контурів. Тож у (2.18) множники при ортах є проекціями вектора $\operatorname{rot}\vec{H}$ на нормалі до елементарних контурів розміщених у координатних площинах. Враховуючи це, вираз (2.17а) можна переписати, як

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} \approx \operatorname{rot}_n\vec{H}\cdot\Delta S \quad \Rightarrow$$

$$\Rightarrow \quad \operatorname{rot}_n\vec{H} \approx \frac{\oint\limits_L \vec{H}\,\mathrm{d}\vec{l}}{\Delta S}.$$

Стягаючи контур у задану точку, отримаємо точний вираз:

$$\operatorname{rot}_n\vec{H} = \lim_{\Delta S\to 0}\frac{\oint\limits_L \vec{H}\,\mathrm{d}\vec{l}}{\Delta S},$$

який по суті є означенням поняття "ротор векторного поля".

Якщо магнітне поле створюється не лише струмами провідності, а й змінним електричним полем, у рівняння (2.19) додасться ще й густина струму зміщення (2.6):

$$\operatorname{rot}\vec{H} = \vec{j} + \frac{\partial\vec{D}}{\partial t}. \tag{2.20}$$

Це рівняння випливає з інтегральної теореми про циркуляцію магнітного поля і є загальним виразом цієї теореми в диференціальній (локальній) формі.

Підставивши вираз (2.20) у рівняння (2.9), отримаємо математичну теорему Стокса

$$\oint\limits_L \vec{H}\,\mathrm{d}\vec{l} = \int\limits_S \operatorname{rot}\vec{H}\cdot\,\mathrm{d}\vec{S}, \tag{2.21}$$

згідно з якою

циркуляція векторного поля по довільному замкненому контуру дорівнює потокові ротора цього поля крізь довільну поверхню обмежену цим контуром.

Теорема Стокса є чинною для будь-якого векторного поля, тож і для електричного:

$$\oint\limits_L \vec{E}\,\mathrm{d}\vec{l} = \int\limits_S \operatorname{rot}\vec{E}\cdot\mathrm{d}\vec{S}.$$

Порівнюючи цей вираз із (12.17а), отримуємо рівняння

$$\operatorname{rot}\vec{E} = -\frac{\partial\vec{B}}{\partial t}, \tag{2.22}$$

яке виражає теорему про циркуляцію електричного поля в диференціальній (локальній) формі.

У випадку, коли електричне поле створюється тільки зарядами і є потенціальним, циркуляція вектора по будь-якому контуру дорівнює нулю (розділ I, п. 5.1). Тому, згідно з формулою Стокса, в усіх точках

$$\operatorname{rot}\vec{E} = 0. \tag{2.23}$$

Цей вираз є диференціальним (локальним) критерієм потенціальності поля. Відповідно до нього потенціальні поля є безвихровими.

Локальний критерій (2.23) є дуже зручним: за ним потенціальність заданого поля $\vec{E}(x, y, z)$ однозначно встановлюється за допомогою не складної операції диференціювання.

2.4. *Рівняння Максвелла в локальній формі*

Розглянуті результати дозволяють описати фундаментальні властивості електромагнітного поля не лише інтегральними рівняннями Максвелла (2.9), а й відповідною системою диференціальних рівнянь, які відображають локальні (в кожній точці) властивості поля. Далі розглядаються:

Диференціальні рівняння Максвелла

Оператор $\vec{\nabla}$

Операторна форма рівнянь Максвелла

Диференціальні рівняння Максвелла. З використанням понять “дивергенція” і “ротор” основні рівняння електромагнітного поля мають вигляд:

$$\begin{aligned} \operatorname{rot} \vec{E} &= -\frac{\partial \vec{B}}{\partial t}; \\ \operatorname{rot} \vec{H} &= \vec{j} + \frac{\partial \vec{D}}{\partial t}; \\ \operatorname{div} \vec{D} &= \rho; \\ \operatorname{div} \vec{B} &= 0. \end{aligned}$$

Відповідно, рівняння електростатики та магнітостатики (коли поля не залежать від часу) в диференціальній формі мають вигляд:

$$\begin{aligned} \operatorname{rot} \vec{E} &= 0, \qquad \operatorname{rot} \vec{H} = \vec{j}, \\ \operatorname{div} \vec{D} &= \rho, \qquad \operatorname{div} \vec{B} = 0. \end{aligned}$$

Оператор $\vec{\nabla}$. Дивергенція і ротор векторів $\vec{E}$ і $\vec{H}$, а також розглянутий раніше (розділ I, п. 3) градієнт потенціалу φ, характеризують у кожній точці не саме поле, а його поведінку при переміщенні в прилеглі точки. Тому для визначення цих величин, над заданою функцією (φ, $\vec{E}$, $\vec{H}$, або ін.) треба виконати відповідний набір математичних дій або “операцій” ((3.4), розділ I), (2.12), або (2.18). Самі по собі вони не є складним, але при необхідності виконання комбінованих операцій, на кшталт $\operatorname{div}(\operatorname{rot} \vec{B})$ чи $\operatorname{rot}(\operatorname{rot} \vec{E})$, прямі викладки є трудомісткими. Проте існує дуже ефективний уніфікований непрямий спосіб обчислень за допомогою векторного диференціального оператора набла $\vec{\nabla}$, інакше — ***оператора Гамільтона***.

На загал, термін “оператор” має різні сенси. Тут це символ, яким позначають дію (“операцію”) чи набір дій, які належить виконати над однією функцією, аби перетворити її на деяку іншу функцію. Оператори дозволяють не лише лаконічно записувати громіздкі математичні вирази, як приміром (2.12) чи (2.18), а й значно скорочують і полегшують математичні перетворення. Це можна проілюструвати на прикладі обчислення градієнта та дивергенції. Для обчислення градієнта потенціалу електричного поля за формулою (розділ I, (3.4)) треба:

— взяти похідні по координатах від заданої функції $\varphi(x, y, z)$;

— помножити їх на орти $(\vec{i}, \vec{j}, \vec{k})$;

— додати утворені складові вектори.

Ця стандартна послідовність математичних операцій являє собою оператор градієнта і позначається символом grad. Аналогічно можна деталізувати й сукупність операцій, які складають оператор дивергенції і шифруються символом div. До прикладу, для визначення дивергенції вектора $\vec{D}$ електричного поля треба за формулою (3.12) виконати наступну послідовність операцій:

— взяти похідні по координатах від відповідних проекцій заданої функції $\vec{D}(x, y, z)$;

— додати їх.

Але, як виявляється, необхідні для визначення обох цих величин дії можна виконувати за допомогою одного оператора, який позначається символом $\vec{\nabla}$. Це можна побачити, якщо у виразах градієнта та дивергенції дещо інакше записати частинні похідні по координатах, увівши “оператори диференціювання” $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$:

$$\operatorname{grad}\varphi = \vec{i}\frac{\partial}{\partial x}\varphi + \vec{j}\frac{\partial}{\partial y}\varphi + \vec{k}\frac{\partial}{\partial z}\varphi$$

і

$$\operatorname{div}\vec{D} = \frac{\partial}{\partial x}D_x + \frac{\partial}{\partial y}D_y + \frac{\partial}{\partial z}D_z.$$

При цьому перша формула виглядає, як вираз деякого вектора $\vec{G}$, що є добутком вектора $\vec{A}$ на скаляр B:

$$\vec{G} = \vec{A}\cdot B = (\vec{i}A_x)B + (\vec{j}A_y)B + (\vec{k}A_z)B.$$

Відповідно, формула для дивергенції нагадує скалярний добуток S якихось векторів $\vec{A}$ і $\vec{C}$:

$$S = \vec{A}\cdot\vec{C} = A_xC_x + A_yC_y + A_zC_z.$$

Це дозволяє увести оператор $\vec{\nabla}$ у вигляді *символічного вектора*

$$\vec{\nabla} = \vec{i}\frac{\partial}{\partial x} + \vec{j}\frac{\partial}{\partial y} + \vec{k}\frac{\partial}{\partial z}, \tag{2.25}$$

і розглядати градієнт як результат символічного множення (тобто дії) “вектора” $\vec{\nabla}$ на скалярну функцію координат φ:

$$\operatorname{grad}\varphi \equiv \vec{\nabla}\varphi. \tag{2.26}$$

Аналогічно, дивергенцію можна трактувати як скалярний добуток векторів $\vec{\nabla}$ і $\vec{D}$:

$$\operatorname{div}\vec{D} \equiv \vec{\nabla}\cdot\vec{D}. \tag{2.27}$$

Можна показати, що й ротор (2.18) можна трактувати як символічний векторний добуток:

$$\operatorname{rot}\vec{H} \equiv [\vec{\nabla}\,\vec{H}]. \tag{2.28}$$

У цьому легко переконатись, якщо розкрити вираз (2.28) за допомогою визначника:

$$[\vec{\nabla}\vec{H}] = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ (\partial/\partial x) & (\partial/\partial y) & (\partial/\partial z) \\ H_x & H_y & H_z \end{vmatrix},$$

як це робиться у векторній алгебрі.

Таким чином, всі основні диференціальні операції теорії поля можна зображувати і виконувати за допомогою одного диференціального векторного оператора $\vec{\nabla}$, властивості котрого одночасно задовольняють і основні правила диференціювання, і аксіоми векторної алгебри. Тому використання оператора $\vec{\nabla}$ є дуже продуктивним, позаяк значно спрощує проведення громіздких диференціальних перетворень скалярних і векторних функцій координат, якими зокрема є характеристики електромагнітного поля.

Операторна форма рівнянь Максвелла. На завершення зведемо в одну систему фундаментальні рівняння електромагнітного поля (рівняння Максвелла), виражені в операторній формі через оператор $\vec{\nabla}$:

$$\begin{aligned} [\vec{\nabla}\vec{E}] &= -\frac{\partial\vec{B}}{\partial t}, \\ [\vec{\nabla}\vec{H}] &= \vec{j} + \frac{\partial\vec{D}}{\partial t}, \\ \vec{\nabla}\cdot\vec{D} &= \rho, \\ \vec{\nabla}\cdot\vec{B} &= 0. \end{aligned} \tag{2.29}$$

У такому вигляді вони найбільш зручні для проведення різноманітних викладок. Для ілюстрації наведемо два приклади.

Приклад 1. Рівняння Пуассона. Електричні поля зарядів (кулонівські поля) є потенціальними, тому вектор напруженості $\vec{E}$ можна визначати не тільки через заряди, а й через скалярну величину — потенціал φ. Тому так само, як у випадку напруженості, важливо знати зв'язок між розподілом заряду та потенціалом поля в кожній точці. З інтегральних рівнянь Максвелла (2.9) цей зв'язок безпосередньо не випливає. Але, враховуючи, що

$$\vec{E} = -\operatorname{grad}\varphi \equiv -\vec{\nabla}\varphi,$$

його не складно отримати з диференціальних рівнянь (2.24) чи (2.29), принаймні, для ізотропних однорідних середовищ, в яких

$$\vec{D} = \varepsilon_0 \varepsilon \vec{E}.$$

Справді, зробивши наведені підстановки в треті рівняння (2.24), або (2.29), отримаємо: $\operatorname{div}\vec{D} = \rho$,

$$\operatorname{div}(\operatorname{grad}\varphi) = -\frac{\rho}{\varepsilon_1 \varepsilon},$$

або

$$\vec{\nabla}(\vec{\nabla}\varphi) = -\frac{\rho}{\varepsilon_0 \varepsilon}.$$

Ліву частину цього рівняння легко перетворити, згадавши, що до оператора $\vec{\nabla}$ можна застосовувати правило скалярного множення векторів і записати:

$$\vec{\nabla}(\vec{\nabla}\varphi) = \nabla_x^2\varphi + \nabla_y^2\varphi + \nabla_z^2\varphi,$$

де за правилами операторного множення

$$\begin{aligned}
\nabla_x^2\varphi &= \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\varphi\right) = \frac{\partial^2}{\partial x^2}\varphi, \\
\nabla_y^2\varphi &= \frac{\partial^2}{\partial y^2}\varphi, \\
\nabla_z^2\varphi &= \frac{\partial^2}{\partial z^2}\varphi.
\end{aligned}$$

Отже,

$$\frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} = -\frac{\rho}{\varepsilon_0}\varepsilon. \tag{2.30}$$

Це рівняння називається рівнянням Пуассона, і відіграє важливу роль в теорії електричного поля.

Рівняння Пуассона можна записати в операторній формі через оператор ∇^2 (набла квадрат) або, як його ще називають, оператор Лапласа і позначають також символом Δ (дельта):

$$\nabla^2 \equiv \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Множення (тобто дія) цього оператора на функцію полягає у взятті та додаванні її других частинних похідних по координатах. Відтак рівняння Лапласа в операторній формі згорнуто записується, як

$$\nabla^2\varphi = -\frac{\rho}{\varepsilon_0\varepsilon}, \quad \text{або} \quad \delta\varphi = -\frac{\rho}{\varepsilon_0\varepsilon}. \tag{2.30а}$$

Приклад 2. Рівняння неперервності. Це рівняння встановлює зв'язок між густиною струму на замкненій поверхні та зарядом носіїв всередині неї і є узагальненим виразом закону збереження електричного заряду. В інтегральній формі воно має вигляд ((1.6), розділ III):

$$\oint\limits_S \vec{j}\,\mathrm{d}\vec{S} = -\int\limits_V \frac{\partial\rho}{\partial t}\mathrm{d}V.$$

Рівняння неперервності в диференціальній формі визначає указаний зв'язок локально (в кожній точці) і дає детальну інформацію про структуру ліній струму. Таку форму рівняння неперервності можна отримати із наведеного інтегрального рівняння за допомогою формули Остроградського — Гаусса (2.15), відповідно

до якої

$$\oint\limits_S \vec{j}\,\mathrm{d}\vec{S} = \int\limits_V \mathrm{div}\,\vec{j}\cdot \mathrm{d}V.$$

Отже,

$$\int\limits_V \mathrm{div}\,\vec{j}\cdot \mathrm{d}V = -\int\limits_V \frac{\partial\rho}{\partial t}\mathrm{d}V.$$

Позаяк вихідні рівняння є чинними для довільної замкненої поверхні S, записана рівність виконується при будь-якому значенні об'єму V, що можливо лише коли в обох інтегралах значення підінтегральних функцій збігаються у кожній точці. Отже,

$$\mathrm{div}\,\vec{j} = -\frac{\partial\rho}{\partial t}. \tag{2.31}$$

Це і є рівняння неперервності в диференціальній формі. З нього випливає, що в тих точках, де $(\partial\rho/\partial t) < 0$, $\mathrm{div}\,\vec{j} > 0$, і навпаки. Згідно з (2.13) і домовленістю про напрям електричного струму, це означає, що лінії струму виходять із точок, з яких витікає додатній заряд, і закінчуються в точках, до яких він стікається. Тому в математиці точки, де дивергенція додатня, називаються витоками, а точки з від'ємною дивергенцією — стоками поля.

Рівняння (2.31) можна отримати і безпосередньо з рівнянь Максвелла в диференціальній формі. Для цього достатньо взяти дивергенцію від лівої та правої частини другого рівняння системи (2.24):

$$\mathrm{div}\,\mathrm{rot}\,\vec{H} = \mathrm{div}\,\vec{j} + \mathrm{div}\,\frac{\partial\vec{D}}{\partial t}.$$

Безпосереднє розкриття лівої частини потребує громіздких перетворень, але за допомогою оператора $\vec{\nabla}$ це робиться досить легко. А саме,

$$\mathrm{div}\ \mathrm{rot}\,\vec{H} = \vec{\nabla}[\vec{\nabla}\vec{H}].$$

У лівій частині маємо символічний мішаний добуток векторів, який за законами векторної алгебри є інваріантним відносно циклічної перестановки множників. Тож можна записати:

$$\vec{\nabla}[\vec{\nabla}\vec{H}] = [\vec{\nabla},\vec{\nabla}]\vec{H} \equiv 0.$$

Відтак маємо:

$$\mathrm{div}\,\vec{j} + \mathrm{div}\,\frac{\partial\vec{D}}{\partial t} = 0.$$

Послідовність диференціювання по незалежних змінних (час і координати) є довільною, тож, урахувавши третє рівняння (2.9), отримуємо рівняння неперервності (2.31):

$$\begin{aligned}\mathrm{div}\,\vec{j} = -\,\mathrm{div}\,\frac{\partial\vec{D}}{\partial t} = -\frac{\partial}{\partial t}(\mathrm{div}\,\vec{D}) \quad \Rightarrow \\ \Rightarrow \quad \mathrm{div}\,\vec{j} = -\frac{\partial\rho}{\partial t}.\end{aligned}$$

В операторній формі воно має вигляд

$$\vec{\nabla}\cdot\vec{j} = -\frac{\partial\rho}{\partial t}. \tag{2.31a}$$

Контрольні запитання

1. Що свідчить про неповноту рівняння (2.1), яке виражає теорему про циркуляцію магнітного поля струмів провідності?

2. Що називається густиною струму зміщення? Чи може існувати струм зміщення у діелектрику та вакуумі?

3. У чому полягає закон повного струму?

4. В яких одиницях вимірюється густина струму зміщення $j_{зм}$? Якщо густина струму зміщення у вакуумі дорівнює 1 $нА/м^2$, то що це означає?

5. Запишіть рівняння Максвелла для електромагнітного поля у вакуумі за відсутності заряджених частинок.

6. Чи є повною система рівнянь Максвелла? Чому?

КОЛИВАННЯ І ХВИЛІ

Зміст

Національний технічний університет України

Вступ

Частина І. КОЛИВАННЯ
- І. Гармонічні коливання
- 2. Механічні гармонічні коливання
- 3. Загасаючі та вимушені механічні коливання
- 4. Вільні електромагнітні коливання
- 5. Вимушені гармонічні коливання в контурі
- 6. Змінний струм

Частина ІІ. ХВИЛІ
- 1. Загальні властивості гармонічних хвиль
- 2. Пружні хвилі
- 3. Електромагнітні хвилі

Вступ

Коливаннями називаються процеси, які багаторазово точно, або наближено повторюються у часі. Приклади таких процесів легко знайти і в природі, і в техніці. Коливальні процеси можуть мати різне походження. Зокрема, важливими типами коливань, які часто спостерігаються в природі і використовуються в техніці, є механічні та електромагнітні коливання. Механічними коливаннями, зокрема, зумовлений звук, а електромагнітними – світло. Як приклади коливань можна навести також добові чи сезонні зміни температури повітря, зміни видимої площі освітленої поверхні Місяця, тощо. Але незалежно від природи коливальні процеси мають спільні загальні властивості, й описуються аналогічними математичними рівняннями та співвідношеннями. Тому теорія коливальних і хвильових процесів різної природи становить єдиний самостійний розділ фізики – фізику коливань.

Загальні властивості коливань визначаються умовами виникнення та законом зміни їхніх характеристик у часі. При цьому всі коливання поділяють на вільні та вимушені. Вільні коливання виникають після одноразового виведення системи із рівноважного стану і далі існують самостійно. Вільні коливання можливі тільки в так званих коливальних системах або осциляторах, в яких при відхиленні від рівноважного стану виникають внутрішні процеси, що спричинюють періодичні зміни фізичних характеристик системи. Натомість вимушені коливання створюються постійним зовнішнім впливом і тому можливі в будь-якій системі.

Що до змін у часі, то найпростішими є гармонічні коливання — процеси, в яких зміна їхніх характеристик у часі визначається гармонічними функціями (синус і косинус).

Усе сказане стосується й хвиль — процесів поширення коливань у просторі з плином часу.

Загальні властивості гармонічних коливальних і хвильових процесів розглядаються в наступних частинах:

I. КОЛИВАННЯ

II. ХВИЛІ

Частина І. КОЛИВАННЯ

У цій частині посібника розглядаються гармонічні коливання. Важливість вивчення саме таких процесів зумовлена двома обставинами. По-перше, вони зустрічаються в природі й використовуються на практиці. А по-друге, як доводить математика, будь-який складний коливальний процес можна розглядати як суперпозицію (сукупність) гармонічних коливань, що є важливим для теорії.

Далі розглянуто так питання:

1. Гармонічні коливання

2. Механічні гармонічні коливання

3. Загасаючі та вимушені механічні коливання

4. Вільні електромагнітні коливання

5. Вимушені коливання в контурі

6. Змінний струм

І. Гармонічні коливання

Серед періодичних коливань для науки і практики найважливішими є гармонічні коливання, які описуються гармонічними функціями і мають прості властивості. Важливість саме таких процесів зумовлена двома обставинами. По-перше, реальні коливання і хвилі, з якими доводиться мати справу, часто є близькими до гармонічних. По-друге, будь-який складний коливальний процес, як доводить математика, можна розглядати як суперпозицію (сукупність) гармонічних коливань. Тому в цей розділ присвячено саме таким процесам. Далі розглянуто:

1. Рівняння гармонічних коливань
2. Зображення гармонічних функцій
3. Додавання гармонічних коливань
4. Диференціальне рівняння гармонічних коливань

Контрольні запитання

1. Рівняння гармонічних коливань

Як було сказано, гармонічні коливання якоїсь величини ξ відбуваються за законом синуса або косинуса, тож визначаються ***рівнянням***:

$$\xi(t) = A\cos\left(\frac{2\pi}{T}t + \varphi_0\right) \text{ або } \xi(t) = A\sin\left(\frac{2\pi}{T}t + \phi_0\right), \tag{1.1}$$

де $$\varphi_0 = \phi_0 + \frac{\pi}{2},$$

і визначаються трьома параметрами — амплітудою, періодом (або частотою) і фазою.

Амплітуда A — додатнє число, що дорівнює максимальному значенню коливної величини ξ, так що всі її можливі значення лежать в інтервалі $\pm A$. Іншими словами, амплітуда задає "розмах" коливань.

За рівнянням (1.1), величина ξ задовольняє умову $\xi(t + nT) = \xi(t)$, отже

період коливань T є найменшим проміжком часу, через який значення величини ξ та характер її зміни повністю повторюються.

Окрім періоду, повторюваність коливального процесу у часі характеризують іще й частотою. При цьому розрізняють лінійну та циклічну (колову) частоту.

Лінійна частота ν є величиною оберненою до періоду:

$$\nu = \frac{1}{T} \tag{1.2}$$

Вона дорівнює кількості коливань, які здійснюються за одну секунду і вимірюється в герцах (Гц). При частоті 1 Гц = 1 $с^{-1}$ за одну секунду здійснюється одне повне коливання.

Циклічна (колова) частота ω пов'язана з лінійною частотою та періодом співвідношеннями

$$\omega = 2\pi\nu = \frac{2\pi}{T} \tag{1.3}$$

і вимірюється в рад/с. (Часто пишуть 1/с або $с^{-1}$).

Використання частоти, замість періоду, дозволяє записати рівняння коливань більш компактно:

$$\xi(t) = A\cos(2\pi\nu t + \varphi_0) \text{ або } \xi(t) = \cos(\omega t + \varphi_0). \qquad (1.4)$$

Якщо коливання описуються функцією синуса, то

$$\xi(t) = A\sin(2\pi\nu t + \phi_0) \text{ або } \xi(t) = A\sin(\omega t + \phi_0). \qquad (1.4а)$$

Задля зручності в теорії найчастіше використовують частоту ω і відповідні рівняння (1.4) або (1.4а).

З рівнянь гармонічних коливань також видно, що при заданій амплітуді A величина $\xi(t)$ у кожен момент часу визначається аргументом гармонічної функції, що називається ***фазою коливань φ.*** Отже,

$$\varphi(t) = \omega t + \varphi_0 \qquad (1.5)$$

Відповідно, величина φ_0 називається ***початковою фазою.*** Одиницею фази є радіан (1 рад).

Слід зауважити, що термін “фаза” вживається не лишень у вказаному формальному сенсі, а й для словесної характеристики стадії коливального процесу або узгодженості коливань – згадаймо такі знайомі вислови, як “фази Місяця” або “коливання відбуваються в однакових (чи протилежних) фазах”.

2. Зображення гармонічних функцій

При розгляді декількох гармонічних коливань, які відбуваються одночасно, доводиться додавати відповідну кількість гармонічних функцій вигляду. Це потребує громіздких перетворень, а при великій кількості коливань може стати технічно нездійсненним без використання комп’ютера. Тому при аналізі коливань часто використовують штучні математичні прийоми, а саме:

2.1. Векторне зображення гармонічних функцій

2.2. Комплексне зображення гармонічних функцій

2.1. *Векторне зображення гармонічних функцій*

Цей спосіб ґрунтується на існуванні відповідності між гармонічними коливаннями та обертальним рухом. Нехай маємо вектор з модулем A, який обертається з кутовою швидкістю ω навколо свого початку (рис. 1.1) Якщо в початковий момент часу цей вектор складає з деякою віссю OX кут φ_0, то в довільний момент t кут між ним і віссю буде рівним

$$\varphi(t) = \omega t + \varphi_0.$$

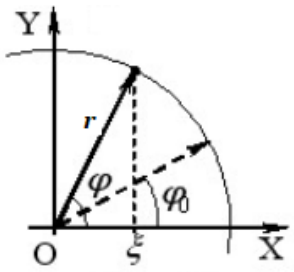


Рис. 1.1

Відповідно, якщо проекцію такого обертового вектора $A_x = A\cos\varphi$ на вісь OX позначити як ξ, то отримаємо:

$$\xi(t) = A\cos(\omega t + \varphi_0),$$

що збігається з (1.4), а спроєктувавши цей вектор на вісь OY, отримаємо рівняння (1.4а). Таким чином,

> ***кожному гармонічному коливанню можна поставити у відповідність вектор із модулем, рівним амплітуді коливань A, який обертається з кутовою швидкістю, рівною циклічній частоті коливань ω, і в початковий момент часу складає з вибраною віссю кут, рівний початковій фазі коливань φ_0.***

Доречі, саме через збіг величини ω із швидкістю обертання вказаного вектора вона й має назву "циклічної", або "колової" частоти коливань.

Таким чином, гармонічні коливання можна зображувати відповідними векторами на рисунку, який називається векторною діаграмою, і оперувати не безпосередньо гармонічними функціями, а векторами, що їх зображують.

2.2. *Комплексне зображення гармонічних функцій*

З математики відомо, що комплексне число $z = a + ib$ $(i = \sqrt{-1})$ можна записати або в показниковій формі, як $z = Ae^{i\varphi}$, або в тригонометричній формі у вигляді

$$z = A\cos\varphi + iA\sin\varphi, \qquad (1.6)$$

де модуль A та аргумент φ комплексного числа визначаються співвідношеннями

$$A = \sqrt{a^2 + b^2} \text{ і } \cos\varphi = \frac{a}{A}, \sin\varphi = \frac{b}{A}.$$

Тому, згідно з (1.6) і (1.4), рівняння гармонічних коливань $\xi(t)$ можна розглядати як дійсну частину комплексної функції

$$\tilde{\xi} = Ae^{\omega t+\varphi_0} = \tilde{A}e^{\omega t} \Rightarrow \xi(t) = Re\left(\tilde{\xi}(t)\right). \qquad (1.7)$$

При цьому комплексне число $\tilde{A} = Ae^{i\varphi_0}$ називається комплексною амплітудою коливань. Дійсна (істинна) амплітуда визначається через квадрат модуля комплексної амплітуди на основі тотожності

$$A^2 = \left|\tilde{A}^2\right| = \tilde{A}\cdot\tilde{A}^* = Ae^{i\varphi}\cdot Ae^{-i\varphi}. \qquad (1.8)$$

(*Нагадаємо, що спряженою до комплексної величини z називається величина z*, що створюється заміною уявної одиниці* ***i*** *у виразі z на* ***–i***.)

Указана відповідність (1.7) зберігається для всіх лінійних математичних операцій: множення на константу, додавання або віднімання, диференціювання, інтегрування. У цьому легко переконатися прямими обчисленнями. Через це розглянутий підхід є дуже продуктивним, оскільки оперувати з експонентами набагато зручніше, ніж із тригонометричними функціями. Зокрема, поширена задача про накладання декількох коливань однакової частоти та напрямку зводиться просто до додавання комплексних чисел (комплексних амплітуд).

3. Додавання гармонічних коливань

На початку говорилося про те, що складні коливальні процеси можна трактувати як суперпозицію (накладання), гармонічних коливань. Цей прийом широко застосовується в оптиці, техніці передачі сигналів та інших областях для аналізу та

формування сигналів необхідної форми. Розглянемо найпростіші випадки — додавання двох гармонічних коливань, які відбуваються в одному та у взаємно перпендикулярних напрямках.

3.1. Додавання коливань одного напрямку

3.2. Додавання взаємно перпендикулярних коливань

3.1. *Додавання коливань одного напрямку*

Розглянемо рух тіла (матеріальної точки), яке бере участь одночасно у двох гармонічних коливаннях одного напрямку. Прикладом можуть бути вертикальні коливання кульки на пружині, прикріпленій до стелі салону автомобіля, який теж коливається на ресорах у вертикальному напрямку. Рух кульки відносно землі складається з коливань відносно автомобіля та коливань автомобіля відносно землі. Відповідно, закон руху кульки буде визначатися сумою рівнянь цих двох коливань. Аналіз показує, що результат такого додавання суттєво залежить від того, однакові чи ні частоти складових коливань.

Частоти однакові. Нехай точка бере участь у двох рухах уздовж однієї осі OX відповідно до рівнянь:

$$\begin{aligned} x_1(t) &= A_1 \cos(\omega t + \alpha_1), \\ x_2(t) &= A_2 \cos(\omega t + \alpha_2). \end{aligned} \qquad (1.9)$$

Для визначення рівняння результуючого руху $x(t) = x_1(t) + x_2(t)$ скористаємося векторним способом зображення гармонічних функцій (1.9). Для цього зобразимо їх в момент t = 0 за допомогою векторів $\vec{A}_1$ і $\vec{A}_2$, що обертаються навколо осі O з кутовою швидкістю ω, як це описано в п. 2.1. На отриманій векторній діаграмі

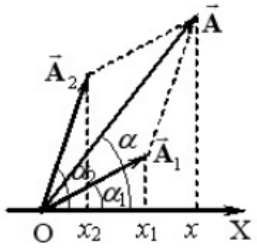


Рис. 1.2

(рис. 1.2) результуючий рух точки зображується вектором $\vec{A} = \vec{A}_1 + \vec{A}_2$, модуль якого за теоремою косинусів задається виразом:

$$A^2 = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\alpha_2 - \alpha_1), \qquad (1.10)$$

а кут з віссю OX, як випливає з рисунка:

$$\operatorname{tg}\alpha = \frac{A_1 \sin\alpha_1 + A_2 \sin\alpha_2}{A_1 \cos\alpha_1 + A_2 \cos\alpha_2}. \qquad (1.11)$$

Таким чином, результатом додавання двох гармонічних коливань однакової частоти та напрямку є гармонічні коливання тієї самої частоти, амплітуда A та початкова фаза α яких визначаються виразами (1.10) і (1.11). Слід зауважити, що при цьому амплітуда результуючих коливань не дорівнює сумі амплітуд коливань, які додаються. Залежно від початкових фаз α_1 і α_2 вона може мати будь-яке значення в межах від $A_1 + A_2$ до $|A_1 - A_2|$.

Частоти різні. Розглянемо тепер випадок, коли точка бере участь у двох коливаннях однакового напрямку, але різної частоти ω_1 і ω_2. Якщо для спрощення викладок прийняти $A_1 = A_2$ і $\alpha_1 = \alpha_2 = 0$, то рівняння цих коливань мають вигляд:

$$x_1 = A_1 \cos \omega_1 t, x_2 = A_1 \cos \omega_1 t. \qquad (1.12)$$

Рівняння результуючого руху точки $x_1(t) + x_2(t)$ запишемо, скориставшися відомою формулою тригонометрії:

$$x(t) = 2A \cos\left(\frac{\omega_1 - \omega_2}{2} t\right) \cos\left(\frac{\omega_1 + \omega_2}{2} t\right). \qquad (1.13)$$

Це рівняння описує складний коливальний процес, наочне уявлення про який можна отримати, побудувавши графік функції (1.13), або шляхом безпосереднього графічного додавання функцій (1.12). На рис. 1.3 показано результат такого додавання для випадку $\alpha = 0$ функцій

$$x_1 = A_1 \cos\left(\omega t - \frac{\pi}{2}\right), x_2 = A_1 \cos 2\omega t.$$

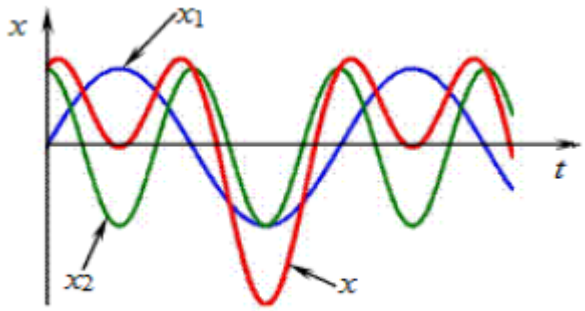


Рис. 1.3

Як для теорії, так і для практики окремий інтерес являє випадок, коли частоти ω_1 і ω_2 є близькими й відрізняються від певного значення ω на величину $\Delta\omega << \omega$ так, що $\omega_1 = \omega + \Delta\omega$ і $\omega_2 = \omega - \Delta\omega$. У такому разі рівняння (1.13) набуває вигляду:

$$x(t) = (2A_1 \cos \Delta\omega t) \cos \omega t. \qquad (1.14)$$

У цьому рівнянні величина $\cos\Delta\omega t$ змінюється з періодом $T = 2\pi/\Delta\omega$, тобто в $\omega/\Delta\omega$ разів повільніше, ніж величина $\cos \omega t$, період зміни якої $T = 2\pi/\omega$. При $\Delta\omega << \omega$ за час T_0 відбувається багато коливань величини x з частотою ω, як показано на рис. 1.4а для випадку $\Delta\omega = \omega/10$. Тому в (1.14) вираз у дужках можна трактувати як залежну від часу амплітуду $A(t)$ і розглядати рух точки як гармонічні коливання з частотою ω, амплітуда яких змінюється теж за гармонічним законом із частотою $\Delta\omega$:

$$x(t) = A(t) \cos \omega t, A(t) = 2A|\cos \Delta\omega t|. \qquad (1.14а)$$

Графік $A(t)$ показаний на рис. 1.4б.

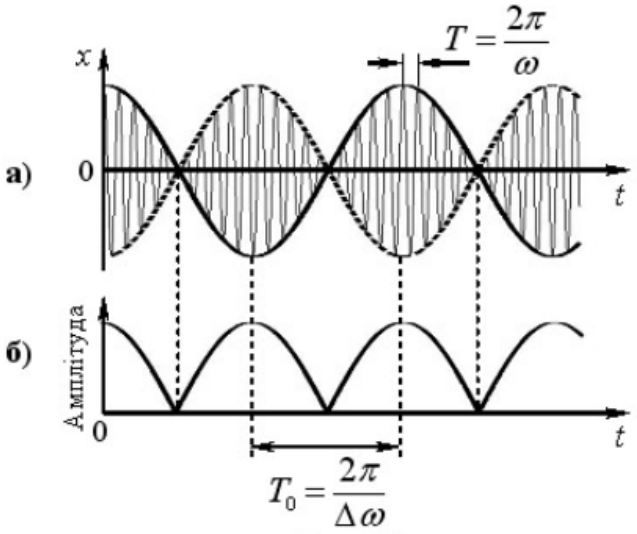


Рис. 1.4

Розглянуті коливання називають ***биттями.*** При накладанні акустичних коливань з близькими частотами ω_1 і ω_2 биття створюють відчуття (не завжди приємне) звукових пульсацій, які повторюються з частотою $|\omega_1 - \omega_2|$. Биття за відповідних умов виникають і при накладанні коливань не механічної природи, зокрема — напруг і струмів у електричних або електронних пристроях.

3.2. *Додавання взаємно перпендикулярних коливань*

Якщо точка одночасно здійснює гармонічні коливання у двох взаємно перпендикулярних напрямках, то на загал її рух не є коливальним і відбувається по складній траєкторії. Зокрема, траєкторія може виявитися незамкненою і ніколи не проходити через одну точку двічі. Проте, коли відношення частот коливань виражається відношенням цілих чисел, тіло рухається по замкненій траєкторії, форма котрої визначається співвідношенням частот, амплітуд і початкових фаз складових коливань.

Частоти однакові. Найпростішим є випадок додавання взаємно перпендикулярних коливань однієї частоти. Якщо коливання точки відбуваються вздовж координатних осей OX і OY, то їх можна описати рівняннями:

$$x = a\cos\omega t,\ y = b\cos(\omega t + \alpha), \qquad (1.15)$$

де початкова фаза коливань координати x вважається нульовою. Ото ж α по суті є різницею фаз цих двох коливань. Аби одержати рівняння траєкторії точки, тобто залежність $y(x)$, з рівнянь (1.15) треба виключити час. Для цього за допомогою відомих формул тригонометрії перетворимо друге рівняння (1.15) наступним чином:

$$\frac{y}{b} = (\cos\omega t\cos\alpha - \sin\omega t\sin\alpha) \Rightarrow \frac{y}{b}\left(\cos\omega t\cos\alpha - \sqrt{1-\cos^2\omega t}\cdot\sin\alpha\right).$$

Підставивши сюди вираз $\cos\omega t = x/a$, який випливає з першого рівняння (1.15), отримаємо:

$$\frac{y}{b} = \frac{x}{a}\cos\alpha - \sin\alpha\cdot\sqrt{1-\left(\frac{x}{a}\right)^2}.$$

Після піднесення до квадрату і нескладних перетворень одержимо шукане рівняння траєкторії:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2xy}{ab}\cos\alpha = \sin^2\alpha. \qquad (1.16)$$

З математики відомо, що рівняння (1.16) є рівнянням еліпса. Отже, на загал точка рухається по еліптичній траєкторії, параметри та просторова орієнтація котрої залежать від амплітуд a, b і різниці фаз α складових коливань. Розглянемо деякі окремі випадки.

1) $\boldsymbol{\alpha = 0}$. З (1.16) випливає, що

$$y = \frac{b}{a}x,$$

тобто еліпс вироджується у відрізок прямої. Отже, точка здійснює гармонічні коливання, напрямок яких пролягає в першій і третій чвертях і визначається відношенням амплітуд складових коливань.

2) $\boldsymbol{\alpha = \pi}$. У цьому випадку ситуація подібна — з (1.16) маємо:

$$y = -\frac{b}{a}x.$$

Тож у такому випадку точка теж здійснює гармонічні коливання, але вздовж лінії, що лежить у другій та четвертій чвертях.

Отримані результати дозволяють розглядати гармонічні коливання вздовж певного напрямку як суму (суперпозицію) двох гармонічних коливань тієї самої частоти, котрі відбуваються у взаємно перпендикулярних напрямках і мають різницю фаз $\boldsymbol{\alpha}$, або $\boldsymbol{\pi}$. Такий підхід плідно використовується, зокрема, в оптиці.

3) $\boldsymbol{\alpha = \pi/2}$. При такій різниці фаз рівняння (1.16) набуває вигляду:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Це означає, що при не однакових амплітудах складових коливань точка рухається по еліпсу, осі котрого орієнтовані вздовж осей координат. При $a = b$ еліпс вироджується в коло з центром у початку координат. Знак різниці фаз $\boldsymbol{\alpha}$ визначає один з двох можливих напрямів руху точки по траєкторії.

Фігури Ліссажу. Якщо точка здійснює взаємно перпендикулярні коливання з різними частотами, що співвідносяться, як цілі числа, траєкторіями руху є складні замкнені криві, що називаються *фігурами Ліссажу*.

У простих випадках рівняння фігур Ліссажу можна отримати за формулами тригонометрії. Наприклад, нехай складові коливання описуються рівняннями $x = A\sin\omega t$ та $y = A\sin 2\omega t$. Тоді

$$y = 2A\sin\omega t \cdot \cos\omega t = 2A\sin\omega t\sqrt{1 - \sin^2\omega t},$$

де $\sin\omega t = x/A$. Зробивши таку заміну, дістанемо рівняння траєкторії точки:

$$y = 2x\sqrt{1 - (x/A)^2}.$$

Її вигляд показує рис. 1.5.

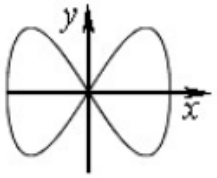


Рис. 1.5

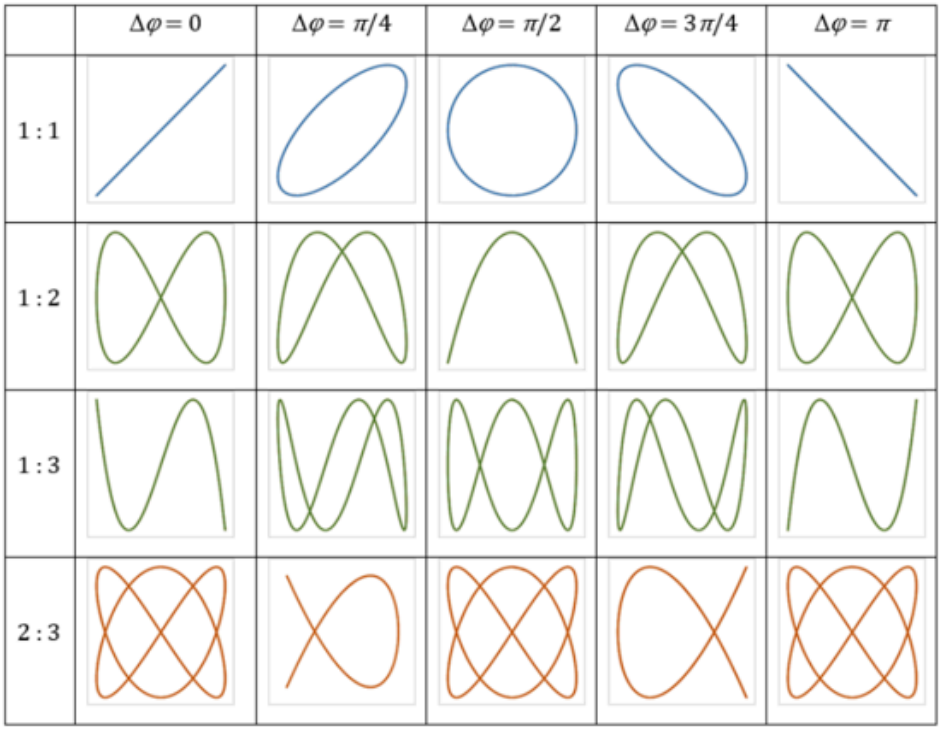


Рис. 1.6

Фігури Ліссажу для деяких інших співвідношень частот і зсувів фаз $\Delta\varphi$ коливань представлені на рис. 1.6. Зазначимо, що фігури Ліссажу при додаванні електричних коливань легко спостерігати на екрані осцилографа. Це використовують на практиці для вимірювання та налаштовування частот електричних сигналів.

4. Диференціальне рівняння гармонічних коливань

Незалежно від умов і фізичної природи величини $\xi(t)$ рівняння гармонічних коливань (1.4) є розв'язком єдиного диференціального рівняння. Його можна установити, знайшовши першу та другу похідні функції $\xi(t)$:

$$\frac{\mathrm{d}\xi}{\mathrm{d}t} = -\omega A \sin(\omega t + \varphi_0) \Rightarrow \frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} = -\omega^2 A \cos(\omega t + \varphi_0).$$

Як видно, вираз другої похідної містить вихідну функцію (1.4), отже,

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} = -\omega^2 \xi. \qquad (1.17)$$

Отримане співвідношення є критерієм (достатньою умовою) гармонічних коливань:

> ***якщо в будь-який момент часу друга похідна даної величини по часу виявляється прямо пропорційною самій величині і має протилежний знак, то ця величина здійснює гармонічні коливання, а коефіцієнт пропорційності дорівнює квадратові циклічної частоти коливань.***

Мовою математики вираз (1.17) є звичайним однорідним диференціальним рівнянням другого порядку із постійними коефіцієнтами, яке прийнято записувати у вигляді:

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} + \omega^2 \xi = 0. \qquad (1.18)$$

У математиці доводиться, що розв'язками такого рівняння можуть бути тільки гармонічні функції. Тому рівняння (1.18) є ***диференціальним рівнянням гармонічних коливань.***

Контрольні запитання

1. Які фізичні процеси називаються коливаннями? За якими ознаками класифікують коливання?

2. Назвіть умови існування вільних і вимушених коливань. Чи може однорідний диск, який закріплений на осі, що проходить через його центр, здійснювати вільні коливання? Вимушені коливання?

3. Які коливання називаються гармонічними та якими величинами вони характеризуються?

4. Що визначає амплітуда коливань? Чи може вона бути від'ємною?

5. Які величини визначають повторюваність коливального процесу в часі? Який вони мають зміст і як пов'язані між собою?

6. Чи можна довільні періодичні коливання характеризувати періодом T? Лінійною частотою ν? Циклічною частотою ω?

7. Чому гармонічні коливання можна зображувати векторами? Як у цьому способі відображається амплітуда коливань? Початкова та поточна фаза?

8. Чому гармонічні коливання можна зображувати комплексними експонентами?

9. Що таке комплексна амплітуда коливань? Яку інформацію вона містить?

10. Коли використання векторного або комплексного способів зображення гармонічних коливань є продуктивним?

11. Що таке биття? За якої умови вони спостерігаються?

12. Що таке фігури Лісажу? За якої умови вони спостерігаються?

13. Сформулюйте критерій гармонічних коливань і запишіть його математичний вираз.

2. Механічні гармонічні коливання

Найбільш наочними й зручними для вивчення є механічні коливання – рух тіла (матеріальної точки), при якому його положення у просторі та інші характеристики руху з плином часу періодично повторюються.

Характер руху всякого тіла визначається властивостями діючих на нього сил. Зокрема, якщо прикладена "*спонукальна*" сила періодично змінює свій напрям на протилежний і здатна подолати сили тертя та опору, то тіло здійснює вимушені коливання. Вільні механічні коливання тіла спостерігаються, коли рівнодійна прикладених сил весь час намагається повернути його в положення рівноваги. Таку силу інколи називають "*повертаючою силою*". Але слід пам'ятати, що цей термін є умовним. У природі не існує ніякої спеціальної повертаючої сили — вона створюється сумісною дією якихось інших сил. До прикладу, при вільних коливаннях кульки на нитці повертаюча сила є рівнодійною сил тяжіння та натягу нитки.

Далі розглянуто:

2.1. Координата, швидкість, прискорення

2.2. Умова існування механічних гармонічних коливань

2.3. Маятники

2.4. Енергія механічних гармонічних коливань

Контрольні запитання

1.1. *Координата, швидкість, прискорення*

Розглянемо характеристики руху тіла (матеріальної точки), що здійснює гармонічні коливання вздовж осі OX.

Рівняння координати. Згідно із виразами (1.4), координату, тіла при гармонічних коливаннях визначимо рівнянням:

$$x(t) = x_m \cos(\omega t + \varphi_0), \qquad (2.1)$$

де $\boldsymbol{x_m}$ ***– амплітуда коливань*** (максимальне відхилення тіла від положення рівноваги).

Рівняння швидкості. Швидкість і напрям руху тіла визначається проєкцією вектора швидкості на вісь OX, тобто похідною координати по часу: $v = \mathrm{d}x/\mathrm{d}t$. Отже, згідно з (2.1), рівняння швидкості тіла при гармонічних коливаннях має вигляд:

$$v(t) = -v_m\omega \sin(\omega t + \varphi_0) \text{ або } v(t) = v_m \cos(\omega t + \varphi_0 + \frac{\pi}{2}), \qquad (2.2)$$

де v_m ***– амплітуда (максимальне значення) швидкості*** пов'язана з амплітудою коливань співвідношенням

$$v_m = x_m\omega. \qquad (2.2a)$$

<u>Примітка</u>. В рівнянні (2.2) мається на увазі ***проєкція*** *вектора швидкості, але, позаяк рух відбувається вздовж однієї осі, індекс проєкції не вказують. Це стосується і всіх інших векторних величин, які використовуються далі для характеристики гармонічних коливань.*

З рівняння (2.2) видно, що коливання швидкості за фазою випереджають коливання координати на $\pi/2$ або на чверть періоду.

Рівняння прискорення отримаємо аналогічно з (2.1) або (2.2), врахувавши, що прискорення є першою похідною по часу від швидкості або другою похідною від координати:

$$a(t) = -A\omega^2 \cos(\omega t + \varphi_0) \text{ або } a(t) = a_m \cos(\omega t + \varphi_0 + \pi). \qquad (2.3)$$

Амплітуда прискорення

$$a_m = x_m\omega^2 = v_m\omega. \qquad (2.3а)$$

Із рівняння (2.3) випливає, що коливання прискорення випереджають (або відстають) коливання координати на π. Інакше говорячи, коливання прискорення відбуваються у протифазі до коливань координати.

На рис. 2.1 зображені графіки координати $x(t)$ швидкості $v(t)$ і прискорення $a(t)$ тіла при гармонічних коливаннях, які наочно показують фазові співвідношення між цими величинами.

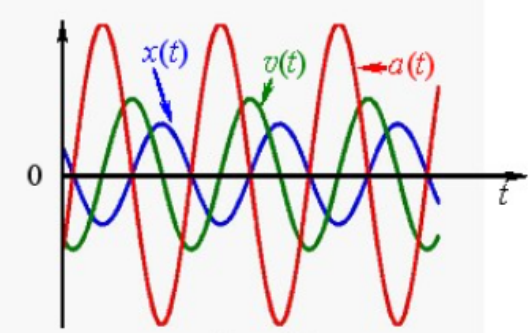


Рис. 2.1

2.2. *Умова існування механічних гармонічних коливань*

Із рівнянь (2.1) і (2.3) випливає, що при гармонічних коливаннях прискорення тіла в будь-який момент часу є прямо пропорційним до координати (зміщення з положення рівноваги):

$$a = -\omega^2 x. \qquad (2.4)$$

Тому за другим законом Ньютона проєкцію на вісь OX рівнодійної сили, що прикладена до тіла, можна записати, як

$$F = -kx, \qquad (2.5)$$

де

$$k = m\omega^2. \qquad (2.6)$$

Із (2.5) видно, що при $x > 0$ проєкція сили $F < 0$ і навпаки. Отже, будь-якої миті сила F напрямлена до положення рівноваги і є прямо пропорційною величині відхилення. Таку властивість має сила пружності. Але в загальному випадку рівнодійна F може створюватись і силами, які за природою не є пружними. Тому її називають ***квазіпружною силою.***

Сказане дозволяє сформулювати критерій (достатню умову) механічних гармонічних коливань:

якщо тіло перебуває під дією пружної або квазіпружної сили, то воно здійснює гармонічні коливання.

Частота ω і період T гармонічних коливань визначаються коефіцієнтом пропорційності k у виразі сили (2.5) і масою тіла m, згідно з формулами:

$$\omega = \sqrt{\frac{k}{m}}, \tag{2.7}$$

$$T = 2\pi\sqrt{\frac{m}{k}}. \tag{2.8}$$

Відмітимо також, що, відповідно до (2.5), сила, яка зумовлює гармонічні коливання (2.1), сама змінюється за гармонічним законом:

$$F = -F_m \cos(\omega t + \varphi_0),$$

де амплітуда сили $F_m = mx_m\omega^2$.

2.3. *Маятники*

Практичне значення розглянутої умови гармонічних коливань (2.5) і співвідношення (2.6) полягає в тому, що, визначивши рівнодійну прикладених до тіла сил, можна не лишень установити, чи здійснює воно гармонічні коливання, а й знайти їхню основну характеристику – частоту або період. Проілюструємо це на прикладі маятника – тіла, що підвішене на пружині чи нитці, або знаходиться на осі, навколо якої може здійснювати вільні коливання.

Пружинний маятник являє собою невеликий тягарець маси m, з'єднаний з невагомою пружиною жорсткості k із закріпленим іншим кінцем і може рухатися або горизонтально (рис.2.2а), або вертикально (рис.2.2*б*). Будемо вважати, що сили тертя та опору відсутні. Тоді, якщо координату x відраховувати від положення рівноваги тягарця, то прикладена до нього рівнодійна сила в обох випадках дорівнює тільки силі пружності $\vec{F}$, що виникає внаслідок зміщення маятника з положення рівноваги (сили тяжіння та реакція опори у випадку а) і сила тяжіння та статична сила пружності при x = 0 у випадку б) є компенсованими і на рисунку не показані):

$$F = -kx,$$

де x – координата маятника.

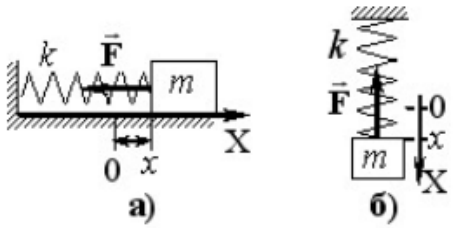


Рис. 2.2

Оскільки сила F задовольняє умову (2.5), маятник (тягарець на пружині), виведений з положення рівноваги і наданий самий собі, в обох випадках здійснює гармонічні коливання з циклічною частотою та періодом, які визначаються жорсткістю пружини k і масою тягарця m, відповідно до формул (2.7) і (2.8).

Математичний маятник. У строгому розумінні математичним маятником називається матеріальна точка, підвішена на невагомій нерозтяжній нитці. Реально – це кулька певної маси m підвішена на довгій тонкій сталевій дротині (струні) довжини l, масою котрої можна нехтувати (рис. 2.3). Силу тертя в точці кріплення та силу опору повітря також можна вважати не істотними.

Коли маятник перебуває в рівновазі, сили тяжіння та натягу нитки є компенсованими. Але, якщо кульку відвести й відпустити, з'явиться рівнодійна $\vec{F} = m\vec{g} + \vec{F}_{\text{н}}$, напрямлена до положення рівноваги (рис. 2.3). Відтак маятник буде здійснювати вільні коливання. Рух маятника можна розглядати як рух матеріальної точки маси m по колу радіуса l навколо горизонтальної осі Z, що проходить через точку кріплення нитки (т. О на рис. 2.3). Тому, відповідно до відомого з механіки рівнянням моментів відносно осі (див. МЕХАНІКА. Розділ 7, п. 7.3), можемо записати:

$$M_z = I\beta_z, \tag{2.9}$$

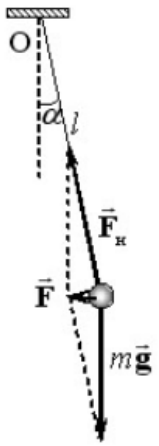


Рис. 2.3

де $M_z,\ I,\ \beta_z$ – сумарний момент сил, момент інерції кульки відносно осі О та кутове прискорення, відповідно.

Оскільки напрям сили натягу $\vec{F}_{\text{н}}$ проходить через вісь Z, момент відносно неї створює тільки сила тяжіння: $M_z = -mgl\sin\alpha$, де α - кут відхилення нитки маятника від вертикалі. Знак "–" у виразі стоїть тому, що кут α відраховують проти годинникової стрілки, а момент сили тяжіння діє за годинниковою стрілкою, тобто у від'ємному напрямку. Крім того $\beta_z = \mathrm{d}^2\alpha/\mathrm{d}t^2$. Відтак можемо записати:

$$-mgl\sin\alpha = I\frac{\mathrm{d}^2\alpha}{\mathrm{d}t^2} \quad \Rightarrow \quad \frac{\mathrm{d}^2\alpha}{\mathrm{d}t^2} + \frac{mgl}{I}\sin\alpha = 0,$$

Підставивши в (2.9) вираз моменту інерції матеріальної точки $I = ml^2$, отримаємо диференціальне рівняння руху математичного маятника:

$$\frac{\mathrm{d}^2\alpha}{\mathrm{d}t^2} + \frac{g}{l}\sin\alpha = 0. \tag{2.9а}$$

Диференціальне рівняння (2.9а) є доволі складним – воно не збігається з рівнянням гармонічних коливань (1.18) і не має розв'язків у елементарних функціях. Це означає, що коливання маятника не є гармонічними. Але при малих амплітудах максимальний кут відхилення нитки маятника від вертикалі малий ($\alpha_m << 1$), і рівнянні (2.9а) можна з достатньою точністю покласти $\sin\alpha = \alpha$. Тоді

$$\frac{\mathrm{d}^2\alpha}{\mathrm{d}t^2} + \omega^2\alpha = 0, \quad \omega^2 = \frac{g}{l}. \tag{2.9б}$$

Отримане спрощене рівняння (2.9б) не відрізняється від загального диференціального рівняння гармонічних коливань (1.18), отже, кут відхилення нитки маятника від вертикалі змінюється за законом (1.4):

$$\alpha(t) = \alpha_m\cos(\omega t + \varphi_0). \tag{2.10}$$

Оскільки при малих кутах дуга практично співпадає з хордою, можна вважати, що кулька маятника рухається вздовж горизонтальної осі, і її координата $x = l\alpha$ змінюється за законом:

$$x(t) = x_m\cos(\omega t + \varphi_0),$$

де амплітуда визначається довжиною та максимальним кутом відхилення маятника від вертикалі: $x_m = l\alpha_m$.

Таким чином, при малих амплітудах математичний маятник здійснює вільні гармонічні коливання з циклічною частотою

$$\omega = \sqrt{\frac{g}{l}} \tag{2.11}$$

і періодом

$$T = 2\pi\sqrt{\frac{l}{g}}. \qquad (2.11а)$$

Останній вираз називається ***формулою Гюйґенса***.

Фізичний маятник. Фізичним маятником називають протяжне тверде тіло, яке може вільно коливатися навколо закріпленої осі О, що не проходить через його центр мас С, (рис. 2.4). У положенні рівноваги центр мас маятника знаходиться на одній вертикалі з точкою підвісу (рис. 2.4а), тому сили тяжіння $m\vec{g}$ та реакції осі $\vec{R}$ є компенсованими. Але після відхилення тіла на деякий кут φ (рис. 2.4б) так само, як у випадку математичного маятника, виникає момент сили тяжіння $M = -mgl\sin\varphi$ (l – відстань між точкою підвісу та центром мас) відносно осі О, який намагається повернути маятник у положення рівноваги. Тому, якщо маятник надати самому собі, він почне коливатися згідно із рівнянням (2.9а), яке не є диференціальним рівнянням гармонічних коливань. Але, як і раніше, при малих кутах відхилення воно спрощується і набуває вигляду (2.9б):

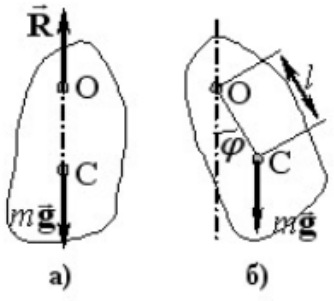


Рис. 2.4

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2} + \omega^2\varphi = 0, \omega^2 = \frac{mgl}{I}.$$

Це означає, що при малих кутах відхилення $\varphi_m \ll 1$, вільні коливання фізичного маятника є гармонічними і здійснюються за законом:

$$\varphi(t) = \varphi_m\cos(\omega t + \phi_0) \qquad (2.12)$$

із циклічною частотою

$$\omega = \sqrt{\frac{mdl}{I}}, \qquad (2.13)$$

чи періодом

$$T = 2\pi\sqrt{\frac{I}{mgl}}. \qquad (2.13а)$$

При розгляді коливань фізичного маятника часто використовують поняття ***зведеної довжини фізичного маятника*:**

$$l_{зв} = \frac{I}{ml}. \qquad (2.14)$$

У такому разі формули (2.13) і (2.13а) трансформуються в

$$\omega = \sqrt{g/l_{зв}} \qquad (2.15)$$

$$T = 2\pi\sqrt{l_{зв}/g}, \qquad (2.15а)$$

які збігаються з відповідними формулами для математичного маятника. Тому можна сказати, що

зведена довжина фізичного маятника дорівнює довжині математичного маятника з таким самим періодом вільних коливань.

2.4. *Енергія гармонічних коливань*

При зміні швидкості та координати тіла, що здійснює гармонічні коливання, відповідно змінюється і його кінетична та потенціальнв енергія.

Кінетична енергія гармонічних коливань (2.1) у довільний момент часу, з урахуванням виразів (2.2), (2.2а) і (2.6), визначається, як:

$$K = \frac{mv^2}{2} = \frac{1}{2}mx_m^2\omega^2 \sin^2(\omega t + \varphi_0) = \frac{1}{2}kx_m^2 \sin^2(\omega t + \varphi_0). \qquad (2.16)$$

Максимальна величина кінетичної енергії

$$K_m = \frac{1}{2}mx_m^2\omega^2 = \frac{1}{2}kx_m^2. \qquad (2.16а)$$

Гармонічні коливаннях тіла відбуваються під дією квазіпружної сили (2.5), котра, як і пружна, є консервативною. Тому, з урахуванням виразу (2.1), ***потенціальна енергія*** гармонічних коливань у довільний момент часу визначається виразом

$$U = \frac{kx^2}{2} = \frac{kx_m^2}{2}\cos^2(\omega t + \varphi_0). \qquad (2.17)$$

Максимальне значення потенціальної енергії

$$U_m = \frac{kx_m^2}{2}. \qquad (2.17а)$$

Із виразів (2.16а) та (2.17а) видно, що

при гармонічних коливаннях максимальні значення кінетичної та потенціальної енергій тіла збігаються:

$$K_m = U_m = W_0. \qquad (2.18)$$

Відтак для повної енергії гармонічних коливань у довільний момент часу, урахувавши (2.16), (2.17) і (2.18), отримуємо:

$$W = K + U = W_0 \sin^2(\omega t + \varphi_0) + W_0 \cos^2(\omega t + \varphi_0) = W_0. \qquad (2.19)$$

Таким чином,

при гармонічних коливаннях повна енергія коливань зберігається, тобто не залежить від часу.

Це означає, що при вільних гармонічних коливаннях механічна енергія не переходить в інші форми, зокрема, в тепло. Тому такі коливання можливі лише за відсутності сил тертя та опору, які насправді в механічних системах завжди є. Отже, реально

вільні коливання ніколи не бувають строго гармонічними.

Але це не стосується вимушених коливань, позаяк у цьому випадку втрати механічної енергії на тертя та опір можуть компенсуватися за рахунок роботи спонукальної сили.

Відмітимо також, що згідно з (2.16а), (2.17) і (2.18),

повна енергія гармонічних коливань є прямо пропорційною квадратові амплітуди.

Наостаок зауважимо, що вирази (2.16) і (2.17) можна записати, як

$$K = \frac{1}{2}K_m\left(1 - \cos(2\omega + 2\varphi_0)\right)$$

та

$$U = \frac{1}{2}U_m\left(1 + \cos(2\omega + 2\varphi_0)\right),$$

звідки видно, що

при гармонічних коливаннях кінетична та потенціальна енергії змінюються з подвоєною частотою по відношенню до частоти самих коливань.

На рис. 2.5 показані графіки кінетичної K, потенціальної U та повної W енергій тіла, що здійснює гармонічні коливання. З них видно, що кінетична та потенціальна енергії змінюються в протифазі. Це пояснює, чому максимальне значення кожної з них дорівнює повній енергії коливань.

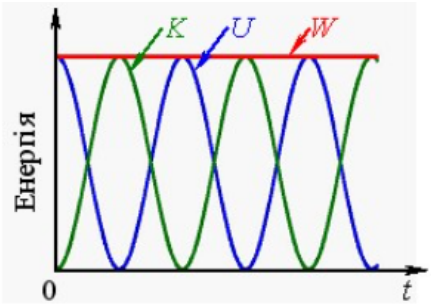


Рис. 2.5

Контрольні запитання

1. Який зв'язок існує між амплітудами швидкості та зміщення точки при гармонічних коливаннях? Між фазами цих величин?

2. Який зв'язок існує між амплітудами прискорення та швидкості точки при гармонічних коливаннях? Між їхніми фазами?

3. Який зв'язок існує між амплітудами прискорення та зміщення точки при гармонічних коливаннях? Між миттєвими значеннями цих величин? Між їхніми фазами?

4. За якої умови вільні коливання тіла є гармонічними? Чи може тіло здійснювати гармонічні коливання під дією не пружних сил?

5. Чи можна сказати, що вільні гармонічні коливання точки здійснюються під дією сили, яка змінюється за гармонічним законом у залежності: а) від часу; б) від координати точки?

6. Чи можна твердити, що вільні коливання маленького тягарця, підвішеного на невагомій нерозтяжній нитці, та коливання тягарця, підвішеного на ідеальній невагомій пружині, відбуваються за однаковим законом?

7. Що таке математичний маятник? За яких умов його вільні коливання є гармонічними?

8. Що таке фізичний маятник? За яких умов вільні коливання фізичного маятника є гармонічними?

9. Від чого і як залежить період вільних коливань фізичного маятника?

10. Що таке зведена довжина фізичного маятника та від чого вона залежить?

11. В якому співвідношенні (K_m : U_m) знаходяться максимальна кінетична та максимальна потенціальна енергія точки при гармонічних коливаннях ? Чому?

12. Який зв'язок існує між частотою гармонічних коливань точки та частотою зміни її кінетичної та потенціальної енергії? Чому?

13. Доведіть прямими розрахунками, що при гармонічних коливаннях потенціальна енергія математичного маятника залежить від його зміщення x із положення рівноваги як $U = kx^2/2$, де $k = mg/l$.

3. Загасаючі та вимушені механічні коливання

Незагасаючі вільні коливання є ідеалізацією. В реальній коливальній системі завжди присутні різноманітні сили тертя і опору, які призводять до втрат енергії коливань. Тому в дійсності вільні коливання є ***загасаючими*** – їхня амплітуда невпинно зменшується від початкового значення аж до нуля. Основні властивості таких коливань розглянуті нижче у наступних питаннях:

3.1. Рівняння вільних загасаючих коливань

3.2. Параметри загасання вільних коливань

3.3. Вимушені механічні коливання. Резонанс

Контрольні запитання

3.1. *Рівняння вільних загасаючих коливань*

Нехай на тіло, крім квазіпружної сили $F = -kx$, діє сила опору, яка зазвичай є пропорційною швидкості руху:

$$F_{\text{оп}} = -\alpha v = -\alpha \frac{\mathrm{d}x}{\mathrm{d}t},$$

де α – стала, яку називають коефіцієнтом опору. Знак "–" зумовлений тим, що сила опору напрямлена протилежно до вектора швидкості. В такому разі, згідно з другим законом Ньютона, рух тіла визначається рівнянням

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx - \alpha\frac{\mathrm{d}x}{\mathrm{d}t} \Rightarrow \frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \frac{\alpha}{m}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{k}{m}x = 0.$$

Уведемо співвідношення $(\alpha/m) = 2\beta$, в якому β називається ***коефіцієнтом загасання***. Врахуємо також, що згідно з (2.6) і (2.7) $k/m = \omega_0^2$, де ω_0 – частота вільних гармонічних коливань під дією квазіпружної сили. Її називають ***власною частотою*** коливальної системи. Після таких замін отримуємо ***диференціальне рівняння загасаючих коливань***:

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2\beta\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = 0. \qquad (3.1)$$

З математики відомо, що загальні розв'язки цього рівняння, отже й рух тіла, залежать від співвідношення між параметрами β і ω_0. Зокрема, за умови $\beta < \omega_0$ рух тіла має коливальний характер. Це зрозуміло, бо за такої умови домінує квазіпружна сила, що спричинює коливання. А от при $\beta \geq \omega_0$ сили опору є настільки великими, що рух тіла після виведенні з рівноваги є ***аперіодичним***: на шляху від крайнього до середнього положення тіло втрачає практично всю свою механічну енергію і зупиняється. Аналіз такого руху не становить інтересу для теми лекції, тому далі розглядається лише випадок $\beta < \omega_0$. У цьому разі з (3.1) випливає наступне загальне ***рівняння координати тіла при загасаючих коливаннях:***

$$xt = A_0 \mathrm{e}^{-\beta t}\cos(\omega t + \varphi), \qquad (3.2)$$

Графік функції (3.2) показаний на рис. 3.1 (штрихові лінії визначають межі, в яких знаходиться зміщення x тіла з положення рівноваги).

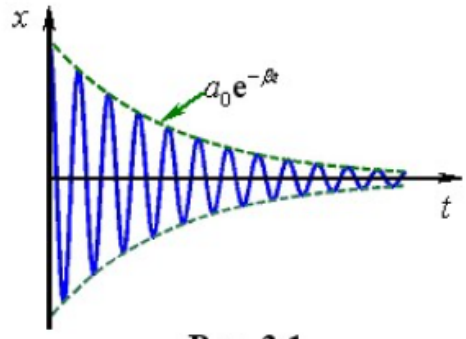


Рис. 3.1.

Циклічна частота загасаючих коливань дорівнює

$$\omega = \sqrt{\omega_0^2 - \beta^2}. \tag{3.3}$$

Відповідно, період

$$T = \frac{2\pi}{\sqrt{\omega_0^2 - \beta^2}}. \tag{3.3a}$$

Амплітуда (розмах) загасаючих коливань із часом спадає, тож крайні значення зміщення та інших характеристик руху з часом ***не повторюються.*** Тому загасаючі коливання не є гармонічними і навіть періодичними в строгому математичному сенсі. Відтак частота і період загасаючих коливань відображають не повне відтворення характеристик руху в часі, а лишень його циклічний характер. Одначе через крайні та нульові положення тіло проходить строго періодично з інтервалом у половину періоду гармонічної функції в рівнянні (3.2). Це дозволяє трактувати загасаючі коливання як гармонічні коливання із залежною від часу амплітудою $A(t)$:

$$x(t) = A(t)\cos(\omega t + \varphi). \tag{3.4}$$

де

$$A(t) = A_0 e^{-\beta t}, \tag{3.5}$$

Зауважимо також, що згідно з виразом (3.3) частота вільних загасаючих коливань ω відрізняється від власної частоти ω_0. Ця відміна тим більша, чим більший коефіцієнт загасання. Зокрема при $\beta \to \omega_0$ маємо $\omega \to 0$ і $T \to \infty$, що означає перехід до аперіодичного руху тіла. Таким чином,

вільні коливання можливі лише за умови $\beta < \omega_0$.

3.2. *Параметри загасання вільних коливань*

Для кількісної характеристики загасання коливань використовують декілька величин. Одна з них – ***коефіцієнт загасання β*** – фігурує безпосередньо в рівнянні (3.2). Що більша величина β, тим швидше спадає амплітуда коливань (рис. 3.2). Використовують також інші параметри загасання: час релаксації, логарифмічний декремент загасання та добротність.

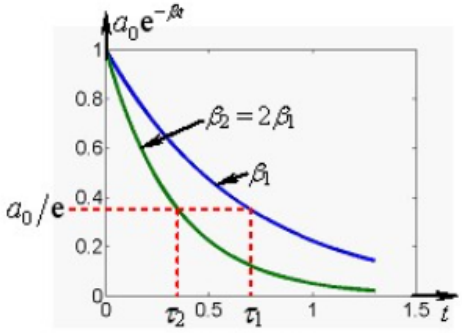


Рис. 3.2

Часом релаксації τ — то є проміжок часу, за який амплітуда коливань зменшується в $e \approx 2{,}72$ разів. Тож узявши до уваги вираз (3.4), отримаємо:

$$\frac{a_0 e^{-\beta t}}{a_0 e^{-\beta(t+\tau)}} = e \Rightarrow e^{\beta\tau} = e \Rightarrow \tau = \frac{1}{\beta}, \tag{3.6}$$

Таким чином, час релаксації є величиною оберненою до коефіцієнта загасання і показує наскільки повільно загасають коливання.

Логарифмічний декремент загасання λ – це логарифм відношення двох амплітуд, які відділені проміжком часу в один період:

$$\lambda = \ln \frac{A(t)}{A(t+T)} = \ln \frac{a_0 e^{-\beta t}}{a_0 e^{-\beta(t+T)}} \Rightarrow \lambda = \beta T = \frac{2\pi\beta}{\omega}. \tag{3.7}$$

З урахуванням (3.6) можна також записати:

$$\lambda = \frac{T}{\tau}. \tag{3.7а}$$

Інколи загасання характеризують кількістю коливань N_e, які відбуваються за час релаксації. Очевидно, що

$$N_e = \frac{\tau}{T} = \frac{1}{\lambda}. \tag{3.8}$$

Що більша величина N_e, тим довше в коливальній системі зберігаються вільні коливання, тобто тим якіснішою вона є. Але зазвичай якість коливальної системи характеризують не безпосередньо числом N_e, а ***добротністю Q*** – величиною, котра пов'язана з N_e або з декрементом загасання λ співвідношеннями:

$$Q = \pi N_e = \frac{\pi}{\lambda}. \tag{3.9}$$

Згідно з (3.7) і (3.7а), добротність можна виразити через частоту та коефіцієнт загасання або через час релаксації та період вільних коливань:

$$Q = \frac{\omega}{2\beta}, \tag{3.9а}$$

$$Q = \pi\frac{\tau}{T}. \tag{3.9б}$$

Добротність характеризує також втрати енергії коливань унаслідок загасання. Можна показати, що ***при слабкому загасанні*** ($\beta << \omega_0$)

$$Q = \frac{2\pi W}{|\Delta W|}, \qquad (3.10)$$

де W – енергія вільних коливань у даний момент часу, а $|\Delta W|$ – утрата енергії коливань за один період, починаючи від цього моменту.

3.3. Вимушені механічні коливання. Резонанс

Коливання різної природи широко використовуються в технічних пристроях і машинах, тому вони мають бути незагасаючими. Для цього до системи треба весь час підводити енергію, аби компенсувати завжди наявні втрати енергії коливань. Отже, незагасаючими можуть бути лише вимушені коливання, що створюються і підтримуються завдяки постійному періодичному зовнішньому впливу на систему.

Рівняння вимушених коливань. Вимушені механічні коливання відбуваються під дією періодичної зовнішньої ***спонукальної (змушуючої) сили***, властивості якої й визначають характер коливань. Найпростішими є вимушені коливання тіла, на яке, крім квазіпружної сили $F_{пр} = -kx$ та сили опору $F_{оп} = -\alpha v$, діє гармонічна спонукальна сила

$$F = F_0 \cos \omega t.$$

У такому разі для тіла маси m рівняння руху (другий закон Ньютона) має вигляд:

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx - \alpha\frac{\mathrm{d}x}{\mathrm{d}t} + F_0 \cos \omega t.$$

Увівши стандартні параметри $\alpha/m = 2\beta$, $k/m = \omega_0^2$ і позначення $f_0 = F_0/m$, це рівняння можна переписати у загально прийнятому вигляді ***диференціального рівняння вимушених коливань*** під дією гармонічної спонукальної сили:

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2\beta\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = f_0 \cos \omega t. \qquad (3.11)$$

У такому вигляді воно є чинним не лише для координати, а й для будь якої іншої фізичної величини.

Із математики відомо, що загальний розв'язок диференціального рівняння такого виду складається з двох доданків:

$$x(t) = a_0 e^{-\beta t} \cos(\omega' t + \varphi) + A \cos(\omega t - \varphi_0),$$

де величина $\omega' = \sqrt{\omega_0^2 - \beta^2}$ збігається з частотою вільних загасаючих коливань (3.3). Отже, можна говорити, що тіло бере участь у двох рухах: 1) загасаючих коливаннях із частотою і 2) незагасаючих коливаннях із частотою спонукальної сили ω. Але внесок від першого доданку з плином часу зменшується, тому амплітуда коливань поступово наближається до незмінної величини A, як показано на рис. 3.3.

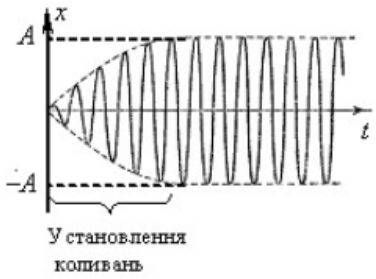


Рис. 3.3

Це відбувається за певний ***час установлення***, що є сумірним із часом релаксації τ вільних загасаючих коливань. Наприклад, за час $t = 5\tau$ амплітуда вимушених коливань досягає величини 99,5% від усталеної величини A. Відтак установлюються гармонічні стаціонарні вимушені коливання, що відбуваються за законом:

$$x(t) = A\cos(\omega t - \varphi_0). \tag{3.12}$$

Амплітуду A та початкову фазу φ_0 вимушених коливань можна визначити різними способами. Наприклад, підставивши функцію (3.12) в рівняння (3.11) і побудувавши відповідну векторну діаграму (див. Розділ I, п. 1.2). Відповідні розрахунки дають:

$$A = \frac{f_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2}}, \tag{3.13}$$

і

$$\mathrm{tg}\varphi_0 = \frac{2\beta\omega}{\omega_0^2 - \omega^2}. \tag{3.14}$$

Ці вирази указують на суттєві відмінності між вимушеними та вільними коливаннями:

1. У той час, як частота вільних коливань визначається параметрами коливальної системи, частота вимушених коливань від них ніяк не залежить і повністю визначається частотою спонукальної сили;

2. Амплітуда і початкова фаза вимушених коливань залежать як від властивостей спонукальної сили (амплітуди F_0 та частоти ω, так і від властивостей коливальної системи (власної частоти ω_0 та коефіцієнта загасання β). У той же час амплітуда та початкова фаза вільних коливань "ні від чого не залежать" і визначаються тільки способом виведення системі із стану рівноваги – величиною початкового відхилення, або початкової швидкості тіла;

3. Амплітуда вимушених коливань є функцією частоти: $A = A(\omega)$.

Резонанс. Проаналізуємо характер залежності амплітуди вимушених коливань від частоти, задля чого будемо подумки поступово збільшувати величину ω у виразі (3.13). Відтак при наближенні величини ω до ω_0 перший доданок під коренем у (3.13) буде швидко зменшуватися, наближаючись до нуля, а по тому – так само швидко збільшуватись. Відповідно, амплітуда A спочатку буде зростати, а потім спадати, проходячи при певному значенні частоти $\omega = \omega_p$ через максимум. Таке зростання амплітуди вимушених коливань при наближенні частоти спонукальної сили до певної величини ω_p називається ***резонансом***, а частота ω_p — ***резонансною частотою***. Резонансну частоту можна знайти, дослідивши підкорінну функцію в (3.13) на мінімум. Відповідний розрахунок дає

$$\omega_p = \sqrt{\omega_0^2 - 2\beta^2}. \qquad (3.15)$$

Підставивши цей вираз у (3.13), визначимо резонансну амплітуду:

$$A_p = \frac{f_0}{2\beta\sqrt{\omega_0^2 - \beta^2}}. \qquad (3.16)$$

Отримані вирази показують, що резонансна частота та резонансна амплітуда залежать від коефіцієнта загасання коливальної системи β. При його зменшенні резонансна частота ω_p наближається до власної частоти ω_0, а резонансна амплітуда A_p зростає, як показано на рис. 3.4.

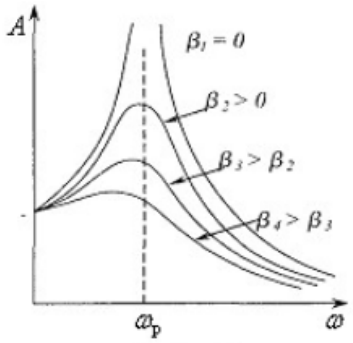


Рис. 3.4

При β = 0 формула (3.15) дає $\omega_p = \omega_0$, а (3.16) – $A_p \to \infty$. Але повністю позбутися загасання неможливо, можна лише зробити його слабким, так, що $\beta << \omega_0$. При цьому резонансна частота практично збігається із власною частотою системи $\omega_p = \omega_0$, але резонансна амплітуда коливань лишається скінченною.

Резонанс є важливим явищем, яке широко використовується в техніці. Приміром, на явищі резонансу ґрунтується дія язичкового частотоміра. Використовують це явище і в резонансних машинах, які призначені для випробування матеріалів на міцність. Широко використовують резонанс в акустиці, електро- та радіотехніці, тощо. Але явище резонансу може бути і шкідливим. До прикладу, відомо, що в масивних обертових вузлах машини при недостатньо точній збалансованості мас під час роботи виникають великі періодичні сили, що діють на підшипники, опору машини, корпус корабля, тощо. Частота таких сил, яка дорівнює частоті обертання валу, може збігтися з власною частотою коливань певної частини машини. При цьому виникнуть сильні резонансні коливання, що може призвести до поломки валу або руйнування фундаменту. Так само через недоліки в конструкції літака при певній швидкості польоту можуть виникнути руйнівні коливання крил. Тому при розробці подібних машин і пристроїв приділяють велику увагу запобіганню можливості виникнення резонансних коливань або, принаймні, їхньому послабленню до безпечного рівня.

Контрольні запитання

1. Чи можливі вільні гармонічні коливання в реальній механічній системі? Чому?

2. Якщо тіло маси m здійснює коливання в середовищі, де на нього діє сила опору

$F = -\alpha v$, то чому дорівнює коефіцієнт загасання коливань?

3. Чи є вільні загасаючі коливання гармонічними? Строго періодичними?

4. Чому частота вільних загасаючих коливань маятника менша, ніж його власна частота?

5. Назвіть ті величини, що визначають швидкість загасання коливань і ті, що характеризують їхню тривалість

6. Що показує час релаксації коливань?

7. Що називається логарифмічним декрементом загасання? Як він пов'язаний з часом релаксації?

8. Яку кількість вільних коливань здійснює точка за час релаксації при логарифмічному декременті загасання λ?

9. Що таке добротність коливальної системи? Як вона виражається через циклічну частоту та коефіцієнт загасання?

10. Який зв'язок існує між втратами енергії загасаючих коливань і добротністю системи?

11. Від чого залежать частота стаціонарних вимушених коливань і частота вільних коливань?

12. Від чого залежать амплітуда стаціонарних вимушених коливань і амплітуда вільних коливань?

13. У чому полягає явище резонансу? Від чого і як залежить резонансна частота вимушених гармонічних коливань точки?

4. Вільні електромагнітні коливання

Уже говорилося, що спостерігаються та використовуються на практиці коливальні процеси різної природи. Зокрема, велику роль у житті людини відіграють електромагнітні коливання – періодичні зміни струмів в електричних колах або характеристик електричного та магнітного полів. Це стосується як вимушених коливань, які створюються спеціальними пристроями, так і вільних коливань, які існують завдяки здатності електричного і магнітного полів до взаємних перетворень (ЕЛЕКТРИКА І МАГНЕТИЗМ. Частина V).

Вільні електромагнітні коливання розглядаються далі в наступних питаннях:

4.1. Коливальний контур

4.2. Вільні коливання в ідеальному контурі

4.3. Вільні загасаючі коливання в контурі

Контрольні запитання

4.1. Коливальний контур

Послідовний контур. Електромагнітні коливання спостерігаються в ***коливальних контурах*** – колах, що включають конденсатори, котушки індуктивності та резистори. Відповідно, електричними параметрами контура є опір R, індуктивність L і ємність C. Зазвичай роль резистора виконують з'єднувальні провідники. Коливальні контури можуть мати різну конфігурацію, але при вивченні загальних властивостей електромагнітних коливань розглядають найпростіші ***послідовні контури,*** в яких всі елементи сполучаються послідовно. Крім того, в теорії конденсатор і котушка контура розглядають як ідеальні. А саме, вважають, що електричне поле конденсатора та магнітне поле котушки повністю зосереджені всередині цих елементів, що виключає можливість випромінювання контуром електромагнітної енергії в простір. Також приймають, що конденсатор заповнено ідеальним діелектриком, а котушка має незмінну індуктивність $L = const,$ і її власний опір r включають в опір резистора R. Окрім цього, в загальному випадку контур включає зовнішнє джерело (генератор) змінної ЕРС $\mathcal{E}$. Принципова електрична схема такого контура показана на рис. 4.1.

Загальне рівняння послідовного контура. Коли в колі відбуваються електричні коливання, струм є нестаціонарним, тобто в різні моменти часу його величина не однакова. Але й у різних точках провідника величина струму в заданий момент часу теж є різною. Це пояснюється тим, що електричне поле, котре створюється джерелом і визначає величину струму в кожній точці провідника, поширюється вздовж нього із скінченною швидкістю. Через це виникає ефект запізнення – у віддалених від джерела точках задане значення струму встановлюється пізніше, ніж на вході кола. Проте через дуже велику швидкість поширення поля ($c = 3 \cdot 10^8$ м/с), ефект запізнення зазвичай є неістотним, і ***струм у контурі можна вважати квазістаціонарним*** (див. ЕЛЕКТРИКА І МАГНЕТИЗМ. Частина III, п. 3.1). За такої умови процеси в контурі є підпорядковані певному диференціальному рівнянню, яке можна встановити на основі загальних законів електричного струму.

Складемо диференціальне рівняння послідовного коливального контура (рис. 4.1).

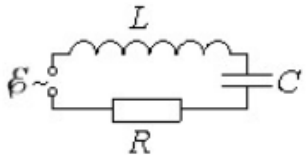


Рис. 4.1

За другим правилом Кірхґофа (ЕЛЕКТРИКА І МАГНЕТИЗМ, частина.III, п. 2.5) сума спадів напруги на всіх ділянках замкненого контура дорівнює сумі всіх ЕРС, які діють у ньому. В коливальному контурі таких ЕРС дві: ЕРС генератора та ЕРС самоіндукції котушки, отож

$$U_R + U_C = \mathcal{E} + \mathcal{E}_C. \tag{4.1}$$

Спад напруги на резисторі визначається законом Ома $U_R = IR$, спад напруги на конденсаторі дорівнює різниці потенціалів між обкладками і визначається його зарядом та ємністю $U_C = q/C$, а ЕРС самоіндукції за законом Фарадея (ЕЛЕКТРИКА І МАГНЕТИЗМ, частина. IV, п. 1.2) $\mathcal{E}_C = -L(\mathrm{d}I/\mathrm{d}t)$. Отже, маємо:

$$IR + \frac{q}{C} = \mathcal{E} - L\frac{\mathrm{d}I}{\mathrm{d}t}. \tag{4.2}$$

При протіканні струму в контурі кожна порція електрики, що проходить провідником за час $\mathrm{d}t$, потрапляє на пластину конденсатора і змінює його заряд на відповідну величину $\mathrm{d}q$. Тому в (4.2) величина I дорівнює першій, а $\mathrm{d}I/\mathrm{d}t$ – другій похідній заряду конденсатора по часу:

$$I = \frac{\mathrm{d}q}{\mathrm{d}t}, \quad \frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\mathrm{d}^2q}{\mathrm{d}t^2}. \tag{4.3}$$

Зробимо ці підстановки в (4.2):

$$L\frac{\mathrm{d}^2q}{\mathrm{d}t^2} + R\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{q}{C} = \mathcal{E} \tag{4.4}$$

і введемо позначення

$$\beta = \frac{R}{2L}, \tag{4.5}$$

$$\omega_0^2 = \frac{1}{LC}. \tag{4.6}$$

Відтак отримаємо ***загальне диференціальне рівнянням послідовного коливального контура***:

$$\frac{\mathrm{d}^2q}{\mathrm{d}t^2} + 2\beta\frac{\mathrm{d}q}{\mathrm{d}t} + \omega_0^2 q = \frac{\mathcal{E}}{L}, \tag{4.7}$$

яке відображає поведінку заряду конденсатора контура. Проінтегрувавши (розв'язавши) його, можна визначити, за яким законом $q(t)$ змінюються з часом заряд конденсатора, а відтак і струм в котушці та напруги на всіх ділянках контура за будь яких заданих умов.

4.2 Вільні коливання в ідеальному контурі

Коливання заряду та напруги на конденсаторі. Власна частота контура. Якщо в контурі R = 0, то він називається ідеальним. Розглянемо процеси в такому контурі за відсутності генератора (рис. 4.2). У цьому випадку в (4.7) $\beta = 0$ і $\mathcal{E} = 0$, отже

$$\frac{\mathrm{d}^2 q}{\mathrm{d}t^2} + \omega_0^2 q = 0. \tag{4.8}$$

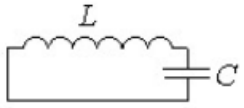


Рис. 4.2

Рівняння (4.8), згідно з (1.18), є рівнянням гармонічних коливань. Це означає, що, коли зарядити конденсатор певним зарядом q_m і надати контур самому собі, в ньому виникнуть ***гармонічні коливання заряду*** $q(t)$ і напруги $U_C(t)$ конденсатора, які можна описати загальними рівняннями (1.4):

$$q(t) = q_m \cos(\omega_0 t + \varphi_0), \tag{4.9}$$

$$U_C(t) = \cos(\omega_0 t + \varphi_0), \tag{4.9a}$$

Частота цих коливань, яку називають ***власною частотою контура***, згідно з (4.6), визначається формулою:

$$\omega_0 = \frac{1}{\sqrt{LC}}, \text{ або } \nu_0 = \frac{1}{2\pi\sqrt{LC}}, \tag{4.10}$$

а період – ***формулою Томсона*:**

$$T_0 = 2\pi\sqrt{LC}. \tag{4.10a}$$

Коливання струму та напруги на котушці індуктивності. Рівняння вільних коливань струму в котушці індуктивності контура знайдемо з виразів (4.3), скориставшись рівнянням (4.9):

$$I(t) = -I_m \sin(\omega_0 t + \varphi_0), \tag{4.12}$$

де ***амплітуда струму***

$$I_m = \omega_0 q_m = \omega_0 C U_{Cm}. \tag{4.12a}$$

Задля зручності перепишемо рівняння (4.12) у вигляді:

$$I(t) = I_m \cos\left(\omega_0 t + \varphi_0 + \frac{\pi}{2}\right). \tag{4.13}$$

Порівнюючи це рівняння з (4.11), бачимо, що

вільні коливання струму з в ідеальному контурі випереджають за фазою коливання напруги конденсатора на π/2 , або на чверть періоду.

Для встановлення рівняння коливань напруги на котушці врахуємо, що при вільних коливанннях в ідеальному контурі в рівнянні (4.2) $R = 0$ і $\mathcal{E} = 0$. Отже,

$$\frac{q}{C} + L\frac{\mathrm{d}I}{\mathrm{d}t} = 0.$$

З іншого боку, за другим правилом Кірхґофа за відсутності зовнішнього джерела,

$$U_C + U_L = 0.$$

Таким чином,

$$U_L = L\frac{\mathrm{d}I}{\mathrm{d}t}. \qquad (4.14)$$

Визначивши похідну $(\mathrm{d}I/\mathrm{d}t)$ з (4.12), отримаємо:

$$U_L(t) = -U_{Lm}\cos(\omega_0 t + \varphi_0), \text{ або}$$

$$U_L(t) = U_{Lm}\cos(\omega_0 t + \varphi_0 + \pi), \qquad (4.15)$$

де ***амплітуда напруги на котушці індуктивності***

$$U_{Lm} = \omega_0 L I_m. \qquad (4.15а)$$

Порівнюючи рівняння (4.15) і (4.13), бачимо, що

вільні коливання струму в ідеальному контурі відстають від коливань напруги на котушці індуктивності на π/2, або на чверть періоду.

Так само і рівнянь (4.15) і (4.13) випливає, що

в ідеальному контурі вільні коливання напруги на конденсаторі та на котушці індуктивності відрізняються за фазою на π, тож відбуваються в протифазі.

Фазові співвідношення між коливаннями струму та напруг в ідеальному контурі наочно показує рис.4.3.

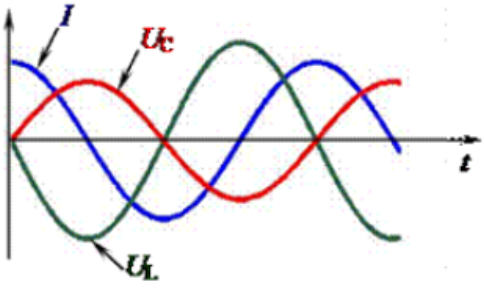


Рис. 4.3

Перетворення енергії в ідеальному коливальному контурі. Заряджений конденсатор і котушка індуктивності із струмом мають певний запас енергії, що зосереджена в їхніх полях – електричному та магнітному. Сума цих енергій складає енергію коливань у контурі. Позаяк в ідеальному контурі немає випромінювання та виділення у провідниках тепла при протіканні струму, то немає жодних каналів перетворення електромагнітної енергії на інші види. Тому

енергія вільних коливань в ідеальному контурі зберігається.

Ці якісні міркування можна обґрунтувати розрахунками. З електродинаміки відомо, що електрична енергія зарядженого конденсатора виражається як

$$W_C = \frac{CU^2}{2} = \frac{q^2}{2C}. \qquad (4.16)$$

Так само магнітна енергія котушки індуктивності визначається її індуктивністю та струмом:

$$W_L = \frac{LI^2}{2}. \qquad (4.17)$$

У коливальному контурі ці енергії, згідно з (4.11) і (4.12) визначаються рівняннями :

$$W_C = W_{Cm}\cos^2(\omega_0 t + \varphi_0), \text{ де } W_{Cm} = \frac{CU_m^2}{2} = \frac{q_m^2}{2C}, \qquad (4.18)$$

$$W_L = W_{Lm}\sin^2(\omega_0 t + \varphi_0), \text{ де } W_{Lm} = \frac{LI_m^2}{2}. \qquad (4.19)$$

Ці рівняння за допомогою відомих формул тригонометрії можна подати інакше:

$$W_C = \frac{1}{2}W_{Cm}(1 + \cos(2\omega_0 t + 2\varphi_0)),$$

$$W_L = \frac{1}{2}W_{Lm}(1 - \cos(2\omega_0 t + 2\varphi_0)).$$

Звідси випливає, що коливання енергій W_C і W_L відбуваються навколо їхніх середніх величин із частотою $\omega' = 2\omega_0$ і в протифазі. Крім того з виразів (4.10) і (4.12а) маємо:

$$I_m = \frac{1}{\sqrt{LC}} \cdot CU_{Cm} \quad \Rightarrow \quad LI_m^2 = CU_m^2, \qquad (4.20)$$

отже,

$$W_{Cm} = W_{Lm} = W_0.$$

Урахувавши це, для повної енергії коливань у контурі з (4.18) і (4.19) маємо

$$W = W_C + W_L = W_0\sin^2(\omega_0 t + \varphi_0) + W_0\cos^2(\omega_0 t + \varphi_0) = W_0. \qquad (4.21)$$

Таким чином, як і говорилося, повна енергія вільних коливань в ідеальному контурі зберігається. З плином часу відбуваються лише взаємні перетворення енергій електричного поля конденсатора та магнітного поля котушки індуктивності.

Енергетичні перетворення в ідеальному контурі ілюструє рис. 4.4:

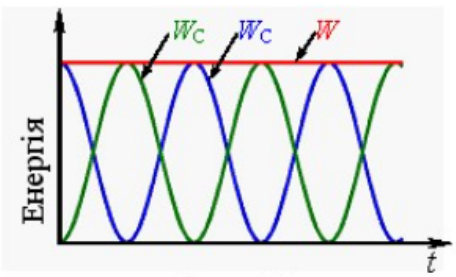


Рис. 4.4

4.3. Вільні загасаючі коливання в контурі

Якщо в контурі $R \neq 0$, то при протіканні струму в резисторі виділяється теплота, що утворюється за рахунок електромагнітної енергії контура. Тому за відсутності генератора (рис. 4.5) вільні коливання в контурі поступово загасають.

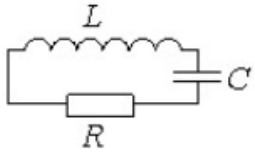


Рис. 4.5

Рівняння та характеристики загасаючих коливань у контурі. Поклавши в (4.7) $\mathcal{E} = 0$, отримаємо диференціальне рівняння вільних коливань у реальному контурі:

$$\frac{\mathrm{d}^2 q}{\mathrm{d}t^2} + 2\beta\frac{\mathrm{d}q}{\mathrm{d}t} + \omega_0^2 q = 0. \qquad (4.22)$$

Це рівняння не відрізняється від рівняння вільних загасаючих механічних коливань (3.1). Тому, якщо подати на конденсатор заряд q_m і надати контур самому собі, то при $\beta < \omega_0$ в ньому виникають вільні загасаючі коливання заряду, що відбуваються за законом:

$$q(t) = q_0 e^{-\beta t}\cos(\omega t + \varphi_0) \qquad (4.23)$$

або

$$q(t) = A(t)\cos(\omega t + \varphi_0), \qquad (4.24)$$

де величину $A(t)$ можна розглядати як залежну від часу амплітуду, що спадає за експоненціальним законом (рис. 4.6):

$$A(t) = q_0 e^{-\beta t}. \qquad (4.25)$$

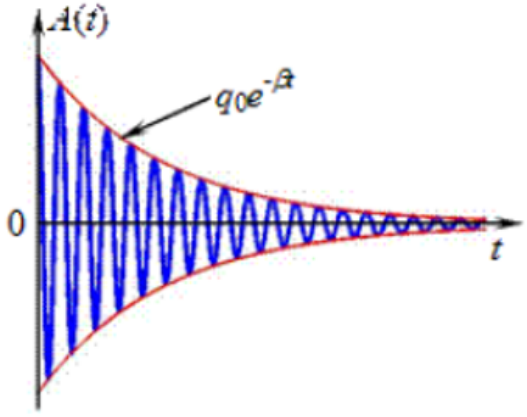


Рис. 4.6

Так само залежать від часу й амплітуди інших величин – напруг на конденсаторі й котушці індуктивності та струму в контурі. Тому можна записати загальний вираз ***амплітуди загасаючих коливань***:

$$A = A_0 e^{-\beta t}, \qquad (4.25а)$$

в якому константу A_0 формально можна трактувати як "початкову" амплітуду.

Циклічна частота ω і період T загасаючих коливань теж відрізняються від характеристик вільних гармонічних коливань ω_0, T_0 і визначаються формулами:

$$\omega = \sqrt{\omega_0^2 - \beta^2}, \tag{4.26}$$

і

$$T = \frac{2\pi}{\sqrt{\omega_0^2 - \beta^2}}. \tag{4.27}$$

Через параметри контура циклічна частота загасаючих коливань, згідно з (4.5) і (4.6), виражається формулою

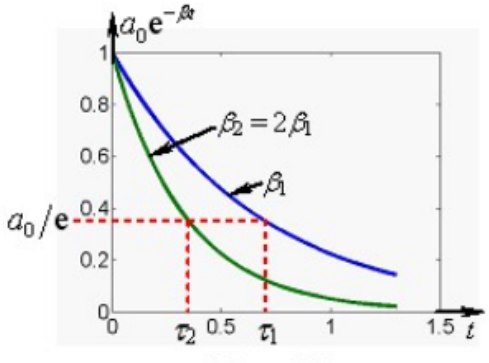
$$\omega = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}. \tag{4.28}$$

Для кількісної характеристики швидкості загасання або тривалості вільних коливань у контурі використовують ті самі величини, що й для механічних коливань: коефіцієнт загасання β, час релаксації τ, логарифмічний декремент загасання λ, добротність Q. Єдине, що варто зауважити, так це те, що частоти електричних коливань у контурах на загал є суттєво вищими, ніж частоти коливань механічних осциляторів, наприклад, маятників. Тому в режимі реального часу загасання вільних коливань в електричних контурах відбувається дуже швидко. Через це на практиці звичайно використовують не прямі міри загасання β і τ, а опосередковані величини λ і Q. Нагадаємо зміст указаних характеристик загасання та розглянемо, як вони визначаються через параметри коливального контура.

Коефіцієнт загасання β визначається формулою (4.5):

$$\beta = \frac{R}{2L}$$

і, згідно з (4.25), показує, як круто спадає амплітуда коливань. Це добре видно з рис. 4.7 і є цілком зрозумілим: що більша величина β, то більший опір контура R і більше електромагнітної енергії переходить у тепло протягом кожного періоду коливань.



Час релаксації τ визначається як

$$\tau = \frac{1}{\beta} = \frac{2L}{R}. \tag{4.29}$$

Згідно з (4.25), величина τ дорівнює проміжку часу, за який амплітуда коливань зменшується в $e \approx 2,72$ разів. Отже, час релаксації показує, як довго зберігаються в контурі вільні коливання. Це також видно з рис. 4.7. Цікаво, що час релаксації є прямо пропорційним індуктивності контура L. Це не випадково, адже вільні коливання в контурі виникають і підтримуються завдяки ЕРС самоіндукції котушки, яка пропорційна L.

Логарифмічний декремент загасання λ є зручною мірою зменшення амплітуди за один період, а отже мірою швидкості загасання коливань. За означенням він дорівнює

$$\lambda = \ln \frac{A(t)}{A(t+T)}.$$

Підставивши з (4.25) вирази амплітуд $A(t)$ і $A(t+T)$, дістанемо:

$$\lambda = \ln \frac{a_0 e^{-\beta t}}{a_0 e^{-\beta(t+T)}} = \beta T = \frac{T}{\tau}. \qquad (4.30)$$

Згідно з (4.5), (4.10) і (4.27), величина λ через параметри контура визначається формулою:

$$\lambda = \frac{\pi}{\sqrt{\frac{L}{CR^2} - \frac{1}{4}}}. \qquad (4.30а)$$

Добротність контура Q за означенням визначається, як

$$Q = \frac{\pi}{\lambda} \qquad (4.31)$$

і є кількісною характеристикою якості контура як коливальної системи, – адже чим більша добротність, тим менший декремент загасання і більший час збереження в контурі вільних коливань. Урахувавши вираз (4.30) і загальне співвідношення $T = 2\pi/\omega$, можна записати:

$$Q = \frac{\omega}{2\beta}. \qquad (4.32)$$

Через параметри контура добротність виражається, як

$$Q = \sqrt{\frac{L}{CR^2} - \frac{1}{4}}. \qquad (4.33)$$

При слабкому загасанні, коли $\beta < \omega_0$, частота загасаючих коливань близька до власної частоти контура ($\omega \approx \omega_0$). Тому для добротності з достатньою точністю можна записати:

$$Q = \frac{\omega_0}{2\beta}. \qquad (4.34)$$

Тоді на підставі (4.5) і (4.10) маємо зручну спрощену формулу для розрахунків:

$$Q = \frac{1}{R}\sqrt{\frac{L}{C}}. \qquad (4.35)$$

Розрахункова формула (4.30а) теж спрощується:

$$\lambda = \pi R \sqrt{\frac{C}{L}}. \qquad (4.36)$$

Слід зауважити, що, по-перше, умова слабкого загасання не є жорсткою, і, по-друге, на практиці коливальні контури завжди мають високу добротність, тобто дуже мале загасання. Тому в усіх реальних випадках спрощені формули (4.35) і (4.36) фактично є точними.

Параметри вільних коливань у контурі суттєво залежать від коефіцієнта загасання β. При цьому, якщо $\beta \to \omega_0$, то, відповідно до (4.26) і (4.27), $\omega \to 0$ і $T \to \infty$. Це означає, що заряд конденсатора та струм у котушці спадають до нуля, не здійснивши

жодного повного коливання. Такий режим називається *аперіодичною розрядкою* контура. Отже, при коефіцієнтах загасання $\beta \geq \omega_0$ вільні коливання в контурі неможливі. Тому величина $\beta_k = \omega_0$ називається ***критичним загасанням.*** Критичному загасанню відповідає ***критичний опір контура*** R_k, який визначити з виразів (4.5) і (4.10):

$$\beta_k = \omega_0 \quad \Rightarrow \quad \frac{R_k}{2L} = \frac{1}{\sqrt{LC}} \quad \Rightarrow \quad R_k = 2\sqrt{\frac{L}{C}}. \qquad (4.37)$$

Вільні коливання в контурі з опором R можливі лише за умови $R < R_k$, тому

критичний опір контура — то є найменший опір, при якому в контурі вже неможливі вільні коливання.

При $R \geq R_k$ всі величини, що характеризують коливання, втрачають зміст. Але їх все одно буває зручно виражати через критичний опір. Наприклад, згідно з (4.35) і (4.37), добротність контура можна подати, як

$$Q = \frac{R_k}{2R}. \qquad (4.38)$$

Загасаючі коливання напруги на конденсаторі та струму. Оскільки напруга на конденсаторі $U_C = q/C$, то поклавши в (4.23) задля зручності $\varphi_0 = 0$, отримаємо наступне рівняння загасаючих коливань напруги на конденсаторі контура:

$$U_C(t) = U_0 e^{-\beta t}\cos\omega t, \text{ де } U_0 = \frac{q_0}{C}. \qquad (4.39)$$

Для отримання рівняння коливань струму визначимо похідну $\mathrm{d}q/\mathrm{d}t$ з рівняння (4.23) за умови $\varphi_0 = 0$:

$$I(t) = \frac{\mathrm{dq}}{\mathrm{d}t} = -q_0 e^{-\beta t}(\beta\cos\omega t + \omega\sin\omega t).$$

Це рівняння можна привести до зручнішого вигляду, помноживши і поділивши його на власну частоту контура ω_0. Відтак

$$I(t) = -\omega_0 q_0 e^{-\beta t}\left(\frac{\beta}{\omega_0}\cos\omega t + \frac{\omega}{\omega_0}\sin\omega t\right). \qquad (4.40)$$

Згідно з (4.26), $\beta^2 + \omega^2 = \omega_0^2$, крім того, $\beta < \omega_0$ і $\omega < \omega_0$. Це дозволяє трактувати числові множники в дужках як гармонічні функції деякого кута, що визначається, скажімо, такими співвідношеннями:

$$\frac{\beta}{\omega_0} = \cos\psi; \quad \frac{\omega}{\omega_0} = \sin\psi, \quad \text{або} \quad \frac{\omega}{\beta} = \mathrm{tg}\psi. \qquad (4.41)$$

Тоді, увівши позначення $\omega_0 q_0 = I_0$ і подавши в (4.40) величину в дужках, як косинус різниці кутів, отримаємо рівняння загасаючих коливань струму в контурі:

$$I(t) = -I_0 e^{-\beta t}\cos(\omega t - \psi),$$

або

$$I(t) = I_0 e^{-\beta t}\cos(\omega t + \varphi), \qquad (4.42)$$

де $\varphi = \pi - \psi$, і, згідно з (4.41),

$$\mathrm{tg}\varphi = \frac{\omega}{\beta}. \qquad (4.43)$$

Таким чином, в контурі із загасанням, на відміну від ідеального, зсув фаз між вільними коливаннями напруги на конденсаторі та струму в котушці $\varphi \neq \pi/2$.

Величину φ за допомогою виразу (4.32) можна визначити й через добротність контура:

$$\mathrm{tg}\varphi = 2Q. \qquad (4.43а)$$

З цього співвідношення випливає, що φ із збільшенням добротності значно зростає і вже при Q = 25 досягає величини 89°. Отже, вказана відміна фазових співвідношень для реальних радіотехнічних контурів є неістотною. Але вона відіграє велику роль у звичайних (не коливальних), колах змінного струму. Детальніше про це говориться в розділі "Змінний струм", п. 6.

Енергія загасаючих коливань. Зрозуміло, що енергія вільних загасаючих коливань у контурі не зберігається, вона монотонно спадає з часом аж до нуля. При цьому через складні фазові співвідношення між напругою на конденсаторі і струмом в котушці індуктивності точний вираз повної енергії коливань у контурі виявляється громіздким і незручним для аналізу. Але, як щойно говорилося, при слабкому загасанні можна вважати, що зсув фаз між напругою (4.39) і струмом (4.42) складає $\varphi = \pi/2$, як в ідеальному контурі. Тоді для струму можна записати

$$I(t) = -I_0 e^{-\beta t} \sin \omega t. \qquad (4.44)$$

Відтак, провівши ті самі викладки, що й для ідеального контура (п. 4.2), отримаємо наступний наближений вираз повної енергії загасаючих коливань у контурі:

$$W(t) = W_0 e^{-2\beta t}, \quad \text{де} \quad W_0 = \frac{CU_0^2}{2} \approx \frac{LI_0^2}{2}. \qquad (4.45)$$

Отже, при слабкому загасанні

енергія вільних коливань у контурі спадає за експоненціальним законом, як і напруга та струм, але вдвічі швидше.

Утрати енергії вільних коливань визначаються загасанням β, від якого залежить і добротність контура Q. Тому можна встановити прямий зв'язок між утратами енергії коливань і добротністю. Для цього визначимо відносну втрату енергії коливань за час одного періоду:

$$\frac{\Delta W}{W} = \frac{W - W_1}{W} = 1 - \frac{W_1}{W}, \qquad (4.46)$$

де

$$W = W_0 e^{-2\beta t} \quad \text{і} \quad W_1 = W_0 e^{-2\beta(t+T)}.$$

Підставивши ці вирази в (4.46), та врахувавши (4.30) і (4.31), дістанемо:

$$\frac{\Delta W}{W} = 1 - e^{-2\lambda} = 1 - e^{2\pi/Q}. \qquad (4.47)$$

При слабкому загасанні декремент $\lambda << 1$, а добротність $Q >> 1$, тобто показник експоненти є суттєво меншим за одиницю. Тому можна скористатися відомим із математики наближеним виразом $e^x \approx 1+x$ при $x << 1$ і, зробивши заміну $x = -2\pi/Q$, отримати з (4.47):

$$\frac{\Delta W}{W} = \frac{2\pi}{Q}, \quad \text{або} \quad Q = 2\pi \frac{W}{\Delta W}. \qquad (4.48)$$

Цей вираз висвітлює ***енергетичний зміст*** добротності:

добротність контура характеризує відношення енергії вільних коливань в будь-який момент часу до втрати цієї енергії за час одного періоду коливань, починаючи від цього моменту.

Це твердження, рівно як і закон (4.45) загасання енергії $W(t)$, є чинними також і для механічних загасаючих коливань.

Контрольні запитання

1. За яких умов у колі можливі електричні коливання?

2. Чи можливі в колі без котушки індуктивності електричні коливання:

а) вільні; б) вимушені?

3. Чи можливі в колі без конденсатора електричні коливання: а) вільні; б) вимушені?

4. Що таке ідеальний коливальний контур?

5. У теорії конденсатор і котушка індуктивності коливального контура розглядаються як ідеальні. Що мається на увазі?

6. Що таке власна частота контура? Чи використовується цей параметр при розгляді коливань у не ідеальному контурі?

7. Співставте вільні електричні коливання в ідеальному контурі з ємністю C та індуктивністю L із вільними гармонічними коливаннями кульки маси m на пружині жорсткості k, і встановіть відповідність між величинами C, L та m, k.

8. Що можна сказаи про частоти вільних коливань у контурах, які мають однакові індуктивності та ємності?

9. Назвіть й охарактеризуйте величини, котрі визначають швидкість загасання вільних коливань у контурі.

10. У скільки разів зменшується амплітуда загасаючих коливань у контурі за час утричі більший, ніж час релаксації контра?

11. Назвіть й охарактеризуйте величини, котрі визначають тривалість існування вільних коливань у контурі.

12. Що таке критичний опір контура? Як він виражається через параметри контура?

13. Який зв'язок є між добротністю та: а) критичним опором контура;

б) енергією вільних коливань у контурі?

14. З якою частотою змінюється енергія конденсатора контура, якщо струм у контурі змінюється з частотою ω?

16. Установіть відповідність між електричною і магнітною енергією ідеального контура з одного боку та кінетичною і потенціальною енергією гармонічних коливань маятника з іншого.

5. Вимушені гармонічні коливання в контурі

Загасаючі коливання становлять інтерес для теорії. А на практиці щонайширше застосування мають незагасаючі електричні коливання – від промислового змінного струму і до найрізноманітніших радіоелектронних приладів і пристроїв. Для створення незагасаючих коливань у контурі необхідно весь час компенсувати втрати електричної енергії на опорі R. Це роблять, включаючи в контур зовнішнє джерело (генератор) змінної напруги, яке створює та підтримує незагасаючі коливання.

Далі розглянуті наступні питання:

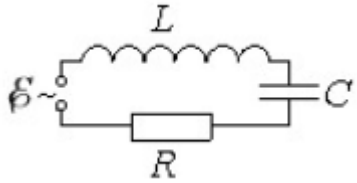
5.1. Амплітуда та фаза коливань заряду

5.2. Коливання напруг і струму. Резонанс

Контрольні запитання

5.1. *Амплітуда та фаза коливань заряду*

Схема контура для дослідження вимушених коливань була наведена в п. 4.1 і має вигляд:



Там же було встановлено й загальне диференціальне рівняння (4.7) послідовного контура:

$$\frac{\mathrm{d}^2q}{\mathrm{d}t^2} + 2\beta\frac{\mathrm{d}q}{\mathrm{d}t} + \omega_0^2 q = \frac{\mathcal{E}}{L},$$

Розглянемо контур, в який включено зовнішній генератор із ЕРС, що змінюється за гармонічним законом $E = E_0 \cos\omega t$. У такому разі наведене рівняння контура має вигляд:

$$\frac{\mathrm{d}^2q}{\mathrm{d}t^2} + 2\beta\frac{\mathrm{d}q}{\mathrm{d}t} + \omega_0^2 q = \frac{\mathcal{E}_0}{L}\cos\omega t. \qquad (5.1)$$

де $\beta = R/2L$ – загасання, і $\omega_0 = 1/\sqrt{LC}$ – власна частота контура.

Із теорії диференціальних рівнянь відомо, що загальні розв'язки цього рівняння можна записати, як

$$q(t) = q_0' e^{-\beta t}\cos(\omega' t + \varphi_0),$$

де $\omega' = \sqrt{\omega_0^2 - \beta^2}$, що збігається з частотою вільних коливань у контурі, а ω — частота генератора.

Така структура функцій $q(t)$ має просте фізичне обґрунтування. В момент увімкнення генератор виводить контур з рівноважного стану. Відтак у ньому крім вимушених коливань, які створюються й підтримуються генератором і визначаються другим доданком у виразі $q(t)$, збуджуються й вільні загасаючі коливання (перший доданок). Але останні через якийсь час практично повністю загасають, і в контурі встановлюються ***стаціонарні вимушені коливання*** заряду на конденсаторі, що визначаються рівнянням:

$$q(t) = q_0\cos(\omega t - \varphi_0). \qquad (5.2)$$

За своїми властивостями такі коливання суттєво відрізняються від вільних. Приміром, частота вимушених коливань зовсім не залежить від параметрів контура і задається генератором. Також, на відміну від вільних коливань, у яких амплітуда та

початкова фаза не залежать від параметрів контура і “задаються” умовами створення, амплітуда й фаза вимушених коливань визначаються не тільки генератором, а й параметрами контура.

Мовою математики амплітуда q_0 і початкова фаза φ_0 є сталими інтегрування рівняння (5.1), які визначаються шляхом підстановки розв’язку у вихідне рівняння. Тож знайдемо з (5.2) похідні $\mathrm{d}q/\mathrm{d}t$ та $\mathrm{d}^2q/\mathrm{d}t^2$ і підставимо їх у (5.1):

$$-\omega^2 q_0\cos(\omega t-\varphi_0)-2\beta\omega q_0\sin(\omega t-\varphi_0)+\omega_0^2 q_o\cos(\omega t-\varphi_0)=\frac{\mathscr{E}_0}{L}\cos\omega t.$$

Далі зробимо заміну

$$-\sin(\omega t-\varphi_0)=\cos(\omega t-\varphi_0+\frac{\pi}{2}),$$

і після елементарних викладок дістанемо:

$$(\omega_0^2-\omega^2)q_0\cos(\omega t-\varphi_0)+2\beta\omega q_0\cos(\omega t-\varphi_0+\frac{\pi}{2})=\frac{\mathscr{E}}{L}\cos\omega t. \qquad (5.3)$$

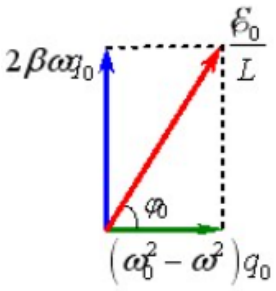


Рис. 5.1

Хоча в рівняння (5.3) входять дві невідомі величини – амплітуда заряду q_0 і початкова фаза φ_0, – з нього за допомогою векторного зображення гармонічних коливань (розділ 1, п. 1.2) можна визначити обидва параметри. При цьому ліва частина рівняння (5.3) зображується двома взаємно перпендикулярними векторами, а права – вектором, який є їхньою сумою, як показано на рис. 5.1. Вектори на цій діаграмі утворюють прямокутний трикутник, з якого легко отримати вирази для амплітуди та початкової фази заряду на конденсаторі контура:

$$\left((\omega_0^2-\omega^2)q_0\right)^2+(2\omega\beta q_0)^2=(\mathscr{E}/L)^2 \quad \Rightarrow$$

$$q_0=\frac{\mathscr{E}_0/L}{\sqrt{\left(\omega_0^2-\omega^2\right)^2+4\beta^2\omega^2}} \qquad (5.4)$$

і

$$\mathrm{tg}\varphi_0=\frac{2\omega\beta}{\omega_0^2-\omega^2}. \qquad (5.5)$$

5.2. Коливання напруг і струму. Резонанс

Вимушені коливання струму та напруг. Вимушені гармонічні коливання в контурі відбуваються за таким самим законом, як і вільні коливання в ідеальному контурі. Тому струм та напруги на конденсаторі і котушці індуктивності визначаються рівняннями й співвідношеннями, що є аналогічними до (4.11), (4.13), (4.15). Тож, увівши в них позначення

$$\frac{\pi}{2} - \varphi_0 = \varphi, \tag{5.6}$$

отримаємо:

$$I(t) = I_m \cos(\omega t - \varphi), \tag{5.7}$$

$$U_C(t) = U_{Cm} \cos\left(\omega t - \varphi - \frac{\pi}{2}\right), \tag{5.8}$$

$$U_L(t) = U_{Lm} \cos\left(\omega t - \varphi + \frac{\pi}{2}\right). \tag{5.9}$$

Початкова фаза коливань струму, згідно з (5.5) і (5.6), визначається, як

$$\mathrm{tg}\varphi = \frac{\omega^2 - \omega_0^2}{2\beta\omega}. \tag{5.10}$$

За фазою коливання напруги на конденсаторі (5.8) відстають, а на котушці індуктивності (5.9) випереджають коливання струму на $\pi/2$. Отже, фазові співвідношення при вимушених коливаннях в контурі такі самі, як і при вільних незагасаючих коливаннях.

Резонанс. Як і при механічних коливаннях (розділ 3, п. 3.3), амплітуди вимушених коливань у контурі залежать від частоти. Ці залежності називають ***амплітудними (або резонансними) характеристиками*** контура. Відповідно до (5.4) і (4.11а), (4.12а), та (4.15а), вони визначаються такими виразами:

$$U_{Cm} = \frac{\mathscr{E}_0 \omega_0^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2}}] \tag{5.11}$$

$$I_m = \frac{\mathscr{E}_0 \omega / L}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2}}, \tag{5.12}$$

$$U_{Lm} = \frac{\mathscr{E}_0 \omega^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2}}. \tag{5.13}$$

Із цих виразів випливає, що для напруг на конденсаторі й котушці та струму в контурі можливе явище резонансу – проходження амплітуди вимушених коливань через максимум при певній резонансній частоті $\omega = \omega_p$.

Резонансну частоту для кожної з указаних величин можна знайти за стандартним алгоритмом, визначивши та прирівнявши до нуля її похідну. Такі обчислення показують, що резонансна частота стуму збігається із власною частотою контура:

$$\omega_I = \omega_0 = \frac{1}{\sqrt{LC}}. \tag{5.14}$$

А от резонансні частоти напруг на конденсаторі ω_c і котушці індуктивності ω_L залежать від загасання β і визначаються виразами:

$$\omega_C = \sqrt{\omega_0^2 - 2\beta^2}, \tag{5.15}$$

$$\omega_L = \frac{\omega_0^2}{\sqrt{\omega_0^2 - 2\beta^2}}. \qquad (5.16)$$

Як видно, ці частоти пов'язані співвідношенням:

$$\omega_C \cdot \omega_L = \omega_0^2$$

Підстановка резонансних частот (5.14) – (5.16) у вирази амплітуд (5.11) – (5.13) дає:

$$I_p = \frac{\mathscr{E}_0}{R}, \qquad (5.17)$$

$$U_{Cp} = U_{Lp} = \frac{\mathscr{E}_0 \omega_0^2}{2\beta\sqrt{\omega_0^2 - \beta^2}}. \qquad (5.18)$$

При слабкому загасанні величиною β^2 під коренем можна знехтувати і, урахувавши (4.34), отримати:

$$U_{Cp} = U_{Lp} = Q\mathscr{E}_0 \quad \Rightarrow \quad Q = \frac{U_{Cp}}{\mathscr{E}_0} = \frac{U_{Lp}}{\mathscr{E}_0}. \qquad (5.19)$$

Ці вирази дають простий спосіб вимірювання добротності та показують ще одне "обличчя" цієї важливої характеристики контура:

добротність можна трактувати як коефіцієнт підсилення напруги на ємності або індуктивності контура при резонансі.

Головні особливості резонансу в коливальному контурі наочно відображають рис. 5.2 і 5.3.

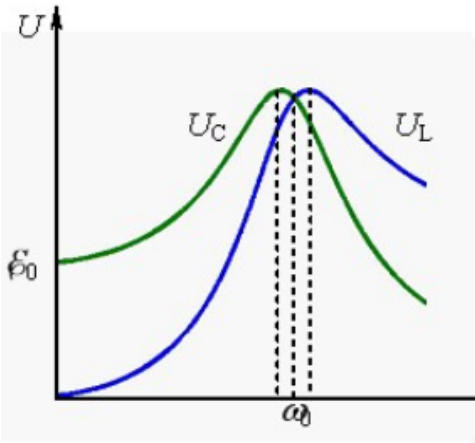


Рис. 5.2

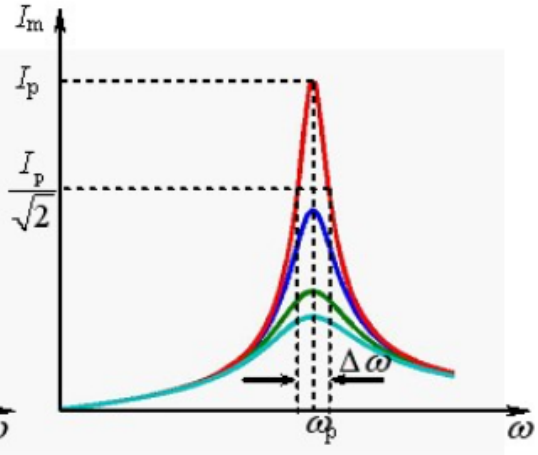


Рис. 5.3

На рис. 5.2 показані резонансні криві напруги на конденсаторі та на котушці індуктивності при заданій ЕРС генератора в контурі з малою добротністю й заданим загасанням, а на рис. 5.3 – резонансні криві струму в контурі при різних загасаннях і однаковій ЕРС генератора. Видно, що при зменшенні загасання і збільшенні добротності резонанс стає гострішим: криві $I_m(\omega)$ стають вищими та вужчими. Можна довести, що при слабкому загасанні добротність контура

$$Q = \frac{\omega_0}{\Delta\omega}, \qquad (5.20)$$

де ω_0 – резонансна частота струму (5.14), а $\Delta\omega = |\omega_1 - \omega_2|$ ширина резонансної кривої, визначена на рівні

$$I_m = \frac{I_p}{\sqrt{2}}, \qquad (5.21)$$

як показано на рис. 5.3.

Велика гострота резонансу в контурі з високою добротністю зумовлює широке практичне застосування коливальних контурів у царині радіотехніки та телекомунікацій, позаяк дозволяє здійснювати селекцію сигналів. Ставлячи на вході приймального пристрою коливальний контур і налаштовуючи його на відповідну резонансну частоту, можна з широкого спектру частот, які приходять з ефіру чи лінії передач, виділяти і далі опрацьовувати сигнали тільки з потрібною частотою. При цьому що вища добротність контура, тим краще він виділяє сигнал заданої частоти. Отже,

добротність виступає ще і як міра селективності контура.

Контрольні запитання

1. Що треба зробити, аби в реальному коливальному контурі створити гармонічні коливання?

2. Від чого залежить частота стаціонарних вимушених коливань у контурі?

3. Зобразіть приблизний вигляд графіка залежності від часу амплітуди вимушених коливань у контурі.

4. Чому дорівнює зсув фаз між коливаннями напруги генератора та заряду конденсатора контура? За якої умови коливання заряду за фазою: а) відстають від коливань напруги генератора; б) випереджають їх?

5. Відстають чи випереджають за фазою вимушені коливання напруги на конденсаторі контура коливання струму? На скільки?

6. Відстають чи випереджають за фазою вимушені коливання напруги на коушці контура коливання струму? На скільки?

7. Відстають чи випереджають за фазою вимушені коливання напруги на конденсаторі контура коливання напруги на котушці? На скільки?

8. За якої умови вимушені коливання струму в контурі за фазою: а) відстають від коливань напруги генератора; б) випереджають коливання напруги генератора; в) збігаються з коливаннями напруги генератора?

9. При якій частоті амплітуда вимушених коливань струму в контурі є максимальною?

10. При якій частоті амплітуда вимушених коливань напруги на конденсаторі контура є максимальною?

11. Чому дорівнює резонансна амплітуда напруги на конденсаторі контура при слабкому загасанні?

12. При якій частоті амплітуда вимушених коливань напруги на котушці контура є максимальною?

13. Чому дорівнює резонансна амплітуда напруги на котушці контура при слабкому загасанні?

6. Змінний струм

Одним із важливих для практики видів вимушених електричних коливань є змінний струм. У широкому значенні змінним струмом називають будь-які періодичні вимушені коливання струму та напруги в електричному колі. Але зазвичай під змінним струмом розуміють синусоїдальний струм, який змінюється за гармонічним законом. Це пояснюється тим, що саме за таким законом змінюється з часом ЕРС потужних індукційних генераторів на електростанціях, які дають більшу частину всієї споживаної людством електричної енергії. Далі розглядаються такі питання:

6.1. Закон Ома для змінного струму

6.2. Резонанс у колі змінного струму

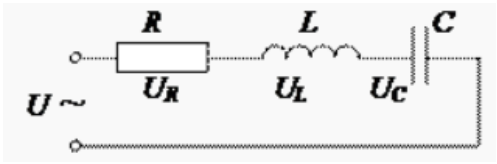
6.3. Потужність у колі змінного струму

Контрольні запитання

6.1. *Закон Ома для змінного струму*

Найпростішим повним за складом колом змінного струму є послідовний контур, який включає резистор R, конденсатор C і котушку індуктивності L (див. схему), на які подається гармонічна напруга

$$u = U_0 \cos\omega t. \qquad (6.1)$$



Але на відміну від коливального контура, в загальному випадку наявність у колі всіх полічених елементів не є обов'язковою. Зокрема це стосується котушки індуктивності. Тому в теорії змінного струму вона розглядається не як джерело ЕРС самоіндукції, а як пасивний елемент кола, що разом з іншими живиться від генератора змінної напруги.

Раніше (ЕЛЕКТРИКА І МАГНЕТИЗМ, частина ІІІ) було показано, що змінні струми практично завжди є квазістаціонарними. Тому, відповідно до другого правила Кірхґофа, для миттєвих значень напруг на ділянках послідовного кола можна записати:

$$u_R + u_C + u_L = U_0 \cos\omega t. \qquad (6.2)$$

При цьому струм у всіх ділянках кола в кожен момент часу однаковий, і виражається загальним рівнянням

$$i = I_0 \cos(\omega t - \varphi), \qquad (6.3)$$

У випадку стаціонарного (постійного) струму зв'язок між струмом та напругою, що подається на послідовне коло, визначається законом Ома. Чи це так і для змінного струму, не є очевидним і потребує спеціального розгляду. Спочатку проаналізуємо зв'язок між струмом і напругою для кожної ділянки кола окремо.

Резистор у колі змінного струму. Активний опір. Резистор по суті – то є просто провідник із певним опором R , для якого в кожну мить виконується закон Ома $u = iR$. Отже, рівняння коливань напруги на опорі R, згідно з (6.3), має вигляд:

$$u_R = I_0 R \cos(\omega t - \varphi). \tag{6.4}$$

Відмітимо, що струм і напруга на опорі R змінюються синфазно, і їхні амплітуди пов'язані співвідношенням:

$$U_{R0} = I_0 R \text{ або } I_0 = \frac{U_{R0}}{R}. \tag{6.5}$$

У колах змінного струму крім резисторів фігурують і котушки індуктивності та конденсатори, які теж чинять опір протіканню струму. Тому в теорії змінного струму опір R, що зумовлений розсіюванням носіїв струму і є причиною виділення у провіднику джоульового тепла, називають ***активним опором***.

Котушка індуктивності та конденсатор у колі змінного струму. Реактивні опори. При протіканні змінного струму по котушці індуктивності напруга на ній визначається як:

$$U_L = L\frac{\mathrm{d}I}{\mathrm{d}t}.$$

Тому, згідно з (6.3), напруга на котушці змінюється за законом:

$$u_L = -I_0 \omega L \sin(\omega t - \varphi) = I_0 \omega L \cos\left(\omega t - \varphi + \frac{\pi}{2}\right). \tag{6.6}$$

Отже,

> ***коливання напруги на котушці за фазою випереджають коливання струму* на $\pi/2$,** ***або чверть періоду,***

і їхні амплітуди пов'язані співвідношенням:

$$U_{L0} = I_0 \omega L \quad \text{або} \quad I_0 = \frac{U_{L0}}{X_L}, \tag{6.7}$$

де уведено позначення

$$X_L = \omega L. \tag{6.8}$$

Установимо тепер зв'язок між струмом $\boldsymbol{i}$ у колі та напругою $\boldsymbol{u_c}$ на конденсаторі. При протіканні струму в колі з конденсатором заряд, який проходить по провідниках, потрапляє на обкладку конденсатора. Тому кожної миті сила струму в колі дорівнює швидкості зміни заряду конденсатора: $i = \mathrm{d}q/\mathrm{d}t$, тож

$$dq = idt \quad \Rightarrow \quad q = \int idt.$$

Відповідно, напруга на конденсаторі $u_C = q/C$ визначається, як

$$\begin{gathered} u_C = \frac{1}{C}\int I_0 \cos(\omega t - \varphi)\mathrm{d}t = \frac{I_0}{\omega C}\sin(\omega t - \varphi) \quad \Rightarrow \\ u_c = U_{C0}\cos\left(\omega t - \varphi - \frac{\pi}{2}\right). \end{gathered} \tag{6.9}$$

Таким чином,

> ***коливання напруги на конденсаторі за фазою відстають від коливань струму* на $\pi/2$,** ***або чверть періоду.***

При цьому їхні амплітуди пов'язані співвідношенням:

$$U_{C0} = I_0 X_C \quad \text{або} \quad I_0 = \frac{U_{C0}}{X_C}, \tag{6.10}$$

де уведено позначення

$$X_C = \frac{1}{\omega C}. \tag{6.11}$$

Порівняння виразів (6.7), (6.10) із (6.5) показує, що котушка індуктивності та конденсатор обмежують амплітуду струму при заданій амплітуді напруги, аналогічно до того, як це робить резистор. Тому можна говорити, що індуктивність і ємність теж створюють опір для протікання змінного електричного струму. Відповідно, величина X_L, яка визначається формулою (6.8), називається ***реактивним індуктивним опором***, а величина X_C (формула (6.11)) — ***реактивним ємнісним опором***.

Отримані співвідношення (6.5), (6.7) і (6.10) показують, що на кожному опорі *амплітуди* струму та напруги пов'язані між собою законом Ома. Але для миттєвих значень він не виконується. Справді, з рівнянь (6.3), (6.6) і (6.9) випливає, що

$$\frac{U_L(t)}{i(t)} = -\omega L \mathrm{tg}(\omega t - \varphi) \neq X_L, \tag{6.12}$$

і

$$\frac{u_C(t)}{i(t)} = \frac{1}{\omega C} \mathrm{tg}(\omega t - \varphi) \neq X_C, \tag{6.12а}$$

тобто

миттєва величина струму не є прямо пропорційною напрузі,

як того вимагає закон Ома. Відповідно, реактивний опір, на відміну від активного, не можна трактувати як коефіцієнт пропорційності між напругою та струмом на котушці або на конденсаторі.

Властивості реактивних опорів є доволі специфічними. З одного боку реактивний опір, подібно до активного, залежить від параметрів відповідного пристрою – котушки індуктивності чи конденсатора. Але на відміну від активного,

реактивні опори (6.8), (6.11) залежать від частоти струму.

Тому, якщо на коло подати постійну напругу (ω = 0), то реактивні опори X_L = 0 і $X_C \to \infty$. Це абсолютно природньо: адже для постійного струму котушка являє собою просто шматок дроту, а конденсатор – розрив кола. Якщо котушки в колі взагалі немає, то знову маємо L = 0, і X_L = 0. Зрозуміло, що за відсутності конденсатора так само X_C = 0, але з (6.11) при цьому випливає досить дивний формальний висновок: $C \to \infty$. Цей "парадокс" має просте пояснення. Відсутність конденсатора означає відсутність у колі розриву провідності, отже, "викидаючи" з кола конденсатор, ми з'єднуємо провідники на місці його розташування. Тому формальний перехід до випадку "коло без конденсатора" можна здійснити, подумки наближаючи пластини конденсатора аж до дотику. При цьому відстань між пластинами конденсатора d = 0, і "все стає на місця": ємність $C = \infty$, і опір $X_C = 0$.

З ємнісним опором пов'язано ще одне питання. Опір X_C ми трактуємо як величину, що пов'язує амплітуди напруги та струму в конденсаторі, тоді як заряджені частинки-носії струму рухатися в діелектрику між пластинами конденсатора не можуть. І дійсно, стаціонарний струм може протікати тільки по замкнених провідниках. Але змінний струм можливий і при розриві провідності, як от у колі з конденсатором. При цьому із загальних рівнянь електродинаміки (ЕЛЕКТРИКА І

МАГНЕТИЗМ, V, п. 2) випливає, що лінії струму провідності, які розриваються на пластинах конденсатора, замикаються всередині лініями струму зміщення, котрий є мірою швидкості зміни електричного поля всередині конденсатора. Отже

струм конденсатора – то є струм зміщення,

котрий в кожен момент часу збігається із струмом провідності в провідниках, які підходять до обкладок.

Імпеданс. Закон Ома для змінного струму. Встановимо тепер, як струм у колі пов'язаний з напругою на вході (напругою генератора). Вище вже говорилося, що в послідовному колі за умови квазістаціонарності напруга на вході дорівнює сумі напруг на всіх ділянках кола, а струм скрізь однаковий, так само, як у випадку стаціонарного (постійного) струму. Тому може здатися, що величина змінного струму, як і у випадку постійного, дорівнює відношенню напруги на вході до суми опорів усіх ділянок кола. Але насправді це не так, оскільки напруги на різних ділянках кола мають різні фази (див. рівняння (6.1),(6.4), (6.6) і (6.9)).

Для встановлення зв'язку між змінним струмом та напругою побудуємо векторну діаграму кола так, як це робилося при розгляді вимушених коливань у контурі (п. 5, рис. 5.1). При цьому врахуємо, що фаза напруги на індуктивному опорі випереджає, а на ємнісному – відстає від фази струму на $\pi/2$. З отриманої векторної діаграми, котра показана на рис. 6.1, випливає, що

$$U_0^2 = U_R^2 \left(U_L^2 - U_C^2\right). \tag{6.13}$$

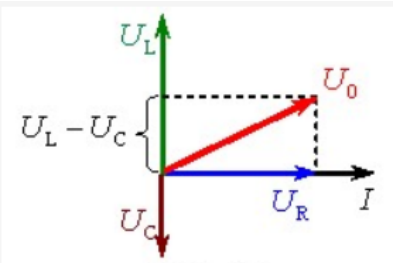


Рис. 6.1

Відтак, урахувавши (6.5), (6.9) і (6.10), одержимо

$$U_0 = I_0\sqrt{R^2 + (X_L - X_C)^2}. \tag{6.14}$$

Величина

$$X = X_L - X_C \tag{6.15}$$

називається ***реактивним опором***, а величина

$$Z = \sqrt{R^2 + X^2} \tag{6.16}$$

називається ***повним опором,*** або ***імпедансом*** послідовного кола змінного струму. Тож співвідношення (6.14) ***між амплітудами*** напруги та струму набуває вигляду:

$$U_0 = I_0 Z, \text{ або } \quad I_0 = \frac{U_0}{Z} \tag{6.17}$$

і є подібним до співвідношення між напругою та силою постійного струму. Тому воно називається ***законом Ома для змінного струму***. З урахуванням (6.8), (6.11) і (6.15), імпеданс послідовного кола та закон Ома в розгорнутому вигляді записуються, як

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}, \qquad (6.18)$$

$$I_0 = \frac{U_0}{\sqrt{R^2 + (\omega L - \frac{1}{\omega C})^2}}. \qquad (6.19)$$

Із векторної діаграми рис. 6.1 також випливає, що зсув фаз φ між струмом і напругою на вході кола визначається виразом:

$$\mathrm{tg}\varphi = \frac{U_L - U_C}{U_R} = \frac{\omega L - 1/\omega C}{R} = \frac{X}{R}. \qquad (6.20)$$

Стосовно закону Ома для змінного струму необхідно зробити наступні зауваження:

1. Поняття реактивних опорів й імпедансу та їхні вирази введені для гармонічних (синусоїдальних) струмів. Тому всі розглянуті співвідношення включно із законом Ома стосуються тільки таких струмів.

2. ***Закон Ома*** (6.17) виконується для амплітудних і ***не виконується для миттєвих значень*** струму та напруги. Це випливає з рівнянь (6.1), (6.3) і виразу (6.17):

$$\frac{u(t)}{i(t)} = Z\frac{\cos\omega t}{\cos(\omega t - \varphi)} \neq Z.$$

Така "невідповідність" пояснюється зсувом фаз φ між струмом і напругою, через що вони ***набувають своїх амплітудних значень у різні моменти часу.*** Тому кожної миті напруга і струм складають не однакові й не фіксовані частки від амплітудних значень. Через це за великим рахунком змінний струм не не задовольняє закон Ома, позаяк сила струму не є прямо пропорційною напрузі. Але для практики це нічого не вартує, бо, через швидку зміну, значущими є лишень усереднені характеристики змінного струму, що залежать тільки від амплітуд і тому підпорядковуються законові Ома.

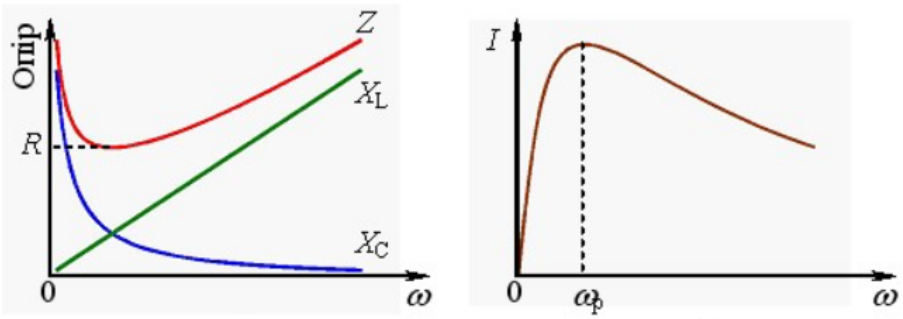
6.2. *Резонанс у колі змінного струму*

Реактивні опори (6.8), (6.11) та імпеданс (6.18) залежать від частоти, як показано на рис. 6.2а. При зростанні частоти індуктивний опір збільшується, а ємнісний зменшується. Тому, коли опори X_L і X_C зрівнюються, то реактивний опір $X_L - X_C$, = 0 й імпеданс, згідно з (6.18), має мінімальне можливе значення $Z = R$. Відповідно, амплітуда струму (6.19) набуває максимального можливого при заданій амплітуді напруги на вході значення

$$I_p = \frac{U_0}{R} \qquad (6.21)$$

і надалі спадає (рис. 6.2б). Тобто, в колі змінного струму спостерігається резонанс. Резонансна частота ω_p визначається умовою $X_L = X_C$, отже згідно з (6.8) і (6.11),

$$\omega_p L = \frac{1}{\omega_p C} \quad \Rightarrow \quad \omega_p = \frac{1}{\sqrt{LC}} = \omega_0. \qquad (6.22)$$



а) б)

Рис. 6.2

При резонансі напруга на активному опорі кола має найбільше можливе значення, рівне напрузі генератора, а сумарна напруга на реактивних опорах дрпівнює нулю. Але напруги окремо на котушці індуктивності $U_L = I_p X_L$ та на конденсаторі $U_C = I_p X_C$ не дорівнюють нулеві. Згідно з (6.8), (6.11), (6.21), і (6.22), вони визначаються, як

$$U_L = U_C = \frac{U_0}{R}\sqrt{\frac{L}{C}}, \qquad (6.23)$$

і можуть виявитися набагато вищими, ніж амплітуда напруги генератора. Це має враховуватися, бо при резонансі може виникнути загроза електричного пробою ізоляції провідників.

Як видно з (6.22), резонансна частота струму в колі збігається із власною частотою коливального контура (4.10), що має такі самі індуктивність і ємність. Всі інші розглянуті співвідношення для кола змінного струму теж збігаються з аналогічними співвідношеннями, що характеризують коливальний контур. У цьому не важко переконатися, зробивши у відповідних виразах для контура заміни:

$$\omega_0 = \frac{1}{\sqrt{LC}}, \quad \omega L = X_L \quad \text{і} \quad \frac{1}{\omega C} = X_C.$$

Указаний збіг є цілком природнім, адже коливальний контур принципово не відрізняється від повного кола змінного струму. Відміна полягає тільки у співвідношенні параметрів: активний опір R у реальних коливальних контурах є дуже малим, а в силових електричних колах змінного струму, напроти, великим.

6.3. *Потужність у колі змінного струму*

Можливості отримання корисної енергії в тому чи іншому електричному пристрої визначаються роботою та потужністю електричного струму в колах цього пристрою. Розглянемо, від чого залежить потужність на окремих ділянках та в усьому колі змінного струму.

Потужність на активному опорі. Діючі значення напруги та струму. Спочатку розглянемо потужність струму, що виділяється на резисторі. Миттєва потужність на будь-якій ділянці кола дорівнює добутку миттєвих значень струму та напруги. Відповідно, на резисторі вона, згідно з (6.3) і (6.4), визначається виразом:

$$P_R = iu_R = I_0 U_{R0} \cos^2(\omega t - \varphi). \qquad (6.24)$$

Але, позаяк змінні струми мають високу частоту, практично значимою є тільки ***середня потужність*** $P = \langle p \rangle$, тобто величина

$$P_R = I_0 U_{R0} \langle \cos^2(\omega t - \varphi) = \frac{I_0 U_{R0}}{2} \langle 1 + \cos\left(2(\omega t - \varphi)\right) \rangle. \quad (6.24а)$$

Середнє значення функцій косинус і синус дорівнює нулеві, тому, з урахуванням (6.5),

$$P_R = \frac{I_0 U_{R0}}{2} = \frac{I_0^2 R}{2} = \frac{U_0^2}{2R}. \quad (6.25)$$

Ці формули можна записати зручніше, увівши так звані “діючі” (або “ефективні”) значення струму та напруги, котрі виражаються через амплітуди як

$$I = \frac{I_0}{\sqrt{2}} \text{ і } U = \frac{U_0}{\sqrt{2}}. \quad (6.26)$$

Відтак формули (6.25) набувають вигляду:

$$P_R = I U_R = I^2 R = \frac{U_R^2}{R}, \quad (6.27)$$

і стають такими самими, як і для постійного струму. Отже, можна сказати, що

> ***діючі значення напруги та сили змінного струму – це такі величини напруги та сили постійного струму, при яких у провіднику виділяється така сама потужність, як і при заданому змінному струмі.***

Іншими словами, діючі значення дорівнюють характеристикам постійного струму, котрий є енергетично еквівалентним даному змінному струмові.

Із означень (6.26) випливає, що всі співвідношення, в які входять амплітуди струму й напруги, зберігаються і для діючих значень. Зокрема, це стосується співвідношень, які виражають закон Ома для всього кола (6.17) та окремих ділянок (6.5), (6.7), (6.10):

$$U = IZ, \quad U_R = IR, \quad U_L = IX_L, \quad U_C = IX_C. \quad (6.28)$$

Діючі значення є дуже зручними і тому загально прийнятими на практиці. Зокрема, всі вимірювальні прилади змінного струму проградуйовані в діючих значеннях. Тому в техніці та побуті, коли мова йде про величину змінного струму чи напруги, терміни “діючі значення” спеціально не вживають, і слова “напруга в мережі дорівнює 220 В” або напис на вилці чи розетці “220 В, 6 А” по замовчуванню означають саме діючі значення. Що ж до амплітуди напруги в освітлювальній мережі, то вона складає $U_0 = 220 \cdot \sqrt{2} \approx 311$ В. Ще одне зауваження про терміни. Діючі значення були уведені через потужність змінного струму, що виділяється на резисторі *R*. На відміну від котушки індуктивності та конденсатора, ця потужність виділяється при будь-якому напрямку і фазі φ струму. Можна сказати, що резистор "активно" споживає енергію джерела струму, тому опір *R* називають ***активним опором***. Слід також зауважити, що активний опір в колі змінного струму не є тотожнім до опору в колі постійного струму. Ці величини збігаються тільки в нерухомих колах, де опір *R* зумовлюється виключно розсіюванням ("гальмуванням") носіїв струму при зіткненнях із іншими частинками провідника і є фактором необоротного перетворення енергії електричного струму на тепло. Але, до прикладу, в електричних двигунах споживана енергія струму переходить не тільки в тепло, а й, головно, в механічну роботу. Тому активний опір працюючого двигуна є набагато більшим, ніж нерухомого.

Потужність на котушці індуктивності та на конденсаторі. Суттєва відмінність між реактивними опорами X_L і X_C та активним опором *R* виявляється в процесах перетворення енергії в колі змінного струму. Аби в

цьому переконатися, знайдемо миттєву та середню потужності струму на індуктивності та на ємності. Вказані миттєві потужності визначаються добутком миттєвих значень струму (6.3) та відповідної миттєвої напруги (6.6) і (6.9):

$$p_L = -\frac{1}{2} I_0 U_{L0} \sin(2\omega t - 2\varphi) = -IU_L \sin(2\omega t - 2\varphi) \qquad (6.29)$$

$$p_C = \frac{1}{2} I_0 U_{C0} \sin(2\omega t - 2\varphi) = IU_C \sin(2\omega t - 2\varphi). \qquad (6.29а)$$

І, позаяк середня величина $\langle \sin(2\omega t - 2\varphi) \rangle = 0$, середні потужності на котушці та на конденсаторі

$$P_L = P_C = 0. \qquad (6.30)$$

Цей, можливо, неочікуваний результат має просте пояснення. Ані в котушці індуктивності, ані в конденсаторі не відбувається ніяких перетворень електричної енергії в інші форми, наприклад у тепло. Через це магнітна енергія, яка накопичується в котушці під час збільшення струму від нуля до амплітудного значення, при подальшому зменшенні струму до нуля повністю повертається до генератора чи інших частин кола. Подібні процеси відбуваються й при збільшенні та зменшенні напруги на конденсаторі. Тому, на відміну від активного опору, стосовно індуктивності та ємності слід говорити не про споживання енергії від генератора, а тільки про енергообмін між ними та генератором. Саме з цієї причини котушку індуктивності та конденсатор називають ***реактивними елементами*** кола, а величини X_L і X_C – ***реактивними опорами***.

Повна потужність у колі змінного струму. Згідно з (6.1) і (6.3), миттєва потужність струму у всьому колі

$p = iu = I_0 U_0 \cos(\omega t - \varphi) \cos \omega t.$

За допомогою відомої формули тригонометрії цей вираз можна записати, як

$$p = \frac{1}{2} I_0 T_0 \left(\cos\varphi + \cos(2\omega t - \varphi)\right) = \frac{I_0 U_0}{2} \cos\varphi + \frac{I_0 U_0}{2} \cos(2\omega t - \varphi). \quad (6.31)$$

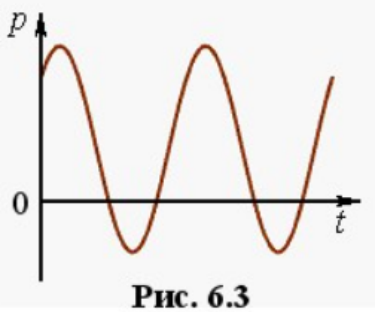


Рис. 6.3

Звідси видно, що миттєва потужність, яка виділяється в колі, коливається навколо рівня $(I_0 U_0 \cos\varphi)/2$ з подвоєною частотою струму. Можливий вигляд графіка $p(t)$, згідно з (6.31), показаний на рис. 6.3. Наявність на графіку ділянок із $p <$ 0 може видатися несподіваною, але легко пояснюється присутністю в колі реактивних опорів. Раніше вже говорилося про те, що, на відміну від резистора, котушка індуктивності та конденсатор не перетворюють енергію струму на інші форми, а лише періодично накопичують, а потім повністю повертають її в коло (див. (6.29), (6.29а)). Причому ці процеси відбувається не синхронно. Тому може статися, що протягом певних частин періоду енергія, яку віддають реактивні елементи перевищує ту, що поглинається на активному опорі. В такому разі "зайву" енергію коло ***віддає*** в генератор, а формальною мовою це означає, що воно ***споживає від'ємну потужність***.

Розглянутий аналіз процесів перетворення енергії в колі є важливими для теорії, але практичне значення має лише ***середня потужність*** струму $P = \langle p \rangle$, яку споживає коло. Тому на практиці під ***потужністю змінного струму***розуміють саме середнє значення величини (6.31):

$$P = \frac{1}{2} I_0 U_0 \langle((cos\varphi + \cos(2\omega t - \varphi))\rangle = \frac{I_0 U_0}{2} \cos\varphi + \frac{I_0 U_0}{2} \langle \cos(2\omega t - \varphi)\rangle.$$

Урахувавши, що $\langle \cos(2\omega t - \varphi)\rangle = 0$, і перейшовши від амплітуд до діючих значень, отримуємо:

$$P = IU \cos\varphi. \qquad (6.32)$$

Отже, потужність, яка виділяється в колі змінного струму, залежить не тільки від сили струму та напруги, а й від зсуву фаз між ними. Відповідно, величина $\cos\varphi$ називається ***коефіцієнтом потужності*** (в інженерній практиці — "косинусом фі". Його вираз можна знайти безпосередньо з векторної діаграми (рис. 6.1) як U_R/U_0, або через формулу (6.20):

$$\cos\varphi = \frac{R}{Z}. \qquad (6.33)$$

Коефіцієнт потужності відіграє важливу роль в електротехніці, оскільки визначає ефективність використання електричної енергії. Справді, якщо $\cos\varphi << 1$, то навіть при великому струмі та напрузі коло майже не споживає енергії з мережі живлення. При цьому генератор працює вхолосту, і необхідні для генерації струму енергетичні ресурси витрачаються марно. Тому коефіцієнт потужності завжди необхідно максимально підвищувати. Для цього треба наближати імпеданс до активного опору, тобто зменшувати реактивну складову імпедансу. З цією метою в коло з великою індуктивністю, яку мають, скажімо, обмотки електричного двигуна, включають компенсуючі конденсатори. Аналогічно в коло з великою ємністю включають додаткові котушки індуктивності. Таким чином, хоча середня потужність на реактивних опорах дорівнює нулеві, ці опори відіграють суттєву роль в енергетиці кола змінного струму. Тому в електротехніці крім потужності (6.32), яку називають ***активною потужністю***, розглядають так звану ***реактивну потужність Q***, яку вимірюють не у ватах, а у "варах" (скорочення від "**в**ольт-**а**мпер **р**еактивний") і визначають за формулою $Q = IU\sin\varphi$. Але детально це питання розглядають не в загальній фізиці, а в електротехніці.

Контрольні запитання

1. Що таке ємнісний опір? Якою формулою він визначається?

2. Що таке індуктивний опір? Якою формулою він визначається?

3. Якою формулою визначається повний реактивний опір кола?

4. Якою формулою визначається імпеданс (повний опір) послідовного кола?

5. Якою формулою визначається імпеданс (повний опір) послідовного кола, в якому немає конденсатора?

6. Якою формулою визначається імпеданс (повний опір) послідовного кола, в якому немає котушки індуктивності?

7. Якою формулою визначається імпеданс (повний опір) ідеального коливального контура?

8. Сформулюйте та запишіть закон Ома для змінного струму.

9. Чи справджується закон Ома для миттєвих значень змінного струму та напруги? Дайте обґрунтовану відповідь.

10. В якому випадку коливання струму в послідовному колі за фазою випереджають коливання напруги генератора?

11. В якому випадку коливання струму в послідовному колі за фазою відстають від коливань напруги генератора?

12. В якому випадку коливання струму в послідовному колі відбуваються в одній фазі з коливаннями напруги генератора?

13. Чи можуть коливання струму за фазою випереджати коливання напруги генератора в колі без конденсатора?

14. Чи можуть коливання струму за фазою відставати від коливань напруги генератора в колі без котушки індуктивності?

15. При якій частоті амплітуда струму в послідовному колі з параметрами *R,L,C* буде максимальною?

16. Від чого та як залежить максимальна можлива амплітуда змінного струму в колі?

17. Отримайте вирази (6.23).

18. Якими формулами визначаються діючі значення змінного струму та напруги та який фізичний зміст вони мають?

19. Чому дорівнює середня потужність змінного струму, що виділяється на активному опорі? Ємнісному опорі? Індуктивному опорі?

20. Якими формулами визначається повна потужність у колі змінного струму?

21. Чи може в колі змінного струму бути від'ємною: а) миттєва потужність струму; б) середня потужність струму?

22. Що таке коефіцієнт потужності та як його можна змінювати для підвищення потужності, що виділяється в колі змінного струму?

Частина ІІ. ХВИЛІ

Коливання як такі є локальним процесом, місце протікання якого обмежується розташуванням і розмірами коливальної системи, до прикладу, маятника чи коливального контура. Але за відповідних умов коливання та інші збурення від місця виникнення поширюються по всіх доступних напрямках, охоплюючи все більшу область простору.

> ***Поширення збурень у просторі з плином часу називається хвильовим процесом, або просто хвилею.***
>
> *(Примітка. Збуреннями називаються різкі локальні відхилення системи від рівноважного тану).*

Визначальною властивістю хвиль є їхня здатність переносити енергію. Достатньо вказати, що цій властивості ми завдячуємо самим своїм існуванням. В усьому іншому хвилі мають багато спільного з локальними коливаннями. Зокрема, хвилі різної природи мають спільні загальні властивості й описуються за допомогою єдиного математичного апарату. Далі розглядаються такі питання:

1. Загальні властивості гармонічних хвиль
2. Пружні хвилі
3. Електромагнітні хвилі

1. Загальні властивості гармонічних хвиль

При єдності загальних властивостей хвилі можуть мати істотні відмінності, пов'язані з їхнім походженням і фізикою процесів, які в них відбуваються. Тому в теорії хвилі класифікуються за відповідними ознаками. До прикладу, окремо розглядають механічні хвилі, з якими, зокрема, пов'язаний звук, та електромагнітні хвилі, що є відповідальні за світло. Крім того, властивості хвиль істотно залежать від умов поширення і можуть бути складними. В даному посібнику розглядаються хвилі в ***ізотропних середовищах***, де характеристики хвиль не залежать від напрямку поширення.

Далі розглядається:

1.1. Поперечні та поздовжні хвилі

1.2. Характеристики та рівняння гармонічних хвиль

1.3. Хвильове рівняння

1.4. Стоячі хвилі

Контрольні запитання

1.1. Поперечні та поздовжні хвилі

Одним з класифікаторів хвилі є напрям коливань, що відбуваються в ній. За цією ознакою хвилі поділяють на поздовжні та поперечні. При цьому

> ***в поздовжній хвилі коливання в кожній точці відбуваються вздовж напрямку поширення.***

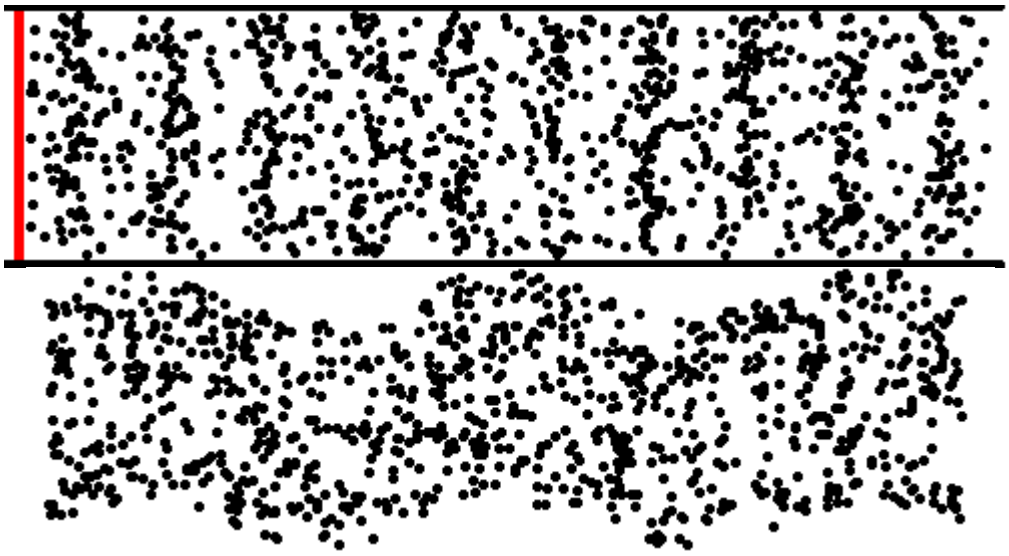
Відповідно,

> ***в поперечній хвилі коливання відбуваються перпендикулярно до напрямку поширення.***

Теорія і дослід свідчать, що електромагнітні хвилі можуть існувати і в речовині, й у вакуумі, і завжди є поперечними. Натомість механічні хвилі, котрі є можливі тільки в речовині, можуть бути як поперечними, так і поздовжніми. При цьому тип хвиль, які можуть виникати, залежить від взаємодії між молекулами середовища. ***У твердому тілі***, де сили зв'язку між атомами є пружними, ***виникають хвилі обох типів***. Проте ***в об'ємі рідини***, де сили взаємодії між молекулами ними є непружними, ***та в газі***, де їх практично немає, ***можуть існувати лише поздовжні хвилі***. Але на поверхні рідини між молекулами діють споріднені з пружними сили поверхневого натягу. Через це на поверхні рідини виникають і поперечні хвилі, які кожен бачив на водоймах.

Нижче наведена “жива картинка”, що схематично зображує поздовжню та поперечну хвилю в середовищі. Якщо уважно до неї придивитися, можна помітити важливу деталь: точки-частинки середовища, не рухаються в напрямку поширення хвилі, а лише здійснюють коливання навколо фіксованих положень. Отже,

> ***у хвилі відбувається не перенесення речовини, а перенесення коливань.***



1.2. *Характеристики та рівняння гармонічних хвиль*

Вище говорилося, що хвилі різної природи мають спільні загальні властивості й описуються за допомогою єдиного математичного апарату. При цьому найпростішими і найважливішими є гармонічні хвилі, що являють собою поширення у просторі гармонічних коливань.

Рівняння плоскої гармонічної хвилі. Уявімо, що в однорідному пружному середовищі розміщена нескінченна площина-мембрана. Якщо вона почне коливатися, то через сили зчеплення частинки прилеглого шару середовища будуть утягувати в рух сусідній шар, він – наступний, і так далі. Як наслідок, коливання від джерела будуть поширюватись, охоплюючи все більшу область простору, тобто, утвориться хвиля. З міркувань симетрії зрозуміло, що в будь-якій площині паралельній до мембрани всі точки середовища будуть рухатись однаково, тому така хвиля називається ***плоскою***.

Нехай мембрана й прилеглі частинки середовища здійснюють коливання вздовж осі OX (рис. 1.1) за законом

$$\xi(0,t) = \xi_0 \cos \frac{2\pi}{T} t,$$

де ξ – зміщення з положення рівноваги, ξ_0 – амплітуда, T - період коливань.

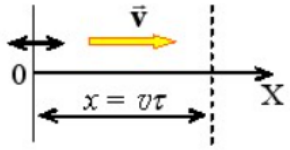


Рис. 7.1

У пружному середовищі відсутні сили тертя, тому з такою самою амплітудою та періодом коливатимуться частинки і в будь-якому іншому шарі. Але, позаяк хвиля поширюється з кінцевою швидкістю, у віддалених шарах частинки почнуть коливатись із деяким запізненням τ , відповідно до рівняння

$$\xi = \xi_0 \cos \frac{2\pi}{T} (t - \tau).$$

Час запізнення дорівнює $\tau = x/v$, де v – швидкість поширення хвилі, x – відстань від джерела до даної точки (див. рис 1.1). Тому зміщення ξ є функцією двох змінних — координати й часу:

$$\xi(x,t) = \xi_0 \cos\frac{2\pi}{T}\left(t - \frac{x}{v}\right) \text{ або } \xi(x,t) = \xi_0 \cos\left(\frac{2\pi}{T}t - \frac{2\pi}{vT}x\right).$$

Це рівняння можна переписати зручніше у вигляді

$$\xi(x,t) = \xi_0 \left(\frac{2\pi}{T}t - \frac{2\pi}{\lambda}x\right), \tag{1.1}$$

або

$$\xi(x,t) = \xi_0 \cos(\omega t - kx), \tag{1.2}$$

де враховано співвідношення $\omega = 2\pi/T$ та введено величини $\boldsymbol{\lambda}$ і $\boldsymbol{k}$, які називаються відповідно ***довжиною хвилі*** та ***хвильовим числом*** і визначаються співвідношеннями:

$$\lambda = vT, \tag{1.3}$$

$$k = \frac{\omega}{v} = \frac{2\pi}{\lambda}. \tag{1.4}$$

Вираз (1.2) (або (1.1)) є ***рівнянням плоскої гармонічної хвилі.*** Воно визначає зміщення ξ частинок середовища у будь-якій точці простору і в будь-який момент часу.

(*Примітка. Для хвилі іншої природи під "зміщенням" слід розуміти збурення тієї фізичної величини, що визначає природу хвилі*).

Розглянута хвиля утворюється праворуч від мембрани й поширюється в додатньому напрямку осі OX. По інший бік мембрани утворюється така сама хвиля, що поширюється в протилежному напрямі. Її рівняння отримаємо з (1.2) заміною v на $-v$:

$$\xi(x,t) = \xi_0 \cos\left(\omega t + \frac{\omega}{v}x\right) = \xi_0 \cos\left(\omega t + \frac{2\pi}{\lambda}x\right) = \xi_0 \cos(\omega t + kx). \tag{1.5}$$

Характеристики гармонічної хвилі. Зрозуміло, що всі характеристики гармонічних коливань одночасно є характеристиками й гармонічної хвилі. Це ***амплітуда ξ_0, період T, циклічна ω або лінійна ν частота та фаза*** (аргумент гармонічної функції у рівнянні хвилі). У цьому зв'язку одразу зауважимо, що коливання в гармонічній хвилі створюються й підтримуються певним джерелом, тобто є вимушеними. Тому

> ***період (частота) хвилі визначається тільки джерелом і не залежить від властивостей середовища.***

Дві інші величини – λ і k – є специфічними саме для хвилі й пов'язані з її поширенням у просторі. Зокрема із (1.3) видно, що

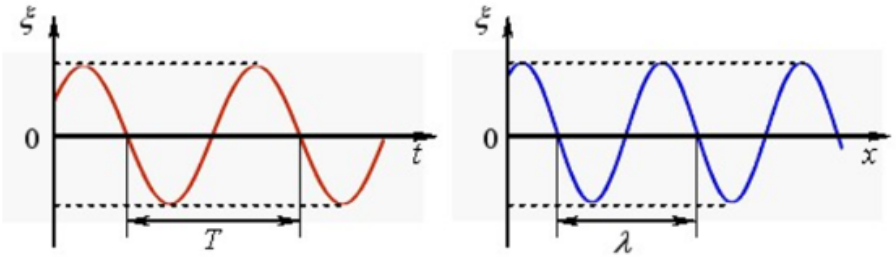
> ***довжина хвилі дорівнює відстані, на яку поширюється хвиля за час одного періоду коливань.***

Для кращого розуміння особливостей хвильового процесу розглянемо залежність величини ξ від часу $\xi(t)$ в заданій точці простору $x = x_0$ та від координати $\xi(x)$ в заданий момент часу $t = t_0$. Згідно з (1.1), ці залежності мають вигляд:

$$\xi(t) = \xi_0 \cos\left(\frac{2\pi}{T}t - \varphi_0\right), \quad \varphi_0 = \frac{2\pi}{\lambda}x_0, \tag{1.6}$$

$$\xi(t) = \xi_0 \cos\left(\frac{2\pi}{T}t - \phi_0\right), \quad \phi_0 = \frac{2\pi}{\lambda}x_0. \tag{1.6a}$$

Графіки $\xi(t)$ та $\xi(x)$ показані на рис. 1.2.



а) б)

Рис. 1.2

Рис. 1.2а показує, що в будь-якій точці простору відбуваються гармонічні коливання величини ξ з періодом T і частотою $\omega = 2\pi/T$, а з рис. 1.2б видно, що в будь-який момент часу значення ξ розподілені у просторі теж за гармонічним законом. При цьому величина λ дорівнює найменшій відстані у напрямку поширення хвилі між точками, в яких коливання відбуваються в однаковій фазі.

(Примітка. Формально зсув фаз складає 2π, але, позаяк частинки в таких точках рухаються однаково, говорять, що вони коливаються в однаковій фазі, або синфазно).

Отже, можна сказати, що

довжина хвилі є просторовим періодом хвилі.

Як і для локальних коливань,

аргумент гармонічної функції в рівняннях (1.1), (1.2) називається фазою хвилі.

Згідно з (1.2), фаза хвилі через частоту та хвильове число визначається, як

$$\varphi = \omega t - kx. \tag{1.7}$$

Фаза визначає і величину ξ, і характер її зміни в різних точках і в різні моменти часу. Тому ***різниця фаз*** $\delta = \varphi_1 - \varphi_2$ в двох точках простору характеризує узгодженість коливань, які збуджуються хвилею в цих точках. Для хвилі в однорідному ізотропному середовищі різниця фаз у двох точках в заданий момент часу дорівнює

$$\delta = k(l_2 - l_1) = \frac{2\pi}{\lambda}\Delta l, \tag{1.7а}$$

де літерою l позначено відстань від джерела до даної точки й ураховано зв'язок (1.4) між хвильовим числом і довжиною хвилі. Співвідношення (1.7а) визначає також і різницю фаз двох різних хвиль у даній точці простору за умови, що вони мають однакові частоти і початкові фази. В цьому випадку величина Δl називається ***різницею ходу*** хвиль; вона є зручною при аналізі різних хвильових явищ.

Хвильові поверхні та фазова швидкість. Фаза хвилі є функцією координат і часу. В даній точці простору вона змінюється з часом. Так само, в даний момент часу фаза змінюється при переході від однієї точки до іншої. При цьому задане значення фази хвиля має не в одній, а в цілій множині точок, які утворюють відповідну "*хвильову поверхню*". Отже,

хвильовою поверхнею називається множина точок у просторі хвилі, в яких у фіксований момент часу фаза хвилі має однакове значення.

В однорідному й ізотропному середовищі, як випливає з (1.7), всі точки хвильової поверхні є рівновіддаленими від джерела. Справді, при заданих величинах $\varphi = \varphi_0$ і $t = t_0$

$$x = \frac{\omega t_0 - \varphi_0}{k} = \text{const}. \qquad (1.8)$$

Через кожну точку з будь-якою координатою x проходить хвильова поверхня із відповідною фазою

$$\varphi_0 = \omega t_0 - kx,$$

отож у просторі хвилі існує нескінченна множина хвильових поверхонь.

Гармонічна хвиля, що тут розглядається, є ідеалізацією: вона займає увесь нескінченний простір і ніяк не обмежена у часі, позаяк у рівнянні (1.2) величини x і t можуть бути якими завгодно. Але реальна хвиля за час від моменту виникнення до заданої миті досягає точок, які віддалені від джерела на певну скінченну відстань. Тому існує хвильова поверхня, що відокремлює область простору, в якій хвиля є, від тієї, куди вона ще не дісталась. Така межова поверхня називається ***фронтом хвилі***.

Із рівнянь (1.7) і (1.8) випливає також, що хвильові поверхні рухаються. Рівняння руху хвильової поверхні отримаємо з (1.8), увівши замість заданого моменту t_0 поточний час t:

$$x = -\frac{\varphi_0}{k} + \frac{\omega}{k}t \Rightarrow v = \frac{\omega}{k}.$$

Отже, хвильові поверхні рухаються із швидкістю

$$v = \frac{\omega}{k}, \qquad (1.9)$$

яка називається ***фазовою швидкістю***.

Зіставлення цього виразу з (1.4) показує, що саме з такою швидкістю поширюється гармонічна хвиля. Тому визначник “фазова” може здатися зайвим. Але в теорії він є необхідним, оскільки реальні хвилі не є строго гармонічними і поширюються з іншою, так званою ***груповою швидкістю***, що збігається фазовою лише в окремих випадках.

Плоскі, циліндричні та сферичні хвилі. Оскільки в однорідному й ізотропному середовищі всі точки хвильової поверхні є рівновіддаленими від джерела, їхня форма визначається геометричною конфігурацією джерела і може бути складною. Але в теорії та практиці найчастіше розглядаються три найпростіші типи хвиль – плоскі, циліндричні та сферичні.

Плоскою хвилею називається така, що має плоскі хвильові поверхні. Саме тому розглянута вище хвиля (1.2), що створюється джерелом у формі нескінченної площини і в усіх точках поширюється в одному напрямку OX, називається плоскою. Відповідно до (1.8), координати точок довільної хвильової поверхні хвилі (1.2) визначаються загальним рівнянням x = const, яке є рівнянням площини, перпендикулярної до осі OX. Отже, хвильові поверхні плоскої хвилі (1.2) складають множину площин, які орієнтовані перпендикулярно до напрямку поширення хвилі (рис. 1.3а).

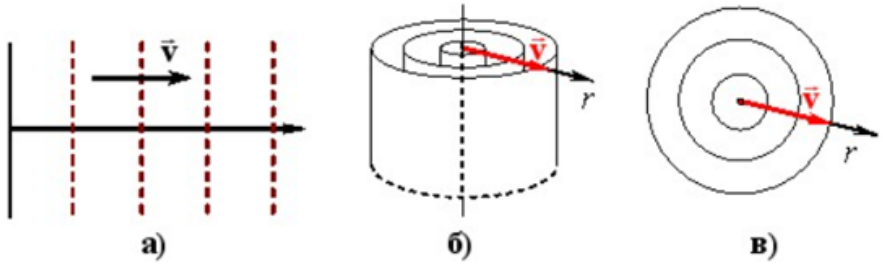


Рис. 1.3

Розглянемо тепер хвилю, що створюється джерелом у вигляді нескінченної прямої пульсуючої нитки. З міркувань симетрії зрозуміло, що така хвиля в однорідному ізотропному просторі поширюється від нитки радіально по всіх напрямках і з однаковою швидкістю. Відповідно, хвильові поверхні такої хвилі являють собою множину коаксіальних циліндрів (рис. 1.3б), через що вона й називається ***циліндричною хвилею***.

Рівняння циліндричної хвилі можна з відповідними корекціями отримати з (1.2). По-перше, циліндрична хвиля поширюється по всіх радіальних напрямках, тому координату x в (1.2) треба замінити на відстань r від даної точки хвилі до нитки:

$$\xi(r,t) = \xi_0 \cos(\omega t - kr).$$

І, по-друге, слід урахувати залежність амплітуди циліндричної хвилі від відстані r до джерела. Джерело має певну потужність і створює у просторі відповідний потік енергії. Тому в середовищі без поглинання крізь будь-яку хвильову поверхню щосекунди проходить однакова енергія Φ_0 .

(Примітка. В ідеальній циліндричній хвилі лінійне джерело вважається нескінченним. Тому під Φ_0 слід розуміти потік енергії, що проходить крізь ділянку циліндричної хвильової поверхні заданої довжини).

Енергія коливань розподіляється по хвильовій поверхні S із поверхневою густиною

$$j = \frac{\Phi_0}{S}.$$

Величина j визначає концентрацію енергії, тож і амплітуду коливань у кожній точці хвильової поверхні. Енергія коливань є прямо пропорційною квадратові амплітуди (див. Частина І, п. 2.4), отже

$$j \sim \xi_0^2 \Rightarrow \xi_0 \sim \sqrt{j} \sim \frac{1}{\sqrt{S}}. \qquad (1.10)$$

У циліндричній хвилі площа ділянки хвильової поверхні довжини l і радіуса r дорівнює $S = 2\pi r l$, Тому можна записати

$$\xi_0 = \frac{a}{\sqrt{r}}, \qquad (1.10а)$$

де a – амплітудний множник залежний від потужності джерела.

Відтак ***рівняння циліндричної хвилі*** набуває вигляду:

$$\xi(r,t) = \frac{a}{\sqrt{r}} \cos(\omega t - kr). \qquad (1.11)$$

Третя з названих простих є ***сферична хвиля.*** Вона створюється в однорідному ізотропному середовищі точковим джерелом і має хвильові поверхні у формі сфер із центром у джерелі (рис.1.3в). Очевидно, що миттєві значення величини $\xi(r, t)$ зменшуються при віддаленні від джерела, тому рівняння сферичної хвилі є подібним до рівняння циліндричної хвилі. Відміна полягає тільки в залежності амплітуди ξ_0 від відстані r до джерела. В сферичній хвилі на відстані r від джерела потік енергії Φ_0 рівномірно розподіляється по поверхні сфери площею $S = 4\pi r^2$, тож густина потоку є обернено пропорційна квадратові відстані: $j \sim (1/r^2)$. Отже, згідно з (1.10) і (1.11),

$$\xi_0 = \frac{a}{r}, \qquad (1.12)$$

і ***рівняння сферичної хвилі*** має вигляд

$$\xi(r,t) = \frac{a}{r}\cos(\omega t - kr). \qquad (1.13)$$

Стосовно отриманих результатів слід зробити деякі зауваження. По-перше, можливі значення координат і часу в рівняннях (1.2), (1.11) і (1.13) формально є нічим не обмежені. Тож вони описують абстрактні хвилі, що охоплюють увесь простір і ніяк не обмежені в часі. Реальні ж хвилі десь і колись починаються та закінчуються і тому мають складніші властивості, зокрема, характеризуються двома швидкостями – фазовою та груповою. Крім того плоске та лінійне джерело не можуть бути нескінченними, тому створити плоску або циліндричну хвилю в усьому просторі є принципово неможливо.

(Правда, за допомогою відповідних технічних засобів їх можна створити в обмеженій області простору. До прикладу, в оптиці це роблять за допомогою лінз або дзеркал).

Неможливим є й необмежене зростання амплітуди ($\xi_0 \to \infty$) циліндричної та сферичної хвиль при наближенні до джерела ($r \to 0$), яке випливає з рівнянь (1.11) і (1.13), оскільки в дійсності лінійне (нитка) та точкове (кулька) джерело мають певний скінченний радіус r_0. Тому рівняння (1.11) і (1.13) реально виконуються тільки *хвильовій зоні* — на відстанях r, які суттєво перевищують r_0.

Хвильовий вектор. Для однієї плоскої хвилі завжди можна вибрати систему координат так, аби хвиля поширювалась уздовж осі OX і описувалася рівнянням (1.2). Але при одночасному розгляді декількох хвиль, які поширюються в різних напрямках, це неможливо. Тому треба встановити загальне рівняння хвилі, тобто визначити функцію $\xi(\vec{r}t)$ у довільній точці з радіусом-вектором $\vec{r}$. Для зручності розглянемо хвилю, що поширюється перпендикулярно до осі OZ у довільному напрямку. Покажемо цей напрямок на рис. 1.4 одиничним вектором (ортом) $\vec{n}$ і проведемо в напрямку $\vec{n}$ допоміжну координатну вісь OX′.

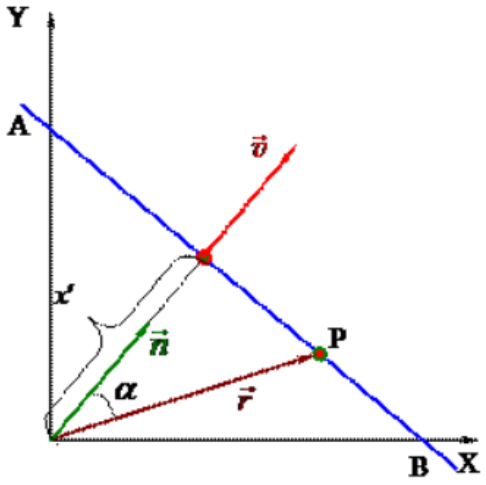


Рис. 7.4

Виберемо тепер якусь точку $\mathrm{P}(\vec{r})$ із радіусом-вектором $\vec{r}$ і проведемо через неї відрізок AB перпендикулярно до осі OX′, який зображує хвильову поверхню, що проходить через точку P і перетинається з віссю OX′ в точці P′. Оскільки обидві точки лежать на одній хвильовій поверхні, коливання в них однакові і, згідно з (1.2), записуються, як:

$$\xi = \xi_0 \cos(\omega t - kx').$$

Із рис. 1.4 видно, що $x' = r\cos\alpha = \vec{n}\cdot\vec{r}$. То ж рівняння коливань в точці $P(\vec{r})$ має вигляд

$$\xi(r,t) = \xi_0 \cos(\omega t - kr\cos\alpha) = \xi_0 \cos(\omega t - k\vec{n}\cdot\vec{r}).$$

Це рівняння є чинним для будь-якої точки та будь-якого напрямку поширення плоскої хвилі й по суті є її загальним рівнянням. Але його можна подати краще, увівши замість хвильового числа $k = 2\pi/\lambda$ ***хвильовий вектор***

$$\vec{k} = \frac{2\pi}{\lambda}\vec{n}, \tag{1.14}$$

який визначає не лише довжину хвилі λ, а й напрям її поширення. Тому ***загальне рівняння плоскої гармонічної хвилі*** у векторній формі записується як

$$\xi(r,t) = \xi_0 \cos(\omega t - \vec{k}\cdot\vec{r}). \tag{1.15}$$

У декартовій системі координат

$$\xi(x,y,z,t) = \xi_o \cos(\omega t - k_x x - k_y y - k_z z). \tag{1.15a}$$

При поширенні плоскої хвилі вздовж осі OX $k_y = k_z = 0$, $k_x = k$, і рівняння (1.15а) трансформується в розглянуте раніше рівняння (1.2). Також, з урахуванням виразів (1.10а) і (1.12), рівняння (1.15) можна застосовувати й для циліндричних та сферичних хвиль.

Загасаючі хвилі. Досі розглядалися хвилі в абсолютно пружному середовищі, де відсутні втрати енергії коливань. Якщо ж середовище поглинає енергію хвилі, то це призводить до додаткового зменшення амплітуди із віддаленням від джерела, і хвиля є загасаючою. Як свідчить дослід, зменшення амплітуди через поглинання відбувається за експоненціальним законом $\xi = \xi_0 e^{-\gamma x}$. Тому рівняння плоскої, сферичної та циліндричної хвиль у середовищі із загасанням, відповідно, мають вигляд:

$$\xi = \xi_0 e^{-\eta x}\cos(\omega t - kx),$$

$$\xi = \frac{A}{r}e^{-\eta r}\cos(\omega t - kr),$$

$$\xi = \frac{A}{\sqrt{r}}e^{-\eta r}\cos(\omega t - kr).$$

1.3. *Хвильове рівняння*

На початку відмічалося, що всі хвильові процеси мають низку спільних властивостей. Це виявляється в тому, що хвильова функція ξ будь-якої хвилі є розв'язком єдиного диференціального рівняння, що називається ***хвильовим рівнянням***. Знайдемо хвильове рівняння незагасаючої хвилі на прикладі плоскої гармонічної хвилі (1.15а). Для цього визначимо другі частинні похідні по часу та по координатах від функції $\xi(x,y,z,t)$:

$$\frac{\partial^2\xi}{\partial t^2} = -\xi_0\omega^2\cos(\omega t - k_x x - k_y y - k_z z) = -\omega^2\xi \Rightarrow$$
$$\xi = -\frac{1}{\omega^2}\frac{\partial^2\xi}{\partial t^2} \tag{1.16}$$

$$\frac{\partial^2\xi}{\partial x^2} = -\xi_0 k_x^2\cos(\omega t - k_x x - k_y y - k_z z) = -k_x^2\xi.$$

Аналогічно отримаємо:

$$\frac{\partial^2\xi}{\partial y^2} = -k_y^2\xi,\ \frac{\partial^2\xi}{\partial z^2} = -k_z^2\xi.$$

Додавши координатні похідні і врахувавши, що $k_x^2 + k_y^2 + k_z^2 = k^2$, одержимо:

$$\begin{gathered}\frac{\partial^2 \xi}{\partial x^2} + \frac{\partial^2 \xi}{\partial y^2} + \frac{\partial^2 \xi}{\partial z^2} = -k^2 \xi \Rightarrow \\ \xi = -\frac{1}{k^2}\left(\frac{\partial^2 \xi}{\partial x^2} + \frac{\partial^2 \xi}{\partial y^2} + \frac{\partial^2 \xi}{\partial z^2}\right).\end{gathered} \tag{1.17}$$

Вираз в дужках записують скорочено, як

$$\frac{\partial^2 \xi}{\partial x^2} + \frac{\partial^2 \xi}{\partial y^2} + \frac{\partial^2 \xi}{\partial z^2} = \Delta \xi = \nabla^2 \xi, \tag{1.18}$$

де символи Δ або ∇^2 позначають ***оператор Лапласа***.

(*Довідка. У математиці терміном "оператор" і відповідним символом позначають задану послідовність математичних дій (операцій), які належить виконати над тою чи іншою функцією. Зокрема, в декартових координатах (1.17) оператор Лапласа передбачає взяття других похідних по координатах від заданої функції $\xi(x, y, z, t)$ вираз буде іншим.*

З урахуванням цього вираз (1.17) можна переписати, як

$$\xi = -\frac{1}{k^2}\nabla^2 \xi.$$

Відтак, порівнюючи його з виразом (1.16), отримуємо:

$$\frac{1}{k^2}\nabla^2 \xi = \frac{1}{\omega^2}\frac{\partial^2 \xi}{\partial t^2}.$$

Звідси випливає ***хвильове рівняння***:

$$\Delta \xi = \frac{1}{v^2} \cdot \frac{\partial^2 \xi}{\partial t^2} \text{ або } \nabla^2 \xi = \frac{1}{v^2} \cdot \frac{\partial^2 \xi}{\partial t^2}, \tag{1.19}$$

в якому параметр $v = (\omega / k)$ є фазовою швидкістю хвилі.

Можна довести, що це диференціальне рівняння виконується не лише для плоскої гармонічної хвилі (1.15а), а й для незагасаючої хвилі будь-якого виду та природи.

1.4 *Стоячі хвилі*

При накладанні хвиль, які поширюються в зустрічних напрямах, за відповідних умов виникає специфічний коливальний процес, відомий як ***стояча хвиля***. Розглянемо утворення стоячої хвилі при накладанні двох плоских гармонічних хвиль, які поширюються у протилежних напрямах осі OX і в яких коливання відбуваються вздовж одного напрямку з однаковою амплітудою ξ_0 і частотою ω, але з певним зсувом фаз α.

Відповідно до (1.2) і (1.5), запишемо рівняння указаних хвиль у вигляді:

$$\xi_1 = \xi_0 \cos(\omega t - kx), \xi_2 = \xi_0 \cos(\omega t + kx + \alpha). \tag{1.20}$$

Відомо, що хвилі задовольняють принцип суперпозиції. Тому при накладанні даних двох хвиль результуючі коливання в кожній точці й у кожен момент часу визначаються як $\xi = \xi_1 + \xi_2$, або розгорнуто:

$$\begin{gathered}\xi = \xi_0 (\cos(\omega t - kx) + \cos(\omega t + kx + \alpha)) = \\ = 2\xi_0 \cos\left(kx + \frac{\alpha}{2}\right) \cos\left(\omega t + \frac{\alpha}{2}\right).\end{gathered} \tag{1.21}$$

Це рівняння можна переписати у вигляді

$$\xi = A\cos\left(\omega t\frac{\alpha}{2}\right), \tag{1.22}$$

де

$$A = 2\xi_0\left|\cos\left(kx + \frac{\alpha}{2}\right)\right|. \tag{1.23}$$

Отже, при накладанні указаних зустрічних хвиль у кожній точці простору виникають гармонічні коливання з частотою ω та залежною від координати амплітудою $A(x)$. Згідно з (1.23), амплітуда коливань періодично змінюється так, що вздовж осі OX утворюється низка ***пучностей*** — точок із максимальною амплітудою $A_п = 2\xi_0$ та ***вузлів*** — точок, у яких амплітуда $A_в = 0$, тобто, коливання відсутні. Визначимо координати таких точок, узявши до уваги, що $k = (2\pi/\lambda)$. Для пучностей маємо:

$$\cos\left(kx + \frac{\alpha}{2}\right) = \pm 1 \Rightarrow kx_п + \frac{\alpha}{2} = n\pi \Rightarrow x_п = \frac{\lambda}{2}\left(n - \frac{\alpha}{2\pi}\right), n = 1, \tag{1.24}$$

2, 3, ...

Виходячи з умови $\cos(kx - (\alpha/2)) = 0$, аналогічно отримаємо координати вузлів:

$$x_в = \frac{\lambda}{2}\left(n + \frac{1}{2} - \frac{\alpha}{2\pi}\right), n = 0, 1, 2, \ldots \tag{1.24а}$$

Координати пучностей і вузлів не залежать від часу. Це означає, що утворені при накладанні зустрічних хвиль коливання (1.22) не поширюються в просторі. Тому такий коливальний процес називається ***стоячою хвилею***. Відповідно, рівняння (1.21) є ***рівнянням плоскої стоячої хвилі***. У цьому контексті хвилю, що поширюється, називають ***біжучою хвилею***. Краще уявити характер коливань у стоячій хвилі та її відміну від біжучої допоможе рис. 1.5, на якому показано графіки $\xi(x)$ при $\alpha = -\pi$ для моментів часу $t_1 < t_2 < t_3$ у випадку біжучої (а) та стоячої (б) хвиль.

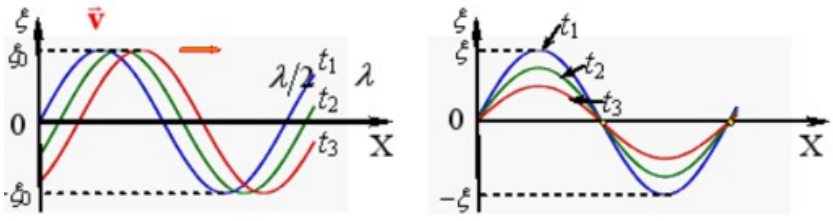


Рис. 1.5

Видно, що в біжучій хвилі коливання зміщуються в напрямку поширення хвилі, а в стоячій хвилі вони не переміщуються в просторі й мають характер пульсацій. Як наслідок, енергія в стоячій хвилі виявляється запертою в областях між вузлами й не переноситься в просторі.

Положення вузлів та пучностей залежить від зсуву фаз α в (1.23), який для двох абстрактних хвиль може бути довільним. Але практично зустрічні хвилі (1.20) утворюють, спрямовуючи плоску хвилю по нормалі на плоску поверхню, що відбиває падаючу хвилю у зворотньому напрямку. В такому разі величина α може дорівнювати тільки 0, або π, залежно від умов відбивання. При відбиванні від більш густого середовища, як, наприклад, при відбиванні звукової хвилі в повітрі від твердої стінки, фаза в точці відбивання змінюється на протилежну, тож $\alpha = \pi$. А при відбиванні від менш густого середовища, приміром, як у випадку звукової хвилі, що поширюється

всередині твердого стержня й відбивається від його торців, фаза не змінюється, і $\alpha = 0$. Тому, відповідно до (1.24) і (1.24а), координати пучностей та вузлів стоячої хвилі у першому випадку визначаються, як

$$x_{\text{п}} = \frac{\lambda}{2}\left(n + \frac{1}{2}\right),\ n = 0, 1, 2, \ldots$$
$$x_{\text{в}} = \frac{\lambda}{2}n,\ n = 0, 1, 2, \ldots \qquad (1.25)$$

а в другому – як

$$x_{\text{п}} = \frac{\lambda}{2}n,\ n = 0, 1, 2, \ldots$$
$$x_{\text{в}} = \frac{\lambda}{2}\left(n + \frac{1}{2}\right),\ n = 0,1,2,\ldots \qquad (1.25а)$$

Із цих виразів випливає, що в точці відбивання ($x = 0$) при відбиванні від більш густого середовища утворюється вузол, а при відбиванні від менш густого середовища – пучність стоячої хвилі. Але за будь-яких умов відстань Δx між сусідніми вузлами або пучностями стоячої хвилі дорівнює

$$\Delta x = \frac{\lambda}{2} = \frac{\pi}{k}. \qquad (1.26)$$

Реально стоячі хвилі утворюються в обмеженому об'ємі, наприклад, при відбиванні від основ пружного стержня. При цьому на обох основах утворюються або вузли, або пучності, тому на довжині стержня вкладається ціла кількість вузлів або пучностей. Це означає, що між двома відбиваючими поверхнями, розміщеними одна від одної на відстані l, можуть утворюватися лише стоячі хвилі з дискретними наборами довжин хвилі λ_n і хвильових чисел k_n, які задовольняють умови:

$$l = n\frac{\lambda_n}{2} \Rightarrow \lambda_n = \frac{2l}{n},$$
$$k_n = \frac{\pi}{l}n,\ n = 0, 1, 2, \ldots \qquad (1.27)$$

Частоти таких хвиль, які інакше називають ***власними частотами***, мають значення

$$\omega_n = \frac{\pi v}{l}n,\ n = 0, 1, 2, \ldots \qquad (1.27а)$$

де v – фазова швидкість хвилі.

Як наслідок, при збудженні й поширенні в обмеженому просторі багатьох коливань із будь-якими можливими частотами, в ньому утворюються система стоячих хвиль із відповідним дискретним набором (спектром) власних частот і довжин хвилі. При цьому частотний спектр стоячих хвиль залежить від форми та розмірів області простору, в якій вони утворюються і у випадку звукових хвиль визначає характерне "забарвлення" або ***тембр*** звуку.

Контрольні запитання

1. Що таке хвиля? Наведіть приклади хвиль різної природи.

2. Які хвилі називаються гармонічними?

3. Що таке довжина хвилі та хвильове число?

4. Як пов'язані між собою довжина хвилі, період коливань і швидкість поширення хвилі? Від чого залежить кожна з цих величин?

5. Що називається фазою хвилі? Чи можна сказати, чому дорівнює фаза хвилі із заданою частотою та довжиною хвилі: а) у даній точці простору ; б) у даний момент часу?

6. Що називається хвильовою поверхнею? Фронтом хвилі?

7. З якою швидкістю рухаються хвильові поверхні? Як ця швидкість називається?

8. Чи визначає фазова швидкість механічної хвилі рух частинок середовища? А групова?

9. Яка відміна існує між термінами "рівняння хвилі" та "хвильове рівняння"? Запишіть хвильове рівняння та загальне рівняння плоскої гармонічної хвилі.

10. Що таке стояча хвиля та за яких умов вона утворюється?

11. Плоска звукова хвиля падає по нормалі на плоску стінку й відбивається. Запишіть координати вузлів та пучностей стоячої хвилі, що утворюється.

2. Пружні хвилі

Поширення механічних коливань на значну відстань від джерела можливе тільки в пружному середовищі, де відсутнє (або є не істотним) перетворення механічної енергії в тепло. Тому механічні хвилі називають ***пружними хвилями***. Пружні хвилі в тілі супроводжуються деформаціями стискання-розтягу або зсуву. При деформаціях першого типу хвилі можуть існувати в будь-якій речовині – твердій, рідкій, чи газоподібній і є поздовжніми. Натомість пружні деформації зсуву можуть виникати тільки в тілах, які мають "пружність форми", тобто здатність зберігати та перешкоджати зміні своєї форми. Така властивість є у твердих тіл, тому в них можливі як поздовжні, так і поперечні хвилі. Рідини, на загал, не зберігають форму, але внаслідок поверхневого натягу таку здатність має їхня вільна поверхня. Тому хвилі в об'ємі рідини є поздовжніми, а на поверхні – поперечними. Що ж до газів, то в них можливі тільки поздовжні хвилі.

Далі розглядаються такі питання:

2.1. Швидкість пружних хвиль

2.2. Енергія та інтенсивність пружних хвиль

2.3. Звукові хвилі

2.4. Ефект Допплера

Контрольні запитання

2.1. *Швидкість пружних хвиль.*

У пружній хвилі переносяться механічні коливання частинок середовища. Але самі частинки не рухаються разом із хвилею, а лише здійснюють коливання навколо фіксованих положень рівноваги. Тому швидкість поширення хвилі в середовищі та швидкість руху його частинок – це різні речі.

У рівнянні механічної хвилі функція ξ визначає зміщення частинок середовища з положення рівноваги, тобто, їхні миттєві положення. Тому швидкість руху частинок середовища можна визначити з рівняння (1.15) шляхом диференціювання по часу:

$$u = \frac{d\xi}{dt} = -\xi_0 \omega \sin\left(\omega t - \vec{k}\vec{r}\right), \qquad (2.1)$$

Частинки середовища у хвилі здійснюють вимушені коливання, частота яких задається джерелом хвилі. Тож і швидкість руху частинок середовища визначається тільки джерелом. Натомість швидкість поширення хвилі (1.9), тобто швидкість передачі коливань частинок середовища від шару до шару, визначається силою пружної взаємодії між частинками та їхньою масою. Тому швидкість поширення хвилі виявляється залежною від властивостей середовища та типу хвилі. Це прямо відображують формули фазової швидкості, які наводимо без доведення.

У твердому тілі фазова швидкість поздовжніх хвиль визначається, як

$$v_1 = \sqrt{\frac{E}{\rho}}, \qquad (2.2)$$

де під коренем стоять табличні константами речовини: – густина і модуль Юнга, який визначає пружні властивості тіла при деформаціях стискання-розтягу. Відповідні розрахунки та вимірювання показують, що для переважної більшості металів і мінералів ця величина потрапляє в інтервал $v_1 = (3 \div 6) \times 10^3$ м/с.

Поперечні хвилі в твердих тілах поширюються із швидкістю

$$v_2 = \sqrt{\frac{G}{\rho}}, \qquad (2.2а)$$

де G — модуль зсуву, що визначає пружні властивості тіла при деформаціях зсуву. Швидкість поперечних хвиль для вказаної вище групи твердих тіл складає v_2 = (2 ÷ 3)×10^3 м/с.

Рідини та гази не зберігають форму, але мають об'ємну пружність, тобто чинять опір зміні об'єму при зміні тиску. При малих деформаціях між зміною тиску dP та відносною зміною об'єму dV/V існує прямо пропорційний зв'язок :

$$\mathrm{d}P = -\kappa \frac{\mathrm{d}V}{V}. \qquad (2.3)$$

Величина κ називається модулем об'ємної пружності. Саме модуль об'ємної пружності разом із густиною визначають швидкість пружних хвиль у рідинах і газах, відповідно до формули:

$$v = \sqrt{\frac{\kappa}{\rho}}. \qquad (2.4)$$

Модуль об'ємної пружності залежить від виду процесу стискання та розширення речовини. При цьому для ідеального газу (газ можна вважати ідеальним при не дуже високих тисках і не наднизьких температурах) його можна визначити теоретично з рівняння процесу. Зокрема, при малій частоті, коли частинки газу рухаються з невеликою швидкістю, процес можна вважати ізотермічним. У такому разі

$$PV = \text{const} \Rightarrow V\mathrm{d}P + P\mathrm{d}V = 0 \Rightarrow \mathrm{d}P = -P\frac{\mathrm{d}V}{V}.$$

Порівнюючи цей вираз із (2.3), маємо $\kappa = P$. При великій частоті хвилі й, відповідно, швидкій зміні тиску процес іде адіабатично. Тому з рівняння адіабати $PV^{\gamma} = \text{const}$ аналогічно виходить $\kappa = \gamma P$, де $\gamma = C_P/C_V$. Урахувавши ці результати та вираз густини ідеального газу $\rho = PM/RT$ через тиск P, температуру T та молярну масу M, отримаємо для швидкості поширення низькочастотних v_1 та високочастотних хвиль у газі:

$$v_1 = \sqrt{\frac{RT}{M}}\ v_2 = \sqrt{\frac{\gamma RT}{M}} \qquad (2.4а)$$

Звукові хвилі мають велику частоту. Тому швидкість звуку в газах визначає друга формула. Зокрема, для звуку в повітрі ($M = 29 \cdot 10^{-3}$кг/моль і $\gamma = 1,4$) при температурі $t = 17^{\circ}$C вона дає 340 м/с, що узгоджується із результатами вимірів. Швидкість звуку в інших газах, за винятком дуже легкого водню, має подібні значення. Рідини порівняно з газами мають великі модулі об'ємної пружності й, відповідно, більші швидкості поширення пружних хвиль. Так, швидкість звуку у воді становить приблизно 1,5·10^3 м/с. В інших рідинах вона лежить в інтервалі (1÷2)·10^3 м/с.

2.2. *Енергія та інтенсивність пружної хвилі*

Об'ємна густина енергії хвилі. Винятково важливою властивістю хвиль є те, що вони переносять енергію від джерела до віддалених точок простору. Достатньо сказати, що практично всі процеси на Землі, включно із самим життям, відбуваються

завдяки енергії, що переноситься від Сонця електромагнітними хвилями. Енергія пружної хвилі складається з кінетичної енергії коливального руху частинок K і потенціальної енергії їхньої взаємодії U у формі енергії пружної деформації середовища, в якому поширюється хвиля.

Розподіл енергії хвилі в просторі визначається ***об'ємною густиною енергії***, тобто, відношенням енергії $\mathrm{d}W$ в елементарному об'ємі $\mathrm{d}V$ до його величини:

$$w = \frac{\mathrm{d}W}{\mathrm{d}V} = \frac{\mathrm{d}K + \mathrm{d}U}{\mathrm{d}V}.$$

Знайдемо величину w на прикладі поздовжньої пружної хвилі. Кінетична енергія $\mathrm{d}K$ в об'ємі $\mathrm{d}V$ визначається швидкістю коливального руху частинок u та їхньою загальною масою $\mathrm{d}m = \rho \mathrm{d}V$, де ρ – густина середовища:

$$\mathrm{d}K = \frac{u^2 \mathrm{d}m}{2} = \frac{\rho u^2}{2} \mathrm{d}V.$$

Для плоскої гармонічної хвилі швидкість u дається виразом (2.1), отже

$$\mathrm{d}K = \frac{\rho \xi_0^2 \omega^2 \mathrm{d}V}{2} \sin^2(\omega t - kx). \qquad (2.5)$$

Потенціальна енергію $\mathrm{d}U$ в об'ємі $\mathrm{d}V$ при пружних деформаціях стиску-розтягу, як відомо з механіки, визначається виразом

$$\mathrm{d}U = \frac{E\varepsilon^2}{2} \mathrm{d}V, \qquad (2.6)$$

де E – модуль пружності (модуль Юнга) і ε – відносна деформація. Для визначення величини ε подумки виділимо в деякій точці x середовища елементарний циліндричний шар із товщиною $\mathrm{d}x$ і основами площею S, які перпендикулярні до напрямку поширення хвилі (рис. 2.1).

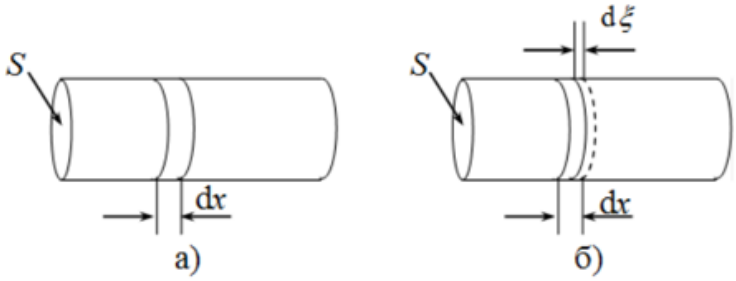


Рис. 2.1

Об'єм цього шару $\mathrm{d}V = S\mathrm{d}x$. Відповідно до рівняння хвилі (1.2), зміщення частинок із положень рівноваги на основах шару будь-якої миті відрізняються на величину

$$|\mathrm{d}\xi| = \left|\frac{\partial \xi}{\partial x}\right| \mathrm{d}x = \xi_0 k \sin(\omega t - kx) \mathrm{d}x,$$

яка є величиною деформації шару. Тоді відносна деформація $\varepsilon = |\mathrm{d}\xi|/\mathrm{d}x$ визначається, як

$$\varepsilon = \frac{\partial \xi}{\partial x} = \xi_0 k \sin(\omega t - kx).$$

Підставляючи цей результат у (2.6), отримуємо:

$$\varepsilon = \frac{\partial \xi}{\partial x},$$

$$\mathrm{d}U = \frac{E\xi_0^2 k^2 \mathrm{d}V}{2}\sin^2(\omega t - kx). \tag{2.7}$$

Виразивши хвильове число як $k = \omega/v$, і взявши для v формулу (2.2) отримаємо для потенціальної енергії:

$$\mathrm{d}U = \frac{\rho\xi_0^2\omega^2 \mathrm{d}V}{2}\sin^2(\omega t - kx). \tag{2.7a}$$

Відтак повна енергія хвилі в об'ємі $\mathrm{d}V$

$$\mathrm{d}W = \mathrm{d}K + \mathrm{d}U = \rho\xi_0^2\omega^2 \mathrm{d}V\sin^2(\omega t - kx),$$

а об'ємна густина енергії

$$w = \frac{\mathrm{d}W}{\mathrm{d}V} = \rho\xi_0^2\omega^2\sin^2(\omega t - kx). \tag{2.8}$$

Густина потоку енергії та інтенсивність хвилі. Із виразу (2.8) видно, що енергія хвилі в будь-який заданий момент часу неоднорідно розподілена в просторі й у будь-якій точці змінюється з часом. При цьому координата точки із заданим значенням густини енергії $w = \mathrm{const}$ задовольняє рівняння

$$\omega t - kx = \mathrm{const} \Rightarrow x = vt - \mathrm{const}\ v = \frac{\omega}{k}.$$

Це означає, що в ідеальній гармонічній хвилі швидкість перенесення енергії від джерела до віддалених точок збігається із фазовою швидкістю (1.9). (Але в реальних хвилях, як говорилося раніше, енергія переноситься з груповою швидкістю, відмінною від фазової).

Енергія, що переноситься хвилею через задану поверхню за заданий проміжок часу залежить як від площі цієї поверхні, так і від величини проміжку часу. Тому кількісною мірою самого процесу перенесення в кожній точці є ***густина потоку енергії***

$$j = \frac{\mathrm{d}W}{\mathrm{d}t\mathrm{d}S_\perp}. \tag{2.9}$$

Можна сказати, що

> ***густина потоку дорівнює енергії, котра переноситься за одиницю часу через одиницю площі поверхні, перпендикулярної до напрямку перенесення.***

Якщо хвиля поширюється із швидкістю v, то за час $\mathrm{d}t$ енергія переноситься на відстань $\mathrm{d}l = v\mathrm{d}t$. Тому при об'ємній густині енергії w через якусь площадку $\mathrm{d}S_\perp$ за час $\mathrm{d}t$ проходить енергія $\mathrm{d}W = w\mathrm{d}S_\perp\mathrm{d}l = w\mathrm{d}S_\perp v\mathrm{d}t$. Відповідно, величина густини потоку енергії дорівнює

$$j = wv. \tag{2.10}$$

В кожній точці енергія переноситься у напрямку вектора швидкості поширення хвилі $\vec{v}$. Аби це відобразити у формулі, густину потоку енергії трактують як вектор

$$\vec{j} = w\vec{v}. \tag{2.11}$$

Розгорнуто він, згідно з (2.8), записується у вигляді:

$$\vec{j} = \rho\xi_0^2\omega^2\vec{v}\sin^2(\omega t - kx). \tag{2.12}$$

Вираз $\vec{j}$ можна подати інакше, використавши тотожність $2\sin^2\alpha = (1 - \cos 2\alpha)$:

$$\vec{j} = \frac{\rho\xi_0^2\omega^2\vec{v}}{2}(1-\cos 2(\omega t - kx)). \qquad (2.12а)$$

Із отриманих виразів випливає, що напрям вектора $\vec{j}$ лишається незмінним, а модуль змінюється з подвоєною частотою хвилі 2ω, тобто перенесення енергії має пульсуючий характер. Але частота хвиль переважно є досить високою, тому реально перенесення енергії визначається

середньою величиною густини потоку енергії, яка називається інтенсивністю хвилі I:

$$I = \langle j \rangle. \qquad (2.13)$$

Згідно з (2.12а), інтенсивність пружної хвилі

$$I = \frac{\rho\xi_0^2\omega^2 v}{2}\langle 1-\cos 2(\omega t - kx)\rangle = \frac{\rho\xi_0^2\omega^2 v}{2}(1-\langle\cos 2(\omega t - kx)\rangle).$$

На інтервалі одного періоду кожному додатньому значенню функції косинус відповідає таке саме по модулю від'ємне значення, тому її середня величина за період дорівнює нулю.

(*Зауважимо, що цей результат є чинним і для будь-якої іншої періодичної знакозмінної функції, графік якої є симетричним відносно осі абсцис*).

Отже, в наведеному виразі другий доданок у дужках дорівнює нулю, й інтенсивність гармонічної хвилі виражається, як

$$I = \frac{1}{2}\rho\xi_0^2\omega^2 v. \qquad (2.14)$$

В однорідному середовищі величини ρ і v скрізь однакові, тому в будь-якій точці інтенсивність хвилі заданої частоти ω визначається тільки квадратом амплітуди:

$$I \sim \xi_0^2. \qquad (2.14а)$$

Тому, коли треба знати не саму інтенсивність, а лише характер її розподілу в просторі, замість величини I (2.14), оперують квадратом амплітуди ξ_0^2.

2.3. *Звукові хвилі*

Характеристики звуку. Пружні хвилі певних частот створюють у людини та інших істот специфічне відчуття звуку й називаються звуковими хвилями. Діапазон частот звукових хвиль, який сприймається людським вухом, складає $(16 \div 20000)$ Гц. Пружні хвилі з частотами меншими за 16 Гц і більшими, ніж 20 кГц називаються інфразвуком та ультразвуком, відповідно.

Суб'єктивне сприйняття звуку залежить від спектрального складу (набору частот) та інтенсивності пружної хвилі, що потрапляє у вухо. Залежно від спектрального складу розрізняють *музичні звуки* та *шуми*. Музичні звуки створюються хвилями з упорядкованим набором частот. Характерною ознакою музичного звуку є ***висота тону*** – суб'єктивне відчуття, завдяки якому ми, до прикладу, легко розрізняємо дзижчання комара та джмеля. За висоту тону є відповідальною частота коливань у звуковій хвилі: чим вона більша, тим вищим є звук. Але й при однаковій висоті тону звуки, що створюються різними джерелами, приміром, голосом чи музичними інструментами, легко пізнаються, бо мають характерне "забарвлення", або ***тембр***. Природа тембру така. При видобуванні звуку, скажімо в скрипці або роялі, через обмежену довжину струн в них виникають власні коливання з різними частотами, які збуджують у порожнині інструменту (резонаторі) відповідні звукові хвилі. Далі при

відбиванні від стінок резонатора й накладанні одні з цих хвиль підсилюються, а інші послаблюються. Відтак у резонаторі встановлюється та випромінюється в простір система звукових хвиль із характерним для даного інструменту (джерела звуку) спектром, тобто, розподілом по частотах та амплітудах. При цьому коливання найнижчої частоти, які називаються ***основним тоном***, мають найбільшу амплітуду й визначають наше сприйняття висоти звуку. Всі інші частоти називаються ***обертонами***, або ***гармоніками***. Набір гармонік і їхній розподіл по амплітудах саме і визначає тембр звуку. Треба зазначити, що хоча людина розрізняє з частотою лише до 20 кГц, більш високочастотні обертони все ж впливають на темброве забарвлення звуку. Тому в пристроях для звукозапису та відтворення звуку намагаються отримати весь діапазон коливань, який дає джерело звуку.

Поряд із музичними звуками існують шуми – звуки з невпорядкованим, переважно неперервним частотним спектром. Такими, зокрема, є звук працюючого двигуна авто чи літака, аплодисменти, приголосні звуки мови, тощо. Шуми сприймаються як беззмістовні звуки, що не мають ані певної висоти, ані тембру. Але суб'єктивне сприйняття шуму все ж таки залежить від характеру розподілу енергії звуку по частотах, подібно до того, як від спектрального складу залежить кольорова гама світла, що сприймається оком. Тому в акустиці, по аналогії з оптикою, шумам умовно приписують різні "кольори". Так, якщо енергія коливань рівномірно розподілена по всьому звуковому діапазону, шум називається білим. Натомість у "рожевому" шумі більш інтенсивними є низькочастотні коливання, а в "блакитному" – високочастотні.

Гучність звуку. Усі звуки характеризуються ***гучністю*** – величиною, що пов'язана з інтенсивністю звукової хвилі , яка потрапляє у вухо. Вухо є дуже чутливим приймачем звуку і характеризується ***порогом чутності*** – інтенсивністю найтихішого звуку, який воно здатне сприймати. Для вуха людини поріг чутності складає всього $I_0 = 10^{-12}$ Вт/м^2. Із ростом інтенсивності звуку зростає і гучність. Але відчуття гучності звуку підсилюється значно повільніше, ніж зростає інтенсивність, так що зв'язок між ними є нелінійним і близьким до логарифмічного. Тому для адекватної характеристики гучності використовують логарифмічну величину, що називається ***рівнем інтенсивності звуку L*** і визначається як

$$L = \lg \frac{I}{I_0}.$$

Вимірюється величина L у ***белах* (Б)**: при L = 1 Б інтенсивність звуку перевищує поріг чутності в 10 разів. Це досить велика різниця, тому на практиці використовують частинну одиницю ***децибел***, котра складає 0,1 бела. То ж

$$L(\text{дБ}) = 10 \lg \frac{I}{I_0}. \qquad (2.15)$$

У цій шкалі рівень інтенсивності звуку, що створюється при тихій розмові на відстані 1 м, складає приблизно 50 дБ, при грі на гітарі на відстані 0,5 м – 70 дБ, на рок-концерті в залі біля сцени – 110 дБ, тощо. Звукові коливання з рівнем інтенсивності > 120 дБ вже не сприймаються як власне звук і створюють у вусі відчуття болю. Тому рівень 120 дБ, якому відповідає інтенсивність 1 Вт/м^2, називають порогом больового відчуття. Такі та більш гучні звуки є шкідливими для здоров'я й можуть бути небезпечним. Зокрема, звук при рівні інтенсивності > 160 дБ може спричинити розрив барабанних перетинок і глухоту, а при рівні > 200 дБ – заподіяти смерть.

Таким чином, діапазон рівнів інтенсивності звуку, що його сприймає людина, простягається від 10^{-12} Вт/м^2 до 1 Вт/м^2, тобто складає дванадцять порядків. Ця цифра вражає. Для порівняння зазначимо, що коли б ми мали фантастичну лінійку з таким

діапазоном вимірювань, то могли би нею визначати відстані, починаючи від розмірів атома (~ 10^{-10} м) й аж до розмірів футбольного поля (~ 100 м).

Рівень інтенсивності є об'єктивною характеристикою сили звуку. Але суб'єктивне сприйняття гучності залежить не тільки від рівня інтенсивності, а й від частоти звуку так, що чутливість вуха є максимальною на частоті приблизно 1000 Гц і помітно зменшується на низьких та на високих частотах. Наприклад, звук із рівнем інтенсивності 30 дБ, який людина добре розрізняє на частоті 1000 Гц, на частоті 100 Гц є вже не чутним. Тому для кількісної характеристики суб'єктивного сприйняття гучності звуку вводять одиницю ***фон***. При цьому приймається, що на частоті 1000 Гц гучність звуку у фонах дорівнює рівню його інтенсивності в децибелах. У такому разі звук з рівнем інтенсивності, скажімо, 60 дБ, на частоті 1000 Гц має гучність 60 фон. Але при такому самому рівні інтенсивності гучність звуку з частотою 50 Гц буде складати вже тільки біля 25 фон.

Відбивання звуку. Пружні хвилі, як і будь-які інші, мають здатність відбиватися від перешкод на своєму шляху й повертатися до місця випромінювання. Цей ефект, який називається ***луною,*** у випадку звукових хвиль сприймається як послаблений повторний звук. В однорідному середовищі, такому, як повітря чи вода, швидкість звуку є сталою. Тому, знаючи величину і вимірюючи час повернення τ відбитого сигналу (відлуння), можна знайти відстань l від джерела звуку до предмета, від якого він відбивається:

$$l = \frac{1}{2} v\tau. \qquad (2.16)$$

На цьому ґрунтується *ехолокація* – спосіб виявлення різних цілей та визначення їхнього положення в просторі за допомогою відбивання ультразвуку.

При відбиванні звуку від декількох перешкод, відстані від яких до джерела суттєво відрізняються, спостерігається декілька послідовних загасаючих лун. Цей ефект створює своєрідний емоційний вплив на слухача й використовується в музичній акустиці шляхом створення в звуковідтворюючій апаратурі послідовності штучних загасаючих сигналів-копій. У приміщеннях, унаслідок многократного відбивання звуку від стін та різних предметів, утворюється велика кількість лун, які розділені дуже малими проміжками часу – (0,01 ÷ 1) мс. Тому вони сприймаються купно, створюючи характерний звуковий шлейф, або відгомін, який називається ***реверберацією***. Завдяки цьому, після припинення дії джерела звук відчувається ще протягом певного ***часу реверберації.*** Реверберація впливає на суб'єктивне сприйняття звуку, створюючи ефект глибини простору. При цьому суттєву роль відіграє час реверберації. Якщо він занадто великий, то послідовні звуки "наповзають" один на одного, що утруднює сприйняття музики й, особливо, мови. Аналогічно, при малому часі реверберації звук стає "сухим" і "тусклим". При цьому кожен різновид звуку має свій оптимальний час реверберації. Наприклад, для камерної та оркестрової музики це приблизно (1÷2) мс, а для органа – набагато більше. Час реверберації залежить від розмірів і конфігурації приміщення та від матеріалу стін і предметів, які знаходяться в ньому. Тому великі приміщення при проектуванні розраховують на оптимальний час реверберації, залежно від їхнього функціонального призначення.

2.4. Ефект Допплера

Раніше відмічалося, що швидкість поширення пружних, зокрема звукових, хвиль визначається тільки властивостями середовища (див. (1.18) - (1.20)), а частота задається джерелом, яке генерує хвилю. При цьому передбачалося, що джерело та

приймач (спостерігач) не рухаються відносно середовища. Але коли джерело або приймач перебувають у русі, частота звуку ν, яку реєструє приймач, на загал, відрізняється від частоти ν_0, яку генерує джерело. Мабуть кожен спостерігав різке пониження висоти звуку сирени локомотива потяга в момент проходження повз спостерігача на великій швидкості

Зміна частоти звуку, яку фіксує приймач, унаслідок його руху відносно джерела, чи джерела відносно нього, називається ефектом Допплера.

Ефект Допплера існує як для звукових, так і для інших, до прикладу, електромагнітних хвиль. Розглянемо теорію класичного ефекту Допплера, коли рух джерела та приймача відбувається з нерелятивістською швидкістю.

Нехай маємо розміщені на осі OX нерухомі приймач Пр і джерело Дж, яке випромінює звукову хвилю у вигляді коротких імпульсів із частотою повторення ν_0 і періодом T_0 і поширюються відносно середовища із заданою швидкістю v. В такому разі відстань між двома послідовними імпульсами (довжина хвилі) складає $\lambda_0 = vT_0 = v/\nu_0$, як схематично показано на рис. 2.2.

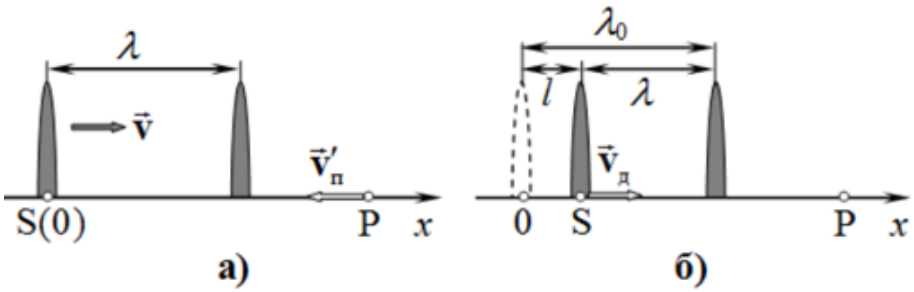


Рис. 2.2

Проаналізуємо, яку частоту звуку буде реєструвати приймач, якщо він і джерело рухаються вздовж осі OX. Розглянемо спочатку окремі випадки.

Випадок 1: рухається тільки приймач. Нехай приймач наближається до нерухомого джерела із швидкістю $u_п$. У такому разі його швидкість відносно імпульсів дорівнює $u'_п = u_п + v$, і відстань λ_0 між послідовними імпульсами приймач долає за час

$$T = \frac{\lambda_0}{u'_п} = \frac{vT_0}{v + u_п}.$$

Цей час є періодом повторення звукових імпульсів, які приходять на приймач, тож частота звуку, яку він реєструє становить

$$\nu = \frac{v + u_п}{v}\nu_0. \qquad (2.17)$$

Випадок 2: рухається тільки джерело (рис. 2.2б). Нехай тепер джерело рухається в бік нерухомого приймача із швидкістю $u_д$, і випускає імпульс 1 у момент t_1. Тоді на момент випускання наступного імпульсу $t_2 = t_1 + T_0$ перший пройде в середовищі відстань $\lambda_0 = vT_0$, а джерело переміститься в тому самому напрямі на відстань $l = u_д T_0 = u_д/\nu_0$. Тож відстань між імпульсами (довжина хвилі) складатиме $\lambda = \lambda_0 - l = (v - u_д)/\nu_0$. Позаяк відносно нерухомого приймача швидкість переміщення імпульсів дорівнює v, вони потраплятимуть на нього з інтервалом часу (періодом)

$$T = \frac{\lambda}{v} = \frac{v - v_д}{v}T_0.$$

Отже, приймач реєструватиме звук із частотою

$$\nu = \frac{v}{(v - u_{\text{д}})}\nu_0. \qquad (2.18)$$

Загальний випадок. Якщо рухаються і приймач, і джерело, то на частоту ν, що реєструється приймачем, впливають обидва фактори. Тому можна записати загальний вираз ν, який вбирає обидві формули (2.17) і (2.18):

$$\nu = \frac{v+u_{\text{п}}}{v - u_{\text{д}}}\nu_0. \qquad (2.19)$$

В обох розглянутих випадках частота звуку, що реєструється приймачем, є більша за частоту звукових імпульсів, які створюються в джерелі. Не важко збагнути, що це зумовлено тим, що в обох випадках відносний рух джерела і приймача був зустрічним. При русі джерела або приймача в протилежному напрямку, у виразах (2.17) – (2.19) треба змінити знак перед відповідною швидкістю. Але цим можна не спеціально не перейматися в кожному окремому випадку, якщо в указаних виразах модулі швидкостей приймача та джерела замінити їхніми проекціями на напрям поширення імпульсів (вісь OX):

$$\nu = \frac{v + u_{\text{п}x}}{v - u_{\text{д}x}}\nu_0. \qquad (2.20)$$

У проведеному аналізі задля більшої наочності розглядалась умовна хвиля у вигляді послідовності імпульсів. Але отримані результати є чинними для хвиль будь-якого вигляду, зокрема для гармонічних хвиль, у яких такими імпульсами слугують “горби”, або “впадини”. Проте слід зазначити, що вираз (2.20) не є загальним, позаяк отриманий лише для випадку, коли джерело та приймач рухаються вздовж однієї прямої. Якщо це не так, зв’язок між частотами є більш складним. Але при не дуже великій відстані між джерелом і приймачем і малій (порівняно із звуковою) швидкості їхнього руху формула (2.20) з достатньою точністю справджується й при довільних напрямах руху джерела чи приймача.

Контрольні запитання

1. Від чого залежить швидкість поширення пружних хвиль? Запишіть формули для швидкості хвиль в твердих тілах, рідинах і газах.
2. У двох рідинах із однаковою густиною, але різною пружністю, поширюються хвилі однакової частоти та амплітуди. В якій рідині інтенсивність хвилі більша?
3. Що таке звук? Укажіть діапазон частот та інтенсивностей звукових хвиль, що їх здатне сприймати вухо людини.
4. Які бувають звуки? Чим відрізняються музичні звуки від шумів?
5. Назвіть основні характеристики музичних звуків. Чим визначаються висота тону та тембр звуку?
6. Чим визначаються гучність та рівень гучності звуку? В яких одиницях вони вимірюються?
7. У чому полягає ефект Допплера? Запишіть загальну формулу, що його виражає.
8. Як змінюється частота звуку, який сприймає нерухомий спостерігач, коли джерело починає наближатися до нього? Віддалятися? Поясніть причину ефекту.
9. Як змінюється частота звуку, що реєструється спостерігачем, який починає рухатися в напрямку до нерухомого джерела? Віддалятися від нього? Поясніть причину ефекту.

3. Електромагнітні хвилі

Із класичної електродинаміки Максвела відомо, що змінні електричне та магнітне поля мають здатність до взаємного перетворення. А саме, якщо в певному місці створити змінне електричне поле, то одразу ж виникне й відповідне змінне магнітне поле, котре створить нове електричне поле, воно — нове магнітне і так далі. Тож змінні електричне та магнітне поля є нерозривно пов'язані між собою й утворюють єдине електромагнітне поле, котре може існувати автономно, незалежно від зарядів або струмів, які були його першоджерелом. При цьому описаний процес взаємоперетворень полів не лишається локалізованим в околі джерела, а охоплює з часом все більшу й більшу область простору, тобто є хвильовим процесом. Отже, за Максвелом, змінне електромагнітне поле існує у формі електромагнітних хвиль (ЕМХ).

> *Зауваження. Таке уявлення є неповним. У дійсності електромагнітна енергія переноситься не неперервною хвилею, а потоком особливих елементарних частинок — квантів, які мають і хвильові, і корпускулярні властивості, (див. ЕЛЕМЕНТИ КВАНТОВОЇ ФІЗИКИ). Тут лише зазначимо, що прояви одних або інших властивостей залежать від частоти випромінювання. Зокрема, радіохвилі практично не показують ознак частинок, а гамма-промені — ознак хвиль. А ось випромінювання оптичного та рентгенівського діапазону в різних явищах проявляє як хвильові, так і квантові властивості.*

Експериментально існування електромагнітних хвиль підтвердив Герц, який першим спостерігав електромагнітні хвилі радіочастотного діапазону, що виникали під час іскрового газового розряду. Електромагнітні хвилі утворюються також і в інших процесах й охоплюють дуже великий інтервал частот, який поділяють на декілька діапазонів, рис. 3.1.

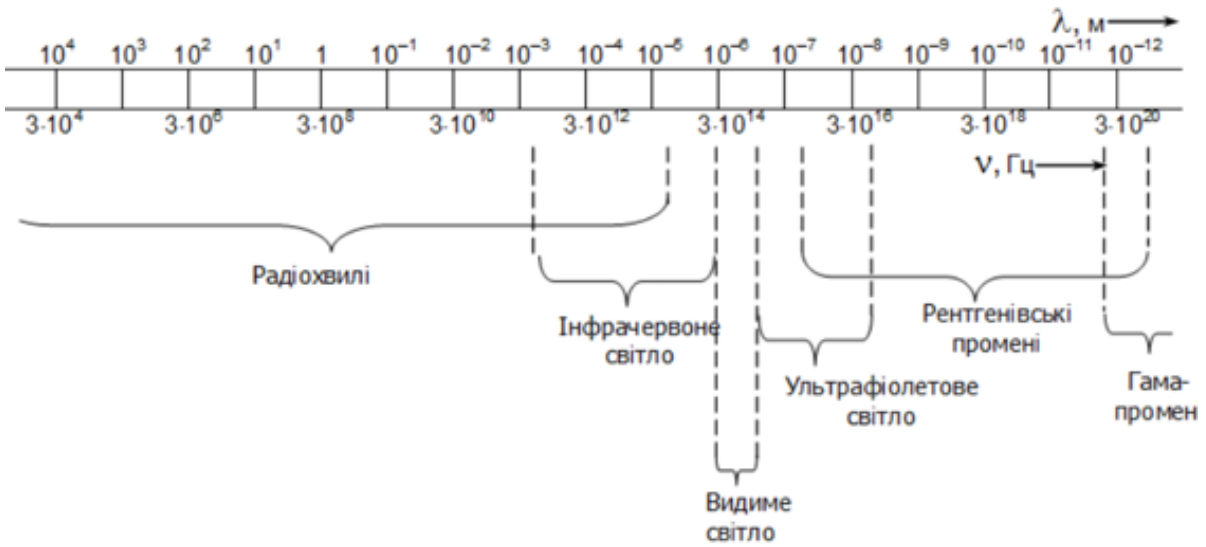


Рис. 3.1

Далі розглядаються наступні питання:

3.1. Характеристики електромагнітних хвиль

3.2. Енергія електромагнітної хвилі

3.3. Властивості електромагнітних хвиль. Додаток

Контрольні запитання

3.1. Характеристики електромагнітних хвиль

Всі властивості електромагнітних хвиль (ЕМХ) випливають із загальних рівнянь електромагнітного поля, але їхній теоретичний аналіз потребує серйозної математичної підготовки. Тому далі основні властивості ЕМХ розглядаються без доведення, а деякі необхідні для формального обґрунтування викладки наводяться в п. 3.3.

Швидкість поширення ЕМХ. У провідному середовищі електричне поле створює електричний струм, енергія якого перетворюється на тепло. Через це коливання електромагнітного поля швидко загасають, тож

електромагнітні хвилі реально можуть існувати тільки в діелектриках і вакуумі.

В цьому випадку з рівнянь Максвелла випливає, що за відсутності вільних зарядів вектори напруженості електричного $\vec{E}$ та магнітного $\vec{H}$ полів задовольняють диференціальні рівняння:

$$\nabla^2 \vec{E} = \varepsilon\varepsilon_0\mu\mu_0 \frac{\partial^2 \vec{E}}{\partial t^2}, \tag{3.1}$$

$$\nabla^2 \vec{H} = \varepsilon\varepsilon_0\mu\mu_0 \frac{\partial^2 \vec{H}}{\partial t^2}, \tag{3.1a}$$

де ε_0, μ_0 — електрична та магнітна сталі, ε, μ — діелектрична та магнітна проникності середовища. Ці рівняння, подібно до (1.4), є хвильовими рівняннями електромагнітного поля. Загальний розв'язок кожного з них $\vec{E}(\vec{r},t)$ або $\vec{H}(\vec{r},t)$ у випадку змінних полів являє собою рівняння відповідної хвилі. Але позаяк електричне та магнітне поля є органічно взаємопов'язані, рівняння (3.1) і (3.1a) в сукупності описують єдину електромагнітну хвилю.

У хвильовому рівнянні (1.4) коефіцієнт при часовій похідній дорівнює оберненому квадрату швидкості, отже у загальному випадку швидкість поширення електромагнітні хвилі

$$v = \frac{1}{\sqrt{\varepsilon\varepsilon_0\mu\mu_0}}, \tag{3.2}$$

а у вакуумі ($\varepsilon = 1$, $\mu = 1$)

$$v_{\text{вак}} = c = \frac{1}{\sqrt{\varepsilon_0\mu_0}}. \tag{3.3}$$

Підстановка у формулу (3.3) значень ε_0 та μ_0 дає $c = 3 \cdot 10^8$ м/с, що збігається із швидкістю світла у вакуумі. Отже, із хвильових рівнянь (3.1) і (3.1a) випливає, що світло має електромагнітну природу. Це теоретичне відкриття Максвелла надалі було підтверджено всім розвитком фізичної оптики.

Враховуючи формулу (3.3), швидкість поширення електромагнітної хвилі у середовищі можна виразити як

$$v = \frac{c}{\sqrt{\varepsilon\mu}}. \tag{3.4}$$

Раніше відмічалося, що електромагнітні хвилі можуть існувати лишень у вакуумі та діелектриках, які переважно є немагнітними середовищами ($\mu = 1$). Тому швидкість електромагнітних хвиль у речовині реально визначається формулою

$$v = \frac{c}{\sqrt{\varepsilon}}. \tag{3.4a}$$

Поперечність ЕМХ і зв'язок між полями. В електромагнітній хвилі коливання напруженостей електричного $\vec{E}$ та магнітного $\vec{H}$ полів відбуваються у взаємно перпендикулярних напрямках, які в той же час є перпендикулярними й до напрямку поширення хвилі (до напрямку вектора $\vec{v}$). Отже,

електромагнітні хвилі є поперечними хвилями.

При цьому вказані вектори утворюють праву трійку, тобто в кожній з послідовностей $(\vec{E},\vec{H},\vec{v})$, $(\vec{H},\vec{v},\vec{E})$ або $(\vec{v},\vec{E},\vec{H})$ останній вектор є спрямований за правим гвинтом відносно перших двох, як показано на рис. 3.2 (вектори, які зображені косо, напрямлені "до нас").

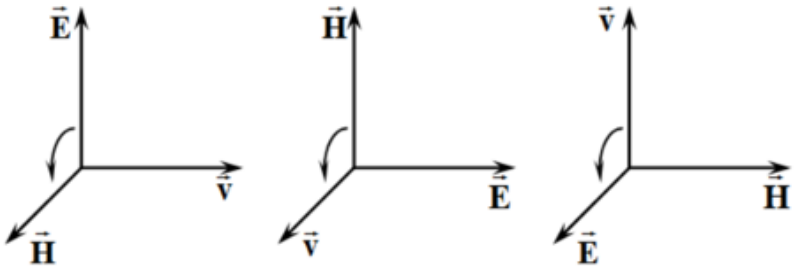


Рис. 3.2

Поперечність електромагнітних хвиль виявляє себе в низці так званих поляризаційних ефектів, які знаходять широке практичне застосування.

Не лише напрямки, а й величини полів в електромагнітній хвилі не є довільними. Величина одного поля однозначно визначає величину іншого так, що за будь-яких умов амплітуди електричного E_0 і магнітного H_0 полів є пов'язані співвідношенням

$$\sqrt{\varepsilon\varepsilon_0}E_0 = \sqrt{\mu\mu_0}H_0. \qquad (3.5)$$

А у вільній хвилі ця умова виконується і для миттєвих значень проекцій напруженостей в кожній точці простору:

$$\sqrt{\varepsilon\varepsilon_0}E(\vec{r},t) = \sqrt{\mu\mu_0}H(\vec{r},t). \qquad (3.5a)$$

Примітка. *Під вільною мається на увазі хвиля, що поширюється у вільному просторі без відбивання від якихось "стінок".*

Це означає, що у вільній електромагнітній хвилі коливання електричного та магнітного полів є синфазними й описуються ідентичними рівняннями. Зокрема, гармонічна електромагнітна хвиля стандартно описується парою рівнянь:

$$\begin{cases} \vec{E}(\vec{r},t) = \vec{E}_0\cos(\omega t - \vec{k}\vec{r}), \\ \vec{H}(\vec{r},t) = \vec{H}_0\cos(\omega t - \vec{k}\vec{r}). \end{cases} \qquad (3.6)$$

Описаний зв'язок між полями ілюструє рис. 3.3, на якому показано розподіл напруженостей полів у якийсь момент часу в напрямку поширення плоскій електромагнітної хвилі, що поширюється в напрямі осі OX.o

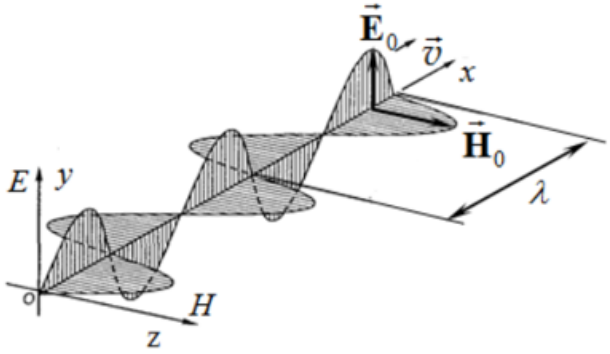


Рис. 3.3

Зауважимо, що завдяки співвідношенню (3.5*а*), кожне з рівнянь (3.6) по суті містить повну інформацію про електромагнітну хвилю. Тому у багатьох задачах можна оперувати лишень одним із них.

3.2. Енергія електромагнітної хвилі

Вектор Пойнтінґа. Електромагнітна хвиля являє собою поширення у просторі змінного електромагнітного поля і тому переносить енергію, що складається з енергії електричного та магнітного полів хвилі. Цей процес у кожній точці простору й у кожен момент часу, як і в механічній хвилі, визначається вектором густини потоку енергії (2.11), який дорівнює добутку об'ємної густини енергії w хвилі та вектора її швидкості $\vec{v}$. Із електродинаміки випливає, що

$$w = w_E + w_H = \frac{\varepsilon\varepsilon_0 E^2}{2} + \frac{\mu\mu_0 H^2}{2}, \tag{3.7}$$

де враховані відомі вирази об'ємної густини енергії електричного та магнітного полів. За допомогою співвідношення (3.5*а*) цей вираз можна записати згорнуто:

$$w = \varepsilon\varepsilon_0 E^2 = \mu\mu_0 H^2 = \sqrt{\mu\mu_0\varepsilon\varepsilon_0}\,EH. \tag{3.7а}$$

Звідси, з урахуванням (3.2), виходить, що

$$w = \frac{1}{v}EH. \tag{3.7б}$$

Відтак величина густини потоку енергії (2.11), яку для електромагнітного поля позначимо літерою Π, визначається, як

$$\Pi = EH. \tag{3.8}$$

Напрям перенесення енергії співпадає з напрямом вектора швидкості поширення хвилі $\vec{v}$, який пов'язаний із напрямками векторів $\vec{E}$ і $\vec{H}$ правилом правого гвинта (рис. 3.2). Тому вектор густини потоку енергії електромагнітної хвилі $\vec{\Pi}$, який називається вектором Пойнтінґа, визначається векторним добутком векторів полів:

$$\vec{\Pi} = [\vec{E}\vec{H}]. \tag{3.8а}$$

Інтенсивність електромагнітної хвилі. У випадку гармонічної хвилі (3.6) модуль вектора Пойнтінґа дорівнює

$$\Pi = E_0 H_0 \cos^2(\omega t - \vec{k}\vec{r}), \tag{3.9}$$

або, з урахуванням (3.5),

$$\Pi = \sqrt{\frac{\varepsilon_0\varepsilon}{\mu_0\mu}}\,E_0^2 \cos^2(\omega t - \vec{k}\vec{r}). \tag{3.9а}$$

Частота електромагнітних хвиль, які знаходять застосування, $\omega \geqslant 10^4$ с$^{-1}$, тож густина потоку енергії дуже швидко змінюється. Тому перенесення електромагнітної енергії в кожній точці простору реально визначається

середньою за період величиною густини потоку, яка називається інтенсивністю хвилі I.

Отже, за означенням $I = \langle\Pi\rangle$ і, згідно з (3.9),

$$I = E_0 H_0 \big\langle \cos^2(\omega t - \vec{k}\vec{r}) \big\rangle = \frac{1}{2} E_0 H_0 \big\langle 1 + \cos\big(2(\omega t - \vec{k}\vec{r})\big) \big\rangle.$$

Оскільки величина

$$\big\langle 1 + \cos\big(2(\omega t - \vec{k}\vec{r})\big) \big\rangle = 1 + \big\langle \cos\big(2(\omega t - \vec{k}\vec{r})\big) \big\rangle = 1,$$

то для інтенсивності електромагнітної хвилі маємо:

$$I = \frac{1}{2} E_0 H_0. \tag{3.10}$$

Аналогічно з (3.9*а*):

$$I = \frac{1}{2} \sqrt{\frac{\varepsilon\varepsilon_0}{\mu\mu_0}} E_0^2. \tag{3.11}$$

Позаяк для вакууму або однорідного середовища величина радикала в усіх точках є однаковою, розподіл інтенсивності в просторі повністю визначається квадратом амплітуди поля:

$$I \sim E_0^2.$$

Це дозволяє в багатьох випадках замість інтенсивності I оперувати величиною E_0^2.

Формула (3.11) дозволяє визначити інтенсивність електромагнітної хвилі тільки через величину електричного поля. Ця обставина широко використовується в оптиці, де більшість ефектів зумовлена саме електричним полем світлової хвилі. Але за допомогою співвідношення (3.5) інтенсивність хвилі при необхідності можна виразити й через амплітуду магнітного поля:

$$I = \frac{1}{2} \sqrt{\frac{\mu\mu_0}{\varepsilon\varepsilon_0}} H_0^2. \tag{3.11*а*}$$

3.3. Властивості електромагнітних хвиль. Додаток

Хвильові рівняння. Існування і загальні властивості електромагнітних хвиль (ЕМХ) можна теоретично довести за допомогою основних рівнянь електромагнітного поля (рівнянь Максвела) в операторній формі (див. ЕЛЕКТРИКА І МАГНЕТИЗМ, частина V, п. 2.4.3). У випадку однорідного ізотропного діелектрика за відсутності вільних зарядів ці рівняння мають вигляд:

$$[\vec{\nabla}\vec{E}] = -\mu_0\mu \frac{\partial \vec{H}}{\partial t}, \tag{3.12}$$

$$\vec{\nabla} \cdot \vec{E} = 0; \tag{3.13}$$

$$[\vec{\nabla}\vec{H}] = \varepsilon\varepsilon_0 \frac{\partial \vec{E}}{\partial t}, \tag{3.14}$$

$$\vec{\nabla} \cdot \vec{H} = 0. \tag{3.15}$$

Наведені рівняння містять тільки перші похідні, тоді як загальне хвильове рівняння (розділ II, п. 1.3) пов'язує другі координатні та часові похідні відповідної характеристики хвилі. Тому спочатку "продиференціюємо по координатах" рівняння (3.12), взявши ротор від лівої та правої частин :

$$\left[\vec{\nabla}[\vec{\nabla}\vec{E}]\right] = -\mu_0\mu \left[\vec{\nabla}, \frac{\partial \vec{H}}{\partial t}\right] \quad \Rightarrow \quad \left[\vec{\nabla}[\vec{\nabla}\vec{E}]\right] = -\mu_0\mu \frac{\partial}{\partial t}[\vec{\nabla}\vec{H}].$$

(*У правій частині використано те, що диференціювання по незалежних змінних можна проводити в довільному порядку.*)

Далі, врахувавши рівняння (3.14), запишемо:

$$\left[\vec{\nabla}[\vec{\nabla}\vec{E}]\right] = -\mu_0\mu \frac{\partial^2 \vec{E}}{\partial t^2}. \tag{3.16}$$

І нарешті розкриємо ліву частину за відомим правилом "бац мінус цаб" для подвійного векторного добутку: $\left[\vec{a}[\vec{b}\vec{c}]\right] = \vec{b}(\vec{a}\vec{c}) - \vec{c}(\vec{a}\vec{b})$. Відтак отримаємо:

$$\left[\vec{\nabla}[\vec{\nabla}\vec{E}]\right] = \vec{\nabla}(\vec{\nabla}\vec{E}) - (\vec{\nabla}\vec{\nabla})\vec{E}.$$

Урахувавши, що в цьому виразі $(\vec{\nabla}\vec{E}) = 0$, згідно з (3.13), і $(\vec{\nabla}\vec{\nabla}) = \nabla^2$ — оператор Лапласа, отримуємо з (3.16) хвильове рівняння (3.1) для вектора $\vec{E}$:

$$\nabla^2\vec{E} = \varepsilon_0\varepsilon\mu_0\mu\frac{\partial^2\vec{E}}{\partial t^2}.$$

Так само виводиться й рівняння (3.1*а*) для магнітного вектора електромагнітної хвилі.

Поперечність ЕМХ. Для спрощення викладок розглянемо довільну плоску хвилю, що поширюється в напрямку координатної осі OX і, отже, має плоскі хвильові поверхні перпендикулярні до вказаної осі. Тому поля $\vec{E}$ і $\vec{H}$ залежать тільки від координати x і їхні похідні по y та по z дорівнюють нулю. Тоді наведені рівняння спрощуються і в координатній формі записуються, як:

$$-\vec{e}_y\frac{\partial E_z}{\partial x} + \vec{e}_z\frac{\partial E_y}{\partial x} = -\mu_0\mu\left(\vec{e}_x\frac{\partial H_x}{\partial t} + \vec{e}_y\frac{\partial H_y}{\partial t} + \vec{e}_z\frac{\partial H_z}{\partial t}\right), \tag{3.17}$$

$$\frac{\partial E_x}{\partial x} = 0, \tag{3.18}$$

$$-\vec{e}_y\frac{\partial H_z}{\partial x} + \vec{e}_z\frac{\partial H_y}{\partial x} = \varepsilon_0\varepsilon\left(\vec{e}_x\frac{\partial E_x}{\partial t} + \vec{e}_y\frac{\partial E_y}{\partial t} + \vec{e}_z\frac{\partial E_z}{\partial t}\right), \tag{3.19}$$

$$\frac{\partial H_x}{\partial x} = 0, \tag{3.20}$$

де $\vec{e}_x, \vec{e}_y, \vec{e}_z$ — орти (одиничні вектори) координатних осей.

Проаналізуємо ці рівняння. Вектори в лівих частинах рівнянь (3.17) і (3.19) не мають x-компонент, отже, не мають їх і вектори в правих частинах. Тому $(\partial H_x/\partial t) = (\partial E_x/\partial t) = 0$, і, врахувавши (3.18) і (3.20), маємо:

$$\begin{cases} \dfrac{\partial E_x}{\partial x} = 0, & \dfrac{\partial E_x}{\partial t} = 0; \\ \dfrac{\partial H_x}{\partial x} = 0, & \dfrac{\partial H_x}{\partial t} = 0. \end{cases}$$

Це означає, що $E_x(x,t) = \text{const}$ і $H_x(x,t) = \text{const}$, тобто складові в напрямку осі OX можуть мати лишень однорідні та стаціонарні (статичні) електричне й магнітне поля. Поля в електромагнітній хвилі такими не є, отже, в електромагнітній хвилі, що розглядається,

$$\begin{cases} E_x = 0 & \Rightarrow \quad \vec{E}\perp\vec{e}_x, \\ H_x = 0 & \Rightarrow \quad \vec{H}\perp\vec{e}_x. \end{cases}$$

Таким чином, обидва поля є перпендикулярними до напрямку поширення електромагнітної хвилі, тобто, вона є поперечною.

Аби з'ясувати взаємну орієнтацію векторів $\vec{E}$ і $\vec{H}$, уявімо, що електричне поле хвилі напрямлене вздовж осі OY: $\vec{E} = \vec{e}_yE_y$, тож $E_x = E_z = 0$. Тоді рівняння (3.17) і (3.19) набувають вигляду:

$$\begin{cases} \vec{e}_z\dfrac{\partial E_y}{\partial x} = -\mu_0\mu\left(\vec{e}_y\dfrac{\partial H_y}{\partial t} + \vec{e}_z\dfrac{\partial H_z}{\partial t}\right), \\ -\vec{e}_y\dfrac{\partial H_z}{\partial x} + \vec{e}_z\dfrac{\partial H_y}{\partial x} = \varepsilon_0\varepsilon\vec{e}_y\dfrac{\partial E_y}{\partial t}. \end{cases} \tag{3.21}$$

Оскільки у векторному рівнянні напрямки векторів у лівій та правій частинах мають бути однакові, робимо висновок, що в (3.21)

$$H_y = 0 \quad \Rightarrow \quad \vec{H} = \vec{e}_z H_z.$$

Отже, якщо $\vec{E} = \vec{e}_y E_y$, то $\vec{H} = \vec{e}_z H_z$. Таким чином, вектори $\vec{E}$ і $\vec{H}$ є взаємно перпендикулярними та перпендикулярними до напрямку поширення хвилі й утворюють праву трійку, як показано на рис. 3.2.

Хвильові рівняння і швидкість поширення. У щойно розглянутому випадку рівняння (3.21) набувають вигляду:

$$\begin{cases} \dfrac{\partial E_y}{\partial x} = -\mu_0\mu \dfrac{\partial H_z}{\partial t}, \\ -\dfrac{\partial H_z}{\partial x} = \varepsilon_0\varepsilon \dfrac{\partial E_y}{\partial t}. \end{cases} \tag{3.21a}$$

Продиференціювавши перше рівняння по координаті, а друге по часу, дістанемо:

$$\begin{cases} \dfrac{\partial^2 E_y}{\partial x^2} = -\mu_0\mu \dfrac{\partial^2 H_z}{\partial x \partial t}, \\ -\dfrac{\partial^2 H_z}{\partial x \partial t} = \varepsilon_0\varepsilon \dfrac{\partial^2 E_y}{\partial t^2} \end{cases} \quad \Rightarrow \quad \frac{\partial^2 E_y}{\partial x^2} = \varepsilon_0\varepsilon\mu_0\mu \frac{\partial^2 E_y}{\partial t^2},$$

а це є хвильове рівняння для електричної компоненти електромагнітної хвилі. Аналогічно отримується й хвильове рівняння для магнітної компоненти електромагнітної хвилі:

$$\frac{\partial^2 H_z}{\partial x^2} = \varepsilon_0\varepsilon\mu_0\mu \frac{\partial^2 H_z}{\partial t^2},$$

При цьому для фазової швидкості виходить вираз (3.2):

$$v = \frac{1}{\sqrt{\varepsilon\varepsilon_0\mu\mu_0}}.$$

Зв'язок між полями. Наостанку на прикладі гармонічної хвилі знайдемо зв'язок між величиною полів E та H. Нехай електричне поле хвилі визначається, як

$$E_y(x,t) = E_0 \cos(\omega t - kx).$$

Тоді з першого рівняння (3.15a) маємо:

$$\frac{\partial H_z}{\partial t} = \frac{kE_0}{\mu_0\mu}\sin(\omega t - kx).$$

Отже, коливання магнітного поля в якійсь заданій точці $x = x_0$ задовольняють рівняння:

$$\frac{\mathrm{d}H_z}{\mathrm{d}t} = \frac{kE_0}{\mu_0\mu}\sin(\omega t - \varphi_0), \qquad \varphi_0 = kx_0.$$

Проінтегрувавши це рівняння, знайдемо залежність магнітного поля хвилі від часу:

$$H_z(t,x_0) = \frac{kE_0}{\mu_0\mu}\int \sin(\omega t - \varphi_0)\,\mathrm{d}t = \frac{k}{\mu_0\mu\omega}E_0\cos(\omega t - \varphi_0) = \frac{k}{\mu_0\mu\omega}E_y(t,x_0).$$

Із виразів (1.4) і (3.2) випливає, що

$$\frac{k}{\mu_0\mu\omega} = \sqrt{\frac{\varepsilon_0\varepsilon}{\mu_0\mu}},$$

отже,

$$H_z(t, x_0) = \sqrt{\frac{\varepsilon_0 \varepsilon}{\mu_0 \mu}} E_y(t, x_0). \tag{3.22}$$

Даний результат є чинним при будь-яких значеннях x_0, тому в будь-якій точці простору й у будь-яку мить між величинами електричного та магнітного полів у гармонічній існує співвідношення (3.15a):

$$\sqrt{\varepsilon \varepsilon_0} E(x, t) = \sqrt{\mu \mu_0} H(x, t).$$

Звернімо увагу й на те, що знаки проекцій E_y і H_z у (3.22) завжди однакові. А це означає, що у вільній електромагнітній хвилі коливання полів відбуваються в однаковій фазі.

На завершення вкажемо, що такі самі результати можна отримати й для вільної хвилі і з довільним законом коливань $E = f\left(\omega t - kx\right)$.

Контрольні запитання

1. Від чого залежить швидкість електромагнітних хвиль? Скориставшись довідковими даними, обчисліть швидкість поширення електромагнітних хвиль у вакуумі.
2. До якого типу належать електромагнітні хвилі? Яка взаємна орієнтація векторів полів $\vec{E}$, $\vec{H}$ і швидкості поширення $\vec{v}$ в електромагнітній хвилі?
3. Запишіть співвідношення між амплітудами електричного та магнітного поля електромагнітної хвилі. Чи завжди воно виконується?
4. Запишіть співвідношення між миттєвими значеннями електричного та магнітного поля електромагнітної хвилі. Чи завжди воно виконується?
5. Який зміст має вектор Пойнтінґа та як він визначається? Що таке інтенсивність електромагнітної хвилі?
6. Запишіть вираз інтенсивності електромагнітної хвилі через амплітуду напруженості електричного поля хвилі. Знайдіть величину та розмірність числового коефіцієнта в цьому виразі.

ЕЛЕМЕНТИ ОПТИКИ

Зміст

ВСТУП

5.1. ВІДБИВАННЯ ТА ЗАЛОМЛЕННЯ СВІТЛА
1.СВІТЛОВІ ХВИЛІ
2. СВІТЛОВІ ХВИЛІ НА МЕЖІ ПОДІЛУ СЕРЕДОВИЩ. ВІДБИВАННЯ ТА ЗАЛОМЛЕННЯ
3. АМПЛІТУДИ І ФАЗИ ВІДБИТОЇ ТА ЗАЛОМЛЕНОЇ ХВИЛЬ

5.2. ІНТЕРФЕРЕНЦІЯ СВІТЛА
1. ДВОПРОМЕНЕВА ІНТЕРФЕРЕНЦІЯ
2. ІНТЕРФЕРЕНЦІЙНА КАРТИНА
3. Інтерференційні максимуми та мінімуми
4. ІНТЕНСИВНІСТЬ СВІТЛА ТА ШИРИНА ІНТЕРФЕРЕНЦІЙНИХ СМУГ
5. ПРОБЛЕМИ КОГЕРЕНТНОСТІ
6. КОНТРОЛЬНІ ЗАПИТАННЯ

5.3. СПОСТЕРЕЖЕННЯ ІНТЕРФЕРЕНЦІЇ СВІТЛА
1. ІНТЕРФЕРЕНЦІЙНІ СХЕМИ
2. ІНТЕРФЕРЕНЦІЯ В ТОНКИХ ПЛАСТИНАХ
КОНТРОЛЬНІ ЗАПИТАННЯ

5.4. ДИФРАКЦІЯ СВІТЛА
1. ЯВИЩЕ ДИФРАКЦІЇ. ПРИНЦИП ГЮЙГЕНСА-ФРЕНЕЛЯ
2. ДИФРАКЦІЯ ФРАУНГОФЕРА НА ЩІЛИНІ
3. ДИФРАКЦІЯ ФРАУНГОФЕРА НА ҐРАТЦІ
4. ДИФРАКЦІЯ РЕНТГЕНІВСЬКИХ ПРОМЕНІВ НА КРИСТАЛАХ
5. КОНТРОЛЬНІ ЗАПИТАННЯ

5.5. ПОЛЯРИЗАЦІЯ СВІТЛА
1. Поляризоване та природне світло
2. ОТРИМАННЯ ЛІНІЙНО ПОЛЯРИЗОВАНАНОГО СВІТЛА
3. ЕЛІПТИЧНА ТА КОЛОВА ПОЛЯРИЗАЦІЯ
4. ОТРИМАННЯ ЕЛІПТИЧНО ТА ЦИРКУЛЯРНО ПОЛЯРИЗОВАНОГО СВІТЛА
5. АНАЛІЗ ПОЛЯРИЗОВАНОГО СВІТЛА
6. КОНТРОЛЬНІ ЗАПИТАННЯ

5.6. ПОШИРЕННЯ СВІТЛА В РЕЧОВИНІ
1. ДИСПЕРСІЯ СВІТЛА
2. ГРУПОВА ШВИДКІСТЬ
3. ПОГЛИНАННЯ ТА РОЗСІЮВАННЯ СВІТЛА
4. КОНТРОЛЬНІ ЗАПИТАННЯ

ВСТУП

Оптика – це розділ фізики, в якому вивчаються властивості світла та його взаємодія з речовиною. А світлом називають електромагнітне випромінювання, що здатне створювати в оці людини та інших істот зорове відчуття.

Електромагнітну природу світла відкрив Максвелл, який теоретично встановив, що швидкість світла у вакуумі збігається із швидкістю поширення електромагнітних хвиль. Відтак закони електродинаміки лягли в основу фізичної оптики і дозволили отримати вичерпне пояснення багатьох хвильових оптичних явищ, зокрема тих, що розглядаються в наступних лекціях. Але відразу зауважимо, що існують і так звані квантові оптичні явища, в яких світло виявляє властивості не хвиль, а частинок (див. Розділ 6). Далі розглядаються такі питання:

5.1. Відбивання і заломлення світла

5.2. Інтерференція світла

5.3. Спостереження інтерференції світла

5.4. Дифракція світла

5.5. Поляризація світла

5.6. Поширення світла в речовині

5.1. ВІДБИВАННЯ ТА ЗАЛОМЛЕННЯ СВІТЛА

Нижче будуть розглянуті такі питання

1. Світлові хвилі

2. Світлова хвиля на межі поділу. Відбивання та заломлення

3. Амплітуди і фази відбитої та заломленої хвиль

1.СВІТЛОВІ ХВИЛІ

Загальні характеристики хвильових процесів, зокрема електромагнітних, а також і світлових хвиль, розглянуті в попередньому розділі (Лекції 4.7 і 4.9). Але для світла, на додачу, існує низка понять і величин, які пов'язані з його суб'єктивним сприйняттям.

Оптичне випромінювання. Інтервали частот *n* і довжин хвилі *l* електромагнітного випромінювання, яке ми сприймаємо як видим світло, дуже вузькі і дещо відрізняються у різних людей. Усереднено вони складають:

$$\nu = (0,75 \div 0,4) \text{ ПГц},$$
$$\lambda = (0,4 \div 0,75) \text{ мкм} = (400 \div 750) \text{ нм}. \qquad (1.1)$$

(Нагадаємо, що 1 ПГц = 10^{15} Гц, 1 мкм = 10^{-6} м, 1 нм = 10^{-9} м).

Суб'єктивно світло сприймається через відчуття яскравості джерела чи освітленої поверхні та через відчуття кольору. Об'єктивно яскравість та освітленість визначається інтенсивністю світлової хвилі, що потрапляє в око людини. За відчуття кольору відповідає частота світла, тож будь-якій частоті з означеного діапазону (1.1) формально відповідає один окремий колір. Тому світло, що характеризується однією частотою, називають монохроматичним («однокольоровим»). Одразу зазначимо, що поняття монохроматичного світла відіграє важливу роль у теорії, але є ідеалізацією. Насправді в будь-якому реальному випромінюванні присутні коливання із безліччю різних частот в певному інтервали $\Delta\nu$, який визначає спектральну ширину цього випромінювання. Тому на практиці за монохроматичне приймають випромінювання з достатньо вузькою спектральною шириною. Слід також зауважити, що око не здатне розрізняти за кольором світло з дуже близькими частотами. Тому умовно виділяють усього сім «чистих» кольорів, які в напрямку зростання частоти світла утворюють відому послідовність: червоний, жовтогарячий (оранжевий, помаранчевий), жовтий, зелений, блакитний, синій та фіолетовий. Проте завдяки досить великій роздільній здатності ока та існуванню трьох типів рецепторів кольору («колбочок») із максимальною чутливістю до різних частот людина здатна розрізняти дуже широку гаму кольорів.

Швидкість світла. Показник заломлення. Теоретично в речовині швидкість електромагнітних хвиль, тож і світла, визначається її діелектричною $varepcilon$ та магнітною μ проникностями (Лекція 4.9, формула (9.4)). Але через сильне загасання електромагнітних коливань у провідних середовищах прозорими для світлових хвиль є тільки діелектрики, що в переважній більшості є немагнітними речовинами ($\mu = 1$. Тому практично швидкість світла в речовині визначається формулою

$$v = \frac{c}{\sqrt{\varepsilon}}, \qquad (1.2)$$

де $c = 3 \cdot 10^8$ м/с – швидкість світла у вакуумі.

Із зміною швидкості поширення світла при переході з одного середовища в інше пов'язаний один з основних оптичних ефектів – заломлення світла. Для його описання використовують величину

$$n = \sqrt{\varepsilon} = \frac{c}{v}, \qquad (1.3)$$

яку називають показником заломлення, або оптичною густиною прозорої речовини. Відтак, швидкість світла в речовині визначається формулою

$$v = \frac{c}{n}. \qquad (1.4)$$

Оскільки частота ν і період T не залежать від середовища, в якому поширюється хвиля, то зміна швидкості v при переході світла з вакууму в речовину супроводжується зміною довжини хвилі. У вакуумі довжина хвилі $\lambda_0 = cT$, а в середовищі $\lambda = cT$, отож

$$\lambda = \frac{\lambda_0}{n} \text{ і } n = \frac{\lambda_0}{\lambda}. \qquad (1.5)$$

Співвідношення (1.3) та (1.5) можна розглядати як означення показника заломлення.

Монохроматична світлова хвиля. В усякі електромагнітній, тож і світловій хвилі, напруженості електричного $\vec{E}(\vec{r},t$ та магнітного $\vec{H}(\vec{r},t)$ полів визначаються ідентичними рівняннями (див. Лекція 4.9, п. 9.1.2). При цьому у вакуумі та вільному (такому, що не містить відбиваючих поверхонь), ізотропному середовищі величини полів пов'язані між собою співвідношенням:

$$\sqrt{\varepsilon_0 \varepsilon} E(r,t) = \sqrt{\mu_0 \mu} H(r,t). \qquad (1.6)$$

До того ж, у переважній більшості оптичних явищ роль магнітного поля є не істотною. Тому зазвичай світлову хвилю описують тільки рівнянням для електричного поля, а вектор $\vec{E}(\vec{r},t)$ називають ***світловим вектором***.

Для теорії велике значення мають монохроматичні світлові хвилі – ідеальні хвилі, в яких коливання полів відбувається за законом косинуса або синуса (Лекція 4.7, п. 7.2.1). Коливання світлового вектора в такій хвилі описується рівнянням

$$\vec{E}(\vec{r},t) = \vec{E}_0 \cos(\omega t - \vec{k}\cdot\vec{r} + \varphi_0), \qquad (1.7)$$

в якому:

- вектор $\vec{E}_0$ визначає амплітуду E_0 і напрям коливань напруженості електричного поля хвилі;

- колова частота ω (рад/с) (або лінійна частота $\nu = \omega/2\pi$ (Гц), чи період $T = 1/\nu = 2\pi/\omega$) визначає часову циклічність змін поля хвилі в заданій точці простору;

- поточна ***фаза*** $\varphi(x,t) = \omega t - \vec{k}\cdot\vec{r} + \varphi_0$ визначає стан поля хвилі в будь-яку мит і в в будь-які точці; відповідно, ***початкова фаза*** визначає фазу в точці $\vec{r} = 0$ у момент часу t = 0;

- [хвильовий вектор $\vec{k}$ визначається виразом

$$\vec{k} = \frac{\omega}{v}\vec{n} = \frac{2\pi}{\lambda}\vec{n}, \quad k = \frac{\omega}{v} = \frac{2\pi}{\lambda}, \qquad (1.8)$$

де v – величина швидкості, $\vec{n}$ – одиничний вектор (орт) напрямку поширення хвилі в даній точці. Модуль хвильового вектора називають хвильовим числом;

- довжина хвилі λ дорівнює найменшій відстані у напрямку поширення, на якій повторюється стан поля хвилів заданий момент часу; довжина хвилі є її просторовим періодом і дорівнює відстані, на яку поширюється хвиля за ча одного періоду коливань поля;

$$\lambda = vT = \frac{v}{\nu} = \frac{2\pi v}{\omega}. \qquad (1.9)$$

У випадку плоскої хвилі (Лекція 4.7, п.7.2.1) напрям поширення та амплітудний вектор $\vec{E}_0$ однакові в усіх точках. Тому, спрямувавши координатну вісь у напрямку поширення, рівняння хвилі можна записати скалярно як

$$E(x,t) = E_0 \cos(\omega t - kx + \varphi_0), \qquad (1.10)$$

де E_0 – амплітуда, а $E(x,t)$ – проекція миттєвої напруженості електричного поля $\vec{E}$ на напрям амплітудного вектора $\vec{E}_0$.

Хвильові поверхні та промені. Фазова швидкість. У загальному випадку розподіл поля та характер поширення хвилі в просторі відображають *хвильові поверхні* – сукупність точок, в яких миттєва фаза хвилі $\varphi(x,t) = const$ (Лекція 4.7 п. 7.2.4). Для плоскої хвилі (1.10) вони мають форму площин перпендикулярних до напрямку поширення хвилі як показано штриховими лініями на рис. 1.1а. Для сферичних та циліндричних хвиль, які описуються рівняннями

$$E(r,t) = \frac{a}{r}\cos(\omega t - kr)$$

і

$$E(r,t) = \frac{a}{\sqrt{r}}\cos(\omega t - kr),$$

(тут *r* – відстань від точки, яка знаходиться на одиничній відстані від джерела, *a* – амплітудний множник) хвильові поверхні мають вигляд концентричних сфер або коаксіальних циліндрів (показані штриховими лініями на рис. 1.1б).

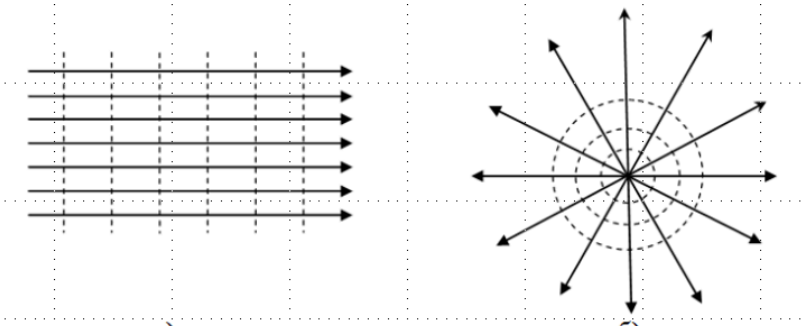


Рис. 1.1.

Зображуючи на рисунку сімейство хвильових поверхонь для заданого моменту часу, можна отримати наочне уявлення про просторовий розподіл фаз, отже й величини поля хвилі в цей момент. У той же час, зображуючи положення хвильової поверхні в різні моменти, можна скласти уявлення про характер поширення хвилі часі. Але зазвичай для цього використовують ***світлові промені*** – лінії, дотичні до яких у кожній точці збігаються з напрямом поширення хвилі в цій точці (суцільні лінії на рис. 1.1). Із наведених прикладів видно, що в кожній точці хвиля поширюється по нормалі

до хвильової поверхні, що проходить через цю точку. Тобто промені та хвильові поверхні розглянутих простих хвиль є взаємно ортогональними. В ізотропних середовищах така взаємні орієнтація зберігається і для більш складних хвиль.

Хвильові поверхні з часом перемішуються в просторі із швидкістю

$$v = \frac{\omega}{k}, \tag{1.11}$$

яка називається ***фазовою швидкістю***.

Порівнюючи вирази (1.11) та (1.8) бачимо, що для ідеальної монохроматичної хвилі фазова швидкість збігається зі швидкістю самої хвилі, тобто швидкість переміщення фази співпадає із швидкістю поширення самих коливань та швидкстю перенесення енергії. Але така хвиля є ідеалізацією. Реальна хвиля завжди характеризуться не однією частотою, а відповідним набором састот (спектром), для кожної з яких є своя фазова швидкістью Тому поширення коливань в реальних хвиялх визначається іншою, так званою, ***груповою швидкістю*** (детальніше див. Лекція 5.6).

Інтенсивність світла. Зорові відчуття людини визначаються світловою енергією, що потрапляє в око за 1 с безпосередньо від джерела, чи після відбивання від освітленої поверхні. Ця енергія залежить від ***інтенсивності*** світлової хвилі I, а інтенсивність світла в кожній точці дорівнює енергії, що переноситься за одиницю часу через одиницю площі поперечного перерізу світлового пучка. Інтенсивність світла можна виразити через показник заломлення та амплітуду світлового вектора:

$$I = \frac{1}{2}\sqrt{\frac{\varepsilon_0}{\mu_0}}\, n E_0^2 \quad \Rightarrow \quad I\, n E_0^2. \tag{1.12}$$

В однорідному середовищі (n = const) розподіл інтенсивностей у просторі визначається тільки амплітудою хвилі. Тому, коли не потрібно знати абсолютні величини, замість інтенсивності оперують квадратом амплітуди світлового вектора:

$$I \sim E_0^2. \tag{1.12a}$$

Зауважимо також, що завдяки співвідношенню (1.6) між полями у світловій хвилі інтенсивність можна аналогічно визначити й через амплітуду магнітного поля.

2. СВІТЛОВІ ХВИЛІ НА МЕЖІ ПОДІЛУ СЕРЕДОВИЩ. ВІДБИВАННЯ ТА ЗАЛОМЛЕННЯ

Визначальною властивістю хвиль є відбивання та заломлення – здатність ділитися на дві хвилі на межі двох різних середовищ. Одна з хвиль (відбита) не проходить, а інша (заломлена) проходить крізь межу поділу. Напрямки поширення відбитої та заломленої хвиль визначаються простими геометричним законами, що не залежать від фізичної природи хвилі. Оптиці ці закони складають основу ***геометричної оптики***, в якій поширення світла досліджується на основі уявлення про падаючі, відбиті та заломлені промені – лінії, що вказують напрям поширення відповідних хвиль в кожній точці простору і спрямовані по нормалі до відповідної хвильової поверхні.

Напрямок падаючого, відбитого чи заломленого променя задається кутом між ним та нормаллю до межі поділу середовищ у точці падіння. Площина, в якій лежать указана нормаль і падаючий промінь називають ***площиною падіння***. В ізотропних середовищах відбитий та заломлений промені лежать у тій самій площині, що випливає з міркувань симетрії. Таким чином,

в ізотропних середовищах падаючий, відбитий та заломлений промені лежать в одній площині – площині падіння.

Це твердження складає одне з основних положень геометричної оптики.

Закони відбивання і заломлення. Ці закони є основними законами геометричної оптики. Їх можна вивести із загальних законів електромагнітної теорії Максвелла. Але до них можна прийти простіше на основі відомого емпіричного принципу Гюйгенса, який спрощено можна сформулювати так:

точки, до яких на дану мить дійшла хвиля (точки фронту хвилі), можна вважати джерелами сферичних вторинних хвиль; положення фронту хвилі, що розглядається, в наступний момент часу визначається обвідною елементарних хвильових поверхонь всіх вторинних хвиль у цей момент.

У такий спосіб за відомими положенням фронту хвилі в даний момент часу можна побудувати його положення в наступні моменти і, відтак, прослідкувати поширення даної хвилі. Зокрема, так можна показати відомий факт прямолінійного поширення світла в однорідному середовищі, який трактується як один з основних законів геометричної оптики. Але слід зауважити, що принцип Гюйгенса не має строго фізичного обґрунтування і є лише емпіричним правилом побудови хвильових фронтів.

Виконаємо побудову Гюйгенса для світлових хвиль на плоскій межі поділу двох середовищ з показниками заломлення n_1 і \n_2\)і відповідним швидкостям поширення v_1 і v_2. Нехай на межу поділу під кутом ϑ_1 падає паралельний світловий пучок (плоска хвиля), обмежений променями 1 і 2 (рис. 1.2).

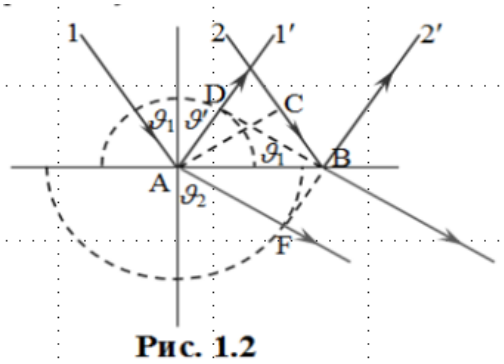


Рис. 1.2

Позначимо кутами ϑ і ϑ' напрямки відбитих та заломлених променів і відрізком AC – положення фронту падаючої хвилі на момент приходу променя 1 в точку A. Промінь 2 потрапляє в точку B пізніше на час проходження ним відстані BC. Таку саме відстань AD = BC за цей час проходить відбитий промінь $1'$. Тому фронт відбитої в точці A вторинної хвилі на цей момент зобразиться півсферою з радіусом AD, а дотична до неї площина показана відрізком BD укаже положення результуючого фронту хвилі. *Примітка. Ця площина є дотичною й до безлічі подібних сфер меншого радіуса, що зображують хвильові поверхні вторинних хвиль відбитих від інших точок ділянки A-B межі поділу середовищ, на яку падає світловий пучок.* Тож відрізок $\mathrm{AD} \perp \mathrm{BD}$ визначає напрям відбитого променя $1'$. Оскільки AD = BC, то DABC = DABD. Звідси, як можна зрозуміти з рис. 1.2, випливає закон відбивання світла:

$$\vartheta' = \vartheta. \tag{1.13}$$

А саме:

кут відбивання дорівнює кутові падіння.

Аналогічною побудовою можна встановити і напрям поширення заломленої хвилі. Падаючий промінь 2 потрапляє на поверхню поділу середовищ пізніше, ніж промінь 1 на час $\tau = \mathrm{BC}/v_1$. За цей час вторинна хвиля від точки A поширюється в другому середовищі на відстань AF так, що

$$\mathrm{AF} = v_2\tau = \mathrm{BC} \cdot \frac{v_2}{v_1} \quad \Rightarrow \quad \frac{\mathrm{BC}}{\mathrm{AF} = \frac{v_1}{v_2}}. \tag{1.14}$$

Площина (показані відрізком BF), що проходить через точку B і є дотичною до сфери радіуса AF, визначає фронт заломленої хвилі, а перпендикуляр AF до нього – напрям заломлених променів. Урахувавши, що в прямокутних трикутниках DABC і DABF катети BC = ABsin ϑ_1 і AF = AB$sin\vartheta_2$ і що $v = c/n$, із співвідношення (1.14) отримаємо закон заломлення:

$$\frac{\sin\vartheta_1}{\sin\vartheta_2} = \frac{n_2}{n_1} \quad \text{або} \quad n_1 \sin\vartheta_1 = n_2 \sin\vartheta_2 \tag{1.15}$$

при переході променя через межу поділу двох середовищ відношення синусів кутів падіння та заломлення дорівнює оберненому відношенню показників заломлення цих середовищ.

Отже, як саме явище заломлення, так і зв'язок між напрямками поширення світлових хвиль у двох середовищах зумовлені відмінністю оптичних густин, тобто швидкостей світла в даних середовищах.

Граничний кут. Із співвідношень (1.15) випливає, що коли $\vartheta_1 = 0$, то й $\vartheta_2 = 0$, тобто при нормальному падінні променя на межу поділу ***ізотропних*** середовищ заломлення в буквальному розумінні немає. Істотно також, що при косому падінні напрям відхилення заломленого променя залежить від співвідношення показників заломлення середовищ.

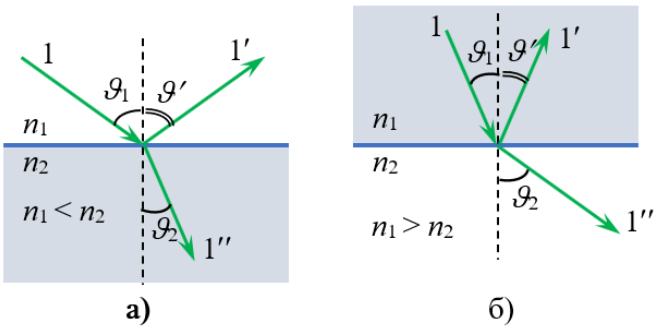


Рис. 1.3

Коли світло переходить в оптично більш густу речовину ($n_1 > n_2$), то згідно з (1.19) $\vartheta_1 < \vartheta_2$ і заломлений промінь відхиляється до нормалі (рис. 1.3а). Якщо ж світло падає на межу поділу із більш густого середовища ($n_1 > n_2$), то $\vartheta_2 > \vartheta_1$ і заломлений промінь відхиляється до поверхні поділу середовищ (рис. 1.3.б). У цьому випадку при певні величині кута падіння $\vartheta_1 = \vartheta_{\text{гр}}$ кут заломлення набуває максимального можливого значення $\vartheta_2 = 90^\circ$. Величина $\vartheta_{\text{гр}}$ називається ***граничним кутом*** (інакше – критичним кутом) для пари середовищ і визначається з (1.15) як

$$\sin\vartheta_{\text{гр}} = \frac{n_2}{n_1} \quad \Rightarrow \quad \vartheta_{\text{гр}} = \arcsin\left(\frac{n_2}{n_1}\right), \quad (n_1 > n_2). \tag{1.16}$$

Як свідчить теорія та експеримент, при збільшенні кута падіння, за будь-яких умов, інтенсивність відбитого променя весь час збільшується, а заломленого – зменшується. При чому, у випадку $n_1 > n_2$, коли кут падіння наближається до значення $\vartheta_1 = \vartheta_{\text{гр}}$, інтенсивність заломленого променя зменшується до нуля. Тому при кутах падіння $\vartheta \geq \vartheta_{\text{гр}}$ енергія падаючого променя повністю відбивається від межі поділу двох прозорих речовин як від ідеального дзеркала. Це явище носить назву ***повного внутрішнього відбивання***. Воно спостерігається у природі і широко використовується в техніці. Достатньо згадати сліпучий виблиск крапель роси на сонці або широке застосування оптичних кабелів у системах телекомунікації.

Нагадаємо ще раз, що повне внутрішнє відбивання при падінні світла із середовища з показником заломлення n_1 намежу поділу із середовищем n_2 спостерігається при одночасному виконанні двох умов:

$$n_1 > n_2 \quad \text{і} \quad \vartheta_1 \geq \vartheta_{\text{гр}}. \tag{1.17}$$

3. АМПЛІТУДИ І ФАЗИ ВІДБИТОЇ ТА ЗАЛОМЛЕНОЇ ХВИЛЬ

При проходженні світлової хвилі крізь межу поділу прозорих середовищ її енергія розподіляється між відбитою та заломленою хвилями. Тому поверхня поділу двох різних середовищ характеризується відповідними ***коефіцієнтами відбивання та пропускання*** (прозорістю).

Указані коефіцієнти можна визначити теоретично із законів електродинаміки за допомогою умов на межі (див. Лекція 3.4, п. 4.4). Ці умови встановлюють зв'язок між векторами електричного $\vec{E}_1$, $\vec{E}_2$ і магнітного $\vec{H}_1$, $\vec{H}_2$ полів на межі поділу двох ізотропних речовин і мають вигляд:

$$\begin{aligned} \varepsilon_1 E_{1n} &= \varepsilon_2 E_{2n}, \quad E_{1\tau} = E_{1\tau}; \\ \mu_1 H_{1n} &= \mu_2 H_{2n}, \quad H_{1\tau} = H_{1\tau}, \end{aligned} \tag{1.18}$$

де індекси n і τ відносять, відповідно, до нормальних та тангенціальних (дотичних до межі поділу) проекцій векторів; ε і μ – діелектричні та магнітні проникності середовищ. Ці співвідношення показують, що на межі поділу змінюються тільки нормальні складові векторів, тож і їхні напрямки. При переході з одного середовища в інше змінюється і напрям поширення хвилі, тобто відбувається її заломлення.

З умов (1.18) можна визначити відносні інтенсивності та потоки енергії відбитої та заломленої хвиль, відтак і коефіцієнти відбивання та пропускання світла на поверхні поділу середовищ при будь-яких кутах падіння променів. Але обмежимося тільки випадком, коли паралельний пучок промені (плоска хвиля) падає на межі поділу по нормалі. Позаяк світлова хвиля є поперечною, то в цьому випадку нормальні складові полів відсутні і модулі тангенціальних проєкцій дорівнюють модулям самих векторів. Тоді умова (1.18) набуде вигляду:

$$E_1 = E_2, \quad H_1 = H_2,$$

де задля зручності опущені індекси тангенціальних проекцій.

Розглянемо ці умови детальніше. Електромагнітне поле в першому середовищі є суперпозицією полів падаючої ($\vec{E}$, $\vec{H}$) і відбитої ($\vec{E}'$, $\vec{H}'$) хвиль:

$$\begin{aligned} \vec{E}_1 &= \vec{E} + \vec{E}', \\ \vec{H}_1 &= \vec{H} + \vec{H}', \end{aligned} \tag{1.18а}$$

а в другому – то є поле заломленої хвилі ($\vec{E}''$, $\vec{H}''$):

$$\begin{aligned} \vec{E}_2 &= \vec{E}'', \\ \vec{H}_2 &= \vec{H}''. \end{aligned} \tag{1.18б}$$

При записі умов (1.18а), (1.18б) у скалярному вигляді треба врахувати, що в електромагнітній хвилі напрямки векторів жорстко узгоджені між собою. В саме, в послідовності $\vec{E}, \vec{H}, \vec{k}$ та при циклічних перестановках ці вектори завжди утворюють праву трійку, тобто при повороті правого гвинта від першого вектора до другого гвинт буде вкручуватися в напрямку третього вектора. Тому зміна напрямку поширення хвилі (напрямку вектора $\vec{k}$) при відбиванні автоматично тягне за собою відповідні зміну напрямку одного з векторів полів.

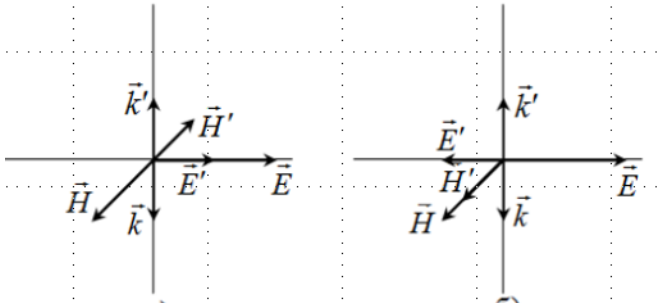


Рис. 1.4

Справді, уявімо, що при відбиванні світловий вектор $\vec{E}'$ не змінив напрям як показано на рис. 1.4а. Тоді вектор $\vec{H}'$ обов'язково змінить напрям на протилежний. Це означає зміну фази коливань магнітного поля в точці відбивання на протилежну. Так само, якщо не змінюється напрям магнітного вектора, то обов'язково «перекидається» електричний вектор (рис. 1.4б). Тому, якщо прийняти для визначеності ситуацію рис. 1.6б, то умови (1.18а) в скалярному вигляді запишуться як

$$\begin{aligned} E + E' &= E'', \\ H - H' &= H''. \end{aligned} \qquad (1.18б)$$

Урахувавши, що в прозорих для світла речовинах $\mu = 1$ і $\sqrt{\varepsilon} = n$, і перейшовши від магнітного поля до електричного на основі співвідношення (1.7), отримаємо систему рівнянь:

$$\begin{cases} E + E' = E'' \\ n_1 E - n_1 E' = n_2 E'', \end{cases}$$

з якої можна визначити тангенціальні складові світлового вектора відбитої та заломленої хвиль:

$$E' = \frac{n_1 - n_2}{n_1 + n_2} E; \quad E'' = \frac{2n_1}{n_1 + n_2}. \qquad (1.19)$$

Таким чином, напруженість електричного та магнітного полів у відбитій та заломленій хвилях визначаються тільки їхніми показниками заломлення (при косому падінні променів вони залежать ще й від кута падіння). Зокрема, для речовин з близькими показниками заломлення $E' \approx 0$, $E'' \approx E$ і межа поділу є практично невидимою. З формул (1.19) також випливає, що знаки E'' і E завжди збігаються. Відповідно, в падаючій заломленій хвилях у точці падіння напрямки полів $\vec{E}''$, $\vec{E}$, і фази коливань однакові. Що ж до відбитої хвилі, то напрям поля $\vec{E}'$ залежить від співвідношення показників заломлення: при $n_1 > n_2$ він збігається з напрямком $\vec{E}$, а при $n_1 < n_2$ – є протилежним. Отже:

> в точці відбивання світла від оптично менш густого середовища ($n_2 > n_1$) фаза світлового вектора $\vec{E}$ не змінюється, в той час як фаза магнітного вектора $\vec{H}$ змінюється на π. Натомість при відбиванні від більш густого середовища ($n_1 < n_2$) не змінюється фаза магнітного вектора $\vec{H}$, а фаза світлового вектора $\vec{E}$ змінюється на π.

При поширенні хвилі фаза змінюється на π на шляху $\lambda/2$. Тому описане «перекидання» фази називають втратою півхвилі на межі поділу двох середовищ. Цей ефект істотно впливає на формування поля світлової хвилі в зоні відбивання і має враховуватися у відповідних задачах оптики.

Вирази (1.19) дозволяють визначити розподіл енергії у відбитій та заломленій хвилях. Цей розподіл визначається коефіцієнтами відбивання ρ та пропускання r («прозорістю»), які дорівнюють відношенню потоків енергії у відбитій Φ' та заломленій Φ'' хвилях до потоку енергії Φ у падаючій хвилі:

$$\rho = \frac{\Phi'}{\Phi}, \quad r = \frac{\Phi''}{\Phi}.$$

У паралельному пучку променів з площею поперечного перерізу S потік енергії $\Phi = IS$, де I – інтенсивність світла. При нормальному падінні площі поперечного перерізу падаючого, відбитого та заломленого променів однакова, тому

$$\rho = \frac{I'}{I}, \quad r = \frac{I''}{I}.$$

Відтак, урахувавши вирази (1.16) та (1.19), отримаємо

$$\rho = \frac{(n_1 - n_2)^2}{(n_1 + n_2)^2} \quad r = \frac{4n_1 n_2}{(n_1 + n_2)^2}. \tag{1.20}$$

Ці формули дають $\rho + r = 1$, як того й вимагає закон збереження енергії.

Слід зауважити, що при косому падінні у формулах (1.20) вираз для ρ зберігається, а для r – ні, бо при косому падінні площі поперечного перерізу падаючого й заломленого пучків не однакові.

Показники заломлення більшості прозорих речовин $n \leq 2$, тому величина ρ є малою, а r – близькою до одиниці. Приміром, при падінні світла з повітря (n_1 = 1) на поверхню води (n_2 = 1,33) або скла (n_2 = 1,5) величина ρ, згідно з (1.20), складає 2 % та 4 %. Проте в оптичних приладах, наприклад, складних об'єктивах, де багато світлові пучки на своєму шляху стрічають багато поверхонь поділу, сумарні втрати на відбивання можуть скласти значну частину падаючої світлової енергії. Тому в оптичному приладобудуванні використовують спеціальні технології для зменшення таких утрат.

5.2. ІНТЕРФЕРЕНЦІЯ СВІТЛА

Нижче розглянуті такі питання

1. Двопроменева інтерференція
2. Інтерференційна картина
3. Проблеми когерентності
4. Контрольні запитання

1. ДВОПРОМЕНЕВА ІНТЕРФЕРЕНЦІЯ

Інтерференція. Інтерференція – це явище, що спостерігається при накладання декількох хвиль. Сутність явища розглянемо на прикладі ***двопроменевої інтерференції***, що за певних умов спостерігається при накладанні двох світлових хвиль.

Нехай в якусь точку приходять дві хвилі із світловими векторами $\vec{E}_1(\vec{r},t)$ і $\vec{E}_2(\vec{r},t)$. За принципом суперпозиції результуюча напруженість електричного поля у цій «*точці спостереження*» $\vec{E}=\vec{E}_1+\vec{E}_2$, тож

$$E^2=(\vec{E}_1+\vec{E}_2)^2=E_1^2+E_2^2+2\vec{E}_1\vec{E}_2.$$

Інтенсивність хвилі I є прямо пропорційною середній величині квадрата напруженості електричного поля $I\sim E^2$, тому провівши усереднення

$$\langle E^2\rangle=\langle E_1^2\rangle+\langle E_2^2\rangle+2\langle\vec{E}_1\vec{E}_2\rangle,$$

для результуючої інтенсивності маємо

$$I+I_1+I_2+2\langle\vec{E}_1\vec{E}_2\rangle.$$

Якщо коливання світлових векторів у точці накладання відбуваються у взаємно перпендикулярних напрямках, то $\vec{E}_1\vec{E}_2=0$, і результуюча інтенсивність $I=I_1+I_2$. Але при інших випадках може виявитися, що середнє значення добутку функцій $\langle\vec{E}_1(\vec{r},t)\cdot\vec{E}_2(\vec{r},t)\rangle\neq(0)$. Тоді

$$I\neq I_1+I_2,$$

у чому й полягає явище інтерференції, а саме,

> ***інтерференцією називають таке накладання хвиль, при якому результуюча інтенсивність не дорівнює сумі інтенсивностей, що створюються в точці накладання кожною хвилею окремо.***

На перший погляд сказане суперечить закону збереження енергії, але це, звісно, не так. Інтенсивність є локальною величиною, що визначає потік енергії лише крізь елементарну площадку в околі даної точки. Якщо ж підрахувати потік крізь скінченну поверхню, то він завжди виявляється рівним сумі потоків, які створюються кожним джерелами окремо. Тож при інтерференції світлова енергія не змінюється, а лише перерозподіляється у просторі так, що в одних місцях результуюча інтенсиваність виявляється більшою, а в інших – меншою за суму інтенсивностей хвиль, які накладаються.

Інтерференційна фомула. Детальний аналіз явища інтерференції почнемо з аналізу накладання ідеальних монохроматичних світлових хвиль від двох точкових (або лінійних) джерел S_1 та S_2 (рис. 2.1), які мають однакову частоту ω і напям коливань світлового вектора. Для загальності будемо також уважати, що хвилі від джерел до точки накладання поширюються в середовищах з різним показниками заломлення, тож і швидкостями швидкостями поширення. За таких умов світлові коливання, що збуджуються кожним джерелом у вибраній точці спостереження P, визначаються (див. рівняння (1.10) з лекції 5.1) як:

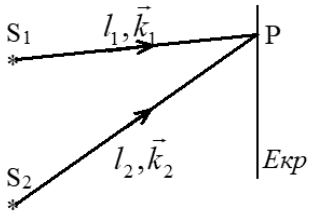


Рис. 2.1

$$E_1 = E_{01}\cos(\omega t - k_1 l_1 + \varphi_{01}) = E_{01}\cos(\omega t - \alpha_1),$$
$$E_1 = E_{01}\cos(\omega t - k_1 l_1 + \varphi_{01}) = E_{01}\cos(\omega t - \alpha_1), \qquad (2.1)$$

де l_1, l_2 – відстані від джерел до точки спостереження, k_1, k_2 – хвильові числа, що визначаються частотою ω, швидкостями поширення v_1, v_2 та довжинами хвиль λ_1, λ_2 як

$$k_1 = \frac{\omega}{v_1} = \frac{2\pi}{\lambda_1} \quad \text{і} \quad k_2 = \frac{\omega}{v_2} = \frac{2\pi}{\lambda_2} \qquad (2.2)$$

Величини φ_{01}, φ_{02} – це початкові фази коливань у джерелах (при відстані $l = 0$, а $\alpha_1 = \varphi_{01} - k_1 l_1$, $\alpha_2 = \varphi_{02} - k_2 l_2$ – початкові фази коливань у точці спостереження.

Результуючу напруженість поля світлових хвиль $\vec{E} = \vec{E}_1 + \vec{E}_2$ у точці спостереження найпростіше визначити за допомогою векторної діаграми, зображуючи кожне коливання відповідним вектором, що обертається з кутовою швидкістю ω, а довжина якого дорівнює модулю відповідного вектра. Миттеве положення вектрів показано на рис. 2.2. За теоремою коснусів визначаємо модуль вектора результуючої напруженості поля в точці спостереження:

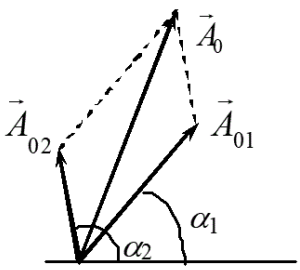


Рис. 2.2

$$E_1^2 = E_{01}^2 + E_{02}^2 + 2E_{01}E_{02}\cos\delta, \qquad (2.3)$$

де $\delta = \alpha_1 - \alpha_2$ – різниця фаз коливань, яка за рівнянням (2.1) складає

$$\delta = (k_2 l_2 - k_1 l_1) - (\varphi_{02} - \varphi_{02}). \qquad (2.4)$$

Відтак, замінивши в (2.3) квадрати амплітуд на інтенсивності, отримаємо:

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta. \qquad (2.5)$$

З виразу (2.5), який називають ***інтерференційною формулою***, видно, що при накладанні двох ідеальних монохроматичних хвиль результуюча інтенсивність $I \neq I_1 + I_2$, тобто – маємо інтерференцію.

Неважко збагнути, що такий результат вийшов тому, що хвилі мають однакову частоту. Через це в кожній точці простору різниця фаз δ падаючих хвиль не залежть від часу, тож вони є жорстко узгодженими. Але в реальних хвилях із різних причин повної узгодженості коливань бути не може і різниця фаз δ, отже і $\cos\delta$, в різні моменти часу мають різну величину. Тому інтенсивність, яку ми сприймаємо візуально або реєструємо приладами, визначається середнім значенням $\langle\cos\delta\rangle$ за час спостереження.

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\langle\cos\delta\rangle. \qquad (2.6)$$

Отже, для можливості реального спостереження інтерференції ведичина $\langle\cos\delta\rangle$ має лишатися незмінною впродовж проміжку часу, який є достатнім для реєстрації ефекту. Інакше кажучи, коливання, що збуджуються хвилями в точці накладання, мають бути в достатній мірі узгодженими. Узгодженість хвильових та коливальних процесів відображається поняттям ***когерентності***:

> ***хвилі (і коливання), різниця фаз яких у даній точці з часом не змінюється або змінюється достатньо повільно, називають когерентними.***

Таким чином,

інтерференцію можна спостерігати лише при накладанні когерентних хвиль.

При цьому ідеальні монохроматичні хвилі однакової частоти, для яких $\cos\delta = const$, є ***повністю когерентними***. Що до реальних хвиль, в яких $\cos\delta \neq const$ але $\langle\cos\delta\rangle \neq 0$, то їх трактують як ***частково когерентні***.

Із виразів (2.4), (2.5) та рис. 2.1 зрозуміло, що при переміщенні точки спостереження P у певному напрямі змінюються: різниця фаз δ – монотонно, а $\cos\delta$ і результуюча інтенсивність I – періодично. Це означає, що при накладанні когерентних хвиль утворюється система упорядковано розташованих областей підвищеної та пониженої інтенсивності світла. Тому на екрані, встановленому в області накладання (зоні інтерференції), спостерігається «***інтерференційна картина***» у вигляді сукупності світлих і темних смуг – інтерференційних ***максимумів*** та ***мінімумів***. Отже, можна дати таке коротке формулювання:

інтерференцією називається утворення максимумів та мінімумів інтенсивності при накладанні когерентних хвиль.

2. ІНТЕРФЕРЕНЦІЙНА КАРТИНА

Різниця ходу. Із формули (2.5) випливає, що інтерференційні максимуми утворюються при $cos\delta = 1$, а мінімуми – при $\cos\delta = -1$. Отже, різниця фаз когерентних хвиль у точках максимумів і мінімумів, відповідно, має задовольняти умови:

$$\begin{matrix} \delta = \pm m \cdot 2\pi, \\ \delta = \pm(2m+1)\pi; \end{matrix} \quad m = 0, 1, 2 \ldots \tag{2.7}$$

Ціле число називається ***порядком інтерференційної смуги***, або просто ***порядком інтерференції***.

Для когерентних хвиль величина δ визначається виразом (2.4). При цьому різниця початкових фаз не впливає на характер розподілу інтенсивностей, до того ж на практиці завжди $\varphi_{01} = \varphi_{02}$. Тому, взявши до уваги вирази (2.2), запишемо:

$$\delta = k_2 l_2 - k_1 l_1 = 2\pi \left(\frac{l_2}{\lambda_2} - \frac{l_1}{\lambda_1} \right),$$

або

$$\delta = \frac{2\pi}{\lambda_0}(n_2 l_2 - n_1 l_1),$$

де l_1, l_2 – відстані від джерел до точки спостереження, λ_0 – довжина світлової хвилі у вакуумі, n_1, n_2 – показники заломлення середовищ, в яких поширюються промені.

Величина $L = nl$ називається ***оптичною довжиною шляху*** або ***оптичним ходом*** променя від джерела до даної точки. Якщо на шляху променя трапляється декілька різних середовищ, то $L = \sum n_i l_i$, а у випадку неоднорідного середовища $L = \int n\mathrm{d}l$. Відповідно, різниця оптичних довжин шляху двох променів

$$\Delta = L_2 - L_1 = n_2 l_2 - n_1 l_1 \tag{2.8}$$

називається ***оптичною різницею ходу*** променів. Якщо обидва промені поширюються в одному середовищі, то $n_1 = n_2 = n$, і

$$\Delta = n(l_2 - l_1) = n\Delta_0, \tag{2.8а}$$

де величина $\Delta_0 = l_2 - l_1$ – ***геометрична різниця ходу***.

Відтак для різниці фаз у загальному випадку маємо:

$$\delta = \frac{2\pi}{\lambda_0}\Delta, \tag{2.9}$$

3. Інтерференційні максимуми та мінімуми

***Умови максимумів та мінімумі**в*. Підставивши вираз (2.9) в (2.7), отримаємо *умови інтерференційних максимумів і мінімумів*:

для максимумів: $$\Delta = \pm 2m\frac{\lambda_0}{2} = \pm m\lambda, \qquad (2.10)$$

для мінімумів: $$\Delta = \pm(2m+1)\frac{\lambda_0}{2} = \pm\left(m+\frac{1}{2}\right)\lambda_0. \qquad (2.10а)$$

В обох випадках $m = 0, 1, 2 \ldots$ – порядок інтерференції.

Згідно з цими умовами

інтерференційні максимуми утворюються в точках, в яких оптична різниця ходу променів дорівнює парній кількості півхвиль або цілій кількості довжин світлової хвилі у вакуумі;

інтерференційні мінімуми утворюються в точках, де оптична різниця ходу променів дорівнює непарній кількості півхвиль або напівцілій кількості довжин світлової хвилі у вакуумі.

Якщо когерентні промені поширюються в однорідному середовищі, то умови (2.10) і (2.10а) можна записати у вигляді через геометричну різницю ходу Δ_0 і довжину світлової хвилі в середовищі λ:

для максимумів: $$\Delta_0 = \pm 2m\frac{\lambda}{2} = \pm m\lambda, \qquad (2.11)$$

для мінімумів: $$\Delta_0 = \pm(2m+1)\frac{\lambda_0}{2} = \pm\left(m+\frac{1}{2}\right)\lambda. \qquad (2.11а)$$

де $\lambda = \lambda_0/n$ – довжина хвилі в речовині, і $\Delta_0 = l_2 - l_1$ – *геометрична різниця ходу* променів.

4. ІНТЕНСИВНІСТЬ СВІТЛА ТА ШИРИНА ІНТЕРФЕРЕНЦІЙНИХ СМУГ

Інтенсивність та ширина інтерференційних смуг. Розглянемо тепер величину інтенсивностей у максимумах I_{max} та мінімумах I_{min} і їхнє положення на екрані при інтерференції від двох точкових когерентних джерел. Згідно з виразом (2.5) та умовами (2.7)

$$I_{max} = (\sqrt{I_1} + \sqrt{I_2})^2, \\ I_{min} = (\sqrt{I_1} - \sqrt{I_2})^2. \tag{2.12}$$

Аби інтерференційні смуги на екрані були чіткими (контрастними), інтенсивність мінімумів має бути якнайменшою, в ідеалі – $I_{min} - 0$. В такому випадку $I_2 = I_1$ і $I_{max} = 4I_1$. Тому на практиці завжди намагаються отримувати когерентні промені з максимально близькими інтенсивностями. При цьому формула (2.5) набуває вигляду:

$$I = 2I_1(1 + \cos\delta) = I_0 \cos^2 \frac{\delta}{2}, \text{ де } I_0 = 4I_1. \tag{2.13}$$

Визначимо із цього виразу розподіл інтенсивностей на екрані $I(x)$ при двопроменевій інтерференції у повітрі (n =1) від ідеальних точкових когерентних джерел S_1 і S_2 з однаковими інтенсивностями $I_1 = I_2$. Для цього спочатку знайдемо різницю ходу $\Delta(x)$, а потім за допомогою формули (2.9а) – різницю фаз $\delta(x)$ променів, які приходять у довільну точку Р, в залежності від її координати x, рис. 2.3.

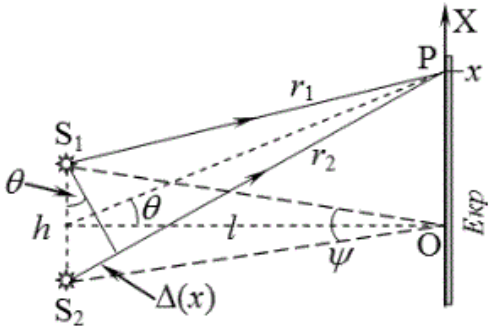


Рис. 2.3

Задля спрощення викладок заздалегідь урахуємо, що інтерференційні смуги на екрані можна візуально спостерігати тільки при малих відстанях між джерелами ($h \ll 1$) і в невеликій області біля центра картини О, тобто при $x \ll l$. У такому разі показані на рис. 2.3 кути θ дуже малі, практично однакові та мають величину $\theta = x/l$. Тож, замінюючи відповідні тригонометричні функції самими кутами, різницю ходу Δ і різницю фаз δ, можна з великою точністю виразити, як:

$$\Delta(x) = h \cdot \theta = h\frac{x}{l} \quad \Rightarrow \quad \delta(x) = \frac{2\pi h}{\lambda l}x\backslash). \tag{2.14}$$

Підставивши цей вираз $\delta(x)$ у формулу (2.13), для інтенсивності отримаємо:

$$I(x) = I_0 \cos^2\left(\frac{\pi h}{\lambda l}\right). \tag{2.15}$$

Графік залежності $I(x)$ показаний на рис.2.4.

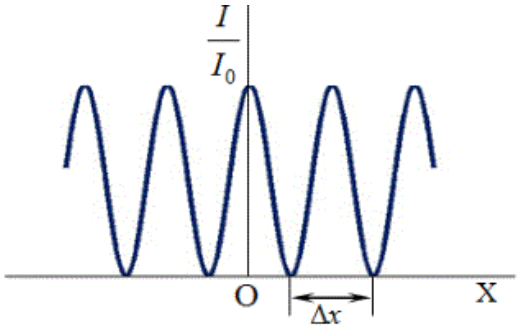


Рис. 2.4

Інтерференційна картина складається системи світлих смуг з інтенсивністю в максимумах $I_0 = 4I_1$ та темних смуг із нульовою інтенсивністю в мінімумах. Смуги розміщуються симетрично по обидва боки від центральної точки О, в якій знаходиться максимум порядку $m = 0$. Він називається центральним максимумом, а відповідна смуга – центральною інтерференційною смугою. Підставивши вираз $\Delta(x)$ з (2.14) в умови (2.11) та (2.11а), отримаємо вирази для координат максимумів та мінімумів при інтерференції:

$$\begin{aligned} x_{m\max} &= \pm m\frac{l\lambda}{h}, \\ x_{m\min} &= \pm\left(m+\tfrac{1}{2}\right)\frac{l\lambda}{h}. \end{aligned} \qquad m = 0, 1, 2\ldots \tag{2.16}$$

З цих формул випливає, що відстань $\Delta x = |x_{m+1} - x_m|$ між сусідніми максимумами чи мінімумами не залежить від порядку інтерференції, тобто смуги розташовані ***еквідистантно***, тобто на однаковій відстані

$$\Delta x = \frac{l\lambda}{h}\backslash), \tag{2.17}$$

або

$$\Delta x = \frac{\lambda}{\Psi}, \tag{2.17}$$

де $\Psi = h/l$ – кутова відстань між когерентними джерелами, тобто кут зору, під яким їх видно з центра інтерференційної картини (див. рис. 2.3).

Між сусідніми мінімумами розташовується інтерференційний максимум, тобто світла смуга. Тому величина називається ***шириною інтерференційної смуги***.

Вираз (2.17) дозволяє зрозуміти, чому інтерференційні смуги можна спостерігати тільки при малих відстанях між когерентними джерелами. Причина полягає в обмеженій роздільній здатності ока (гостроті зору), що визначається найменшим кутом зору Ψ_0, при якому дві близькі риски ще бачаться роздільно (не зливаються). Для нормального ока $\Psi_0 \approx 1' \approx 3 \cdot 10^{-4}$ рад, тож при відстані до джерел $l = 1$ м інтерференцію можна візуально спостерігати, коли відстань між когерентними джерелами $h \leq 0,3$ мм.

5. ПРОБЛЕМИ КОГЕРЕНТНОСТІ

Некогерентність природнього світла. Дослід свідчить, що при накладанні світлових пучків від двох звичайних джерел світла (маються на увазі всі природні та штучні джерела, крім оптичних квантових генераторів - лазерів) інтерференція ніколи не спостерігається. Це означає, що хвилі від двох незалежних джерел світла є некогерентними. Причина цього лежить у самому механізмі випромінювання. Світло випромінюється атомами тіла не у вигляді неперервної хвилі, а окремими «порціями» (причини цього пояснюються в квантовій механіці), які у хвильовій теорії називаються ***цугами***.

Кожен цуг являє собою «відрізок хвилі» тривалістю $\tau \sim 10^{-8}$ с і протяжністю $l \sim 3$ м, який має задані початкову фазу та напрям коливань світлового вектора, але характеризується не заданою частотою ω, а спектром – неперервним набором частот у деякому інтервалі $\Delta\omega$ в околі заданого значення ω. Тож хвильовий цуг не є строго монохроматичним. Тому на практиці монохроматичним називають випромінювання не з однією частотою, а з вузькою спектральною шириною $\Delta\omega \ll \omega$. Іншою особливістю цугів є те, що їхні початкові фази, так само, як і напрямки коливань світлового вектора, ніяк не узгоджені: у кожного наступного цугу вони можуть як завгодно відрізнятися від попереднього. Тому в світловому промені від реального джерела початкова фаза та напрям коливань дуже швидко і хаотично змінюються. Відповідно так само швидко і невпорядковано змінюється різниця фаз і взаємна орієнтація напрямків коливань променів, які приходять у дану точку від двох незалежних джерел. Тому світлові пучки від незалежних джерел є повністю некогерентними і при накладанні не інтерферують. Проте руйнівний вплив нестабільності початкової фази на інтерференцію можна виключити, використовуючи для отримання когерентних пучків не два незалежних джерела, а одне. (Як це робиться, розглянемо в наступній лекції). Але через принципову немонохроматичність випромінювання та наявність у джерел скінченних лінійних розмірів добитися повної когерентності реальних світлових пучків все одно неможливо, що суттєво обмежує умови спостереження інтерференції світла. Проаналізуємо вплив кожного із указаних двох факторів окремо.

Часова когерентність. Розглянемо спочатку спрощену ситуацію, коли кожне з двох точкових джерел випромінює монохроматичну хвилю з початковою фазою φ_0 і з тільки двома близькими частотами ω та $\omega + \Delta\omega$ та довжинами хвилі λ і $\lambda + \Delta\lambda$. Тоді розподіл інтенсивностей на екрані можна змоделювати як суперпозицію двох інтерференційних картин, одна з яких створюється променями з довжиною хвилі λ а інша – $\lambda + \Delta\lambda$. Ці картини характеризуються різною шириною смуги (формула (2.17)) і поступово «розповзаються», як показано на рис. 2.5а. Як наслідок, результуюча інтенсивність у максимумах буде поступово зменшуватись, а в мінімумах – збільшуватись аж до зникнення смуг (рис. 2.5б). Це станеться тоді, коли максимум якогось порядку m' для довжини хвилі $\lambda + \Delta\lambda$ співпаде з відповідним мінімумом для довжини хвилі λ. Отже, згідно з формулами (2.16):

$$m' \cdot \frac{l(\lambda + \Delta\lambda)}{h} = \left(m' + \frac{1}{2}\right) \cdot \frac{l\lambda}{h} \quad \Rightarrow \quad m' = \frac{\lambda}{2\Delta\lambda}, \tag{2.18}$$

де m' – порядок інтерференції в місці, де зникають смуги.

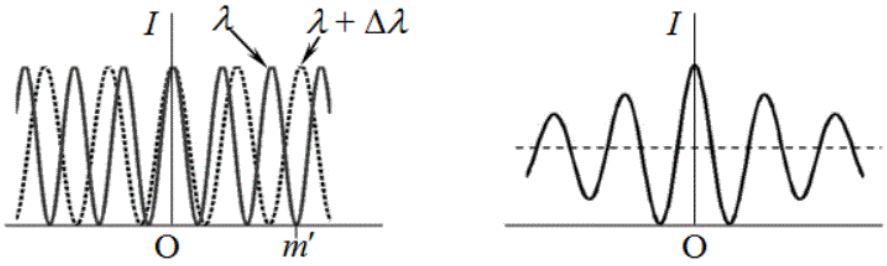


Рис. 2.5

Формально число m' можна трактувати як максимальний порядок інтерференції при накладанні умовних променів, кожен із яких включає дві довжини хвилі. Насправді ж у випромінювання присутні всі можливі довжини хвилі в інтервалі від λ до $\lambda + \Delta\lambda$, тож можна вважати, що весь спектр складається з безлічі пар променів з різницею довжин хвилі $\Delta\lambda/2$ кожна. Тому, підставивши цю величину в (2.18) замість $\Delta\lambda$, отримаємо більш коректний вираз:

$$m_{max} = \frac{\lambda}{\Delta\lambda}, \qquad (2.19)$$

в якому величина

$$\frac{\lambda}{\Delta\lambda} = \frac{\omega}{\Delta\omega}$$

є мірою наближеності світла до строго монохроматичного.

Таким чином, неповна монохроматичність світла обмежує кількість інтерференційних смуг, які можна реально спостерігати, величиною $N \approx 2m_{max}$. Наприклад, при використанні білого світла і скляних світлофільтрів $m_{max} \sim 10$, тому інтерференція спостерігається на екрані тільки в невеликій центральній зоні, про що говорилося в п. 2.2. При цьому різниця ходу променів не повинна перевищувати максимальної допустимої величини $\Delta_{max} = m_{max}\lambda$, яка називається ***довжиною когерентності***: $l_{\text{ког}} = \Delta_{max}$. Згідно з умовами (2.10) і (2.19),

$$l_{\text{ког}} = l_{max}\lambda = \frac{\lambda^2}{\Delta\lambda}. \qquad (2.20)$$

Поняття довжини когерентності дозволяє сформулювати наступний критерій:

інтерференцію світла можна спостерігати лише, коли різниця ходу променів не перевищує довжину когерентності цього світла, тобто за умови:

$$\Delta \le l_{\text{ког}}, \text{ тобто } \Delta \le \frac{\lambda^2}{\Delta\lambda}. \qquad (2.21)$$

Таким чином, випромінювання не сповна монохроматичних джерел є когерентним не скрізь, а тільки в обмеженій області простору, в якій різниця ходу променів задовольняє умову (2.21). Але це питання можна розглядати і з іншої точки зору. Якщо хвилі від точкових джерел S_1 і S_2 (рис.2.1.) є монохроматичними, але мають дещо різні частоти $\omega_1 = \omega$ і $\omega_2 = \omega + \Delta\omega$, то різниця фаз променів у точці накладання змінюється з часом:

$$\delta(t) = (\omega_2 t - \alpha_2) - (\omega_1 t - \alpha_1) = \Delta\omega \cdot t - (\alpha_2 - \alpha_1).$$

Тому при накладанні таких двох хвиль інтерференція буде спостерігатися лише протягом часу

$$\tau' = \frac{\pi}{\Delta\omega},$$

за який різниця фаз змінюється на π, і умова підсилення інтенсивності переходить в умову послаблення, чи навпаки. Величина τ' визначає граничний проміжок часу, протягом якого дані дві хвилі можна вважати когерентними, і тому називається ***часом когерентності.***

Для не повністю монохроматичних світлових пучків з частотами в усьому інтервалі від *omega* до $\omega + \Delta\omega$ умови спостереження інтерференції покращуються, оскільки в них присутні промені з різницею частот не лише $\Delta\omega$, а й з усіма меншими значеннями аж до 0. Тому, поставивши у вираз τ' замість $\Delta\omega$ середню величину $\Delta\omega/2$, отримаємо більш коректний вираз для часу когерентності $\tau_{\text{ког}}$ не повністю монохроматичних хвиль:

$$\tau_{\text{ког}} = \frac{2\pi}{\Delta\omega}. \qquad (2.22)$$

Звідси випливає, що чим меншою є величина $\Delta\omega$, тобто чим вищий ступінь монохроматичності хвиль, тим довше вони лишаються когерентними. Тому

когерентність хвиль, зумовлену близькістю їхніх частот, називають часовою когерентністю.

Відповідно, величина $\tau_{\text{ког}}$ є мірою часової когерентності не повністю монохроматичних хвиль.

Час когерентності можна виразити й через інтервал довжин хвиль $\Delta\lambda$ присутніх у випромінюванні. Для цього досить взяти до уваги, що $\omega = 2\pi c/\lambda$ і що за умови $\Delta\omega \ll \omega$ величина $\Delta\omega = (\mathrm{d}\omega)/\mathrm{d}\lambda)\Delta\lambda$. Тоді виходить:

$$\Delta\omega = \frac{2\pi c}{\lambda^2}\Delta\lambda. \qquad (2.23)$$

і зі співвідношень (2.22) і (2.20) виходить:

$$\tau = \frac{1}{c}\cdot\frac{\lambda^2}{\Delta\lambda} \quad \Rightarrow \quad l_{\text{ког}} = c\tau_{\text{ког}}. \qquad (2.24)$$

Отже, між часом і довжиною когерентності є прямий зв'язок: довжина когерентності дорівнює відстані, на яку поширюється хвиля за час когерентності. Тому часову когерентність не сповна монохроматичних хвиль можна характеризувати як часом $\tau_{\text{ког}}$, так і довжиною когерентності $l_{\text{ког}}$. Але слід зауважити, що обидві ці величини є лишень оціночними, бо чіткої межі між можливістю та неможливістю спостерігати інтерференцію не існує.

Підводячи загальний підсумок, зазначимо, що причиною неповної часової когерентності є не повна визначеність частоти випромінювання реальних джерел світла, а наслідком – обмежена кількість інтерференційних смуг область простору, де їх можна спостерігати.

Просторова когерентність. На можливість спостерігати інтерференцію негативно впливають і лінійні розміри джерел світла, наявність яких зумовлює неоднозначність різниці ходу когерентних променів, які приходять у точку спостереження від різних точок джерел. У променів, які виходять із близьких точок джерел різниця ходу менша, а в променів від віддалених точок вона більша. Це погіршує умови спостереження та якість інтерференційних смуг. Може навіть статися, що промені від одних ділянок джерел підсилюються, а від інших – послаблюються й інтерференційні смуги взагалі не спостерігаються. Це можна трактувати як те, що

випромінювання може бути більш або менш когерентним залежно від розмірів джерела. Тому когерентність, обумовлену лінійними розмірами джерел, називають ***просторовою когерентністю***.

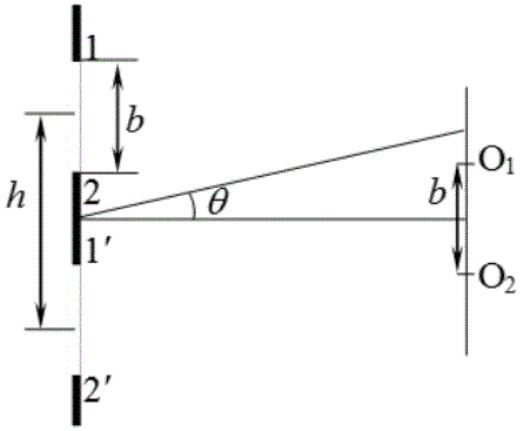


Рис. 2.6

Зрозуміло, що за будь-яких умов просторова когерентність випромінювання погіршується при збільшені розмірів джерел. Але якоїсь єдиної формули-критерію тут немає, бо просторова когерентність залежить не тільки від розмірів джерел, а й від взаємного розташування джерел і точки спостереження. Як приклад розглянемо вплив неповної просторової когерентності на інтерференційну картину від двох джерел у вигляді однакових щілин ширини b розміщених на відстані h одна від одної (рис. 2.6), які випромінюють монохроматичне світло з довжиною хвилі λ. Будемо вважати, що ці джерела складаються з безлічі пар гранично вузьких щілин – від $1-1'$ до $2-2'$, які розміщені на відстані h одна від одної. Всі вони утворюють на екрані однакові елементарні смуги з шириною Δx (формула (2.17)). Але центри інтерференційних картин від кожної пари цих віртуальних щілин зміщені одна відносно одної й розташовані між точками O_1 і O_2 всередині області шириною b, рис.2.7а.

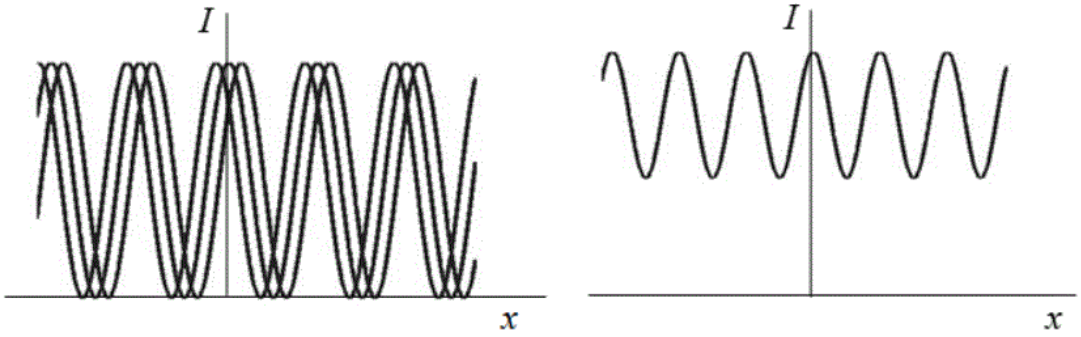


Рис. 2.7

Через це результуюча інтенсивність у максимумах зменшується, а в мінімумах збільшується, що зменшує чіткість смуг, як схематично показано на рис. 2.7 б. Тому при поступовому збільшенні ширини щілин b інтерференційні смуги по всьому полю зору будуть менш і менш виразними і в решті зовсім зникнуть. (Оскільки око може розрізняти смуги за яскравістю лише при певній мінімальній різниці інтенсивностей, при збільшенні ширини щілин буде зменшуватись не тільки чіткість, а й кількість смуг, які спостерігаються)/ Так станеться, коли відстань O_1O_2 = зрівняється із шириною інтерференційної смуги Δx, тож інтерференція буде спостерігатися тільки за умови $b \le \Delta x$. Цю умову можна розглядати як критерій просторової когерентності у схемі рис. 2.6. Відповідно до формул (2.17) і (2.17а), його можна записати, як:

$$\Delta x = \frac{l\lambda}{h} \text{ або } \Delta x = \frac{\lambda}{\Psi}, \qquad (2.25)$$

де $\Psi = h/l$ – кутова відстань між когерентними джерелами (див. рис. 2.3), тобто кут зору, під яким їх видно із зони інтерференції на екрані.

На завершення ще раз наголосимо, що від неповної просторової та часової когерентності принципово неможливо позбутися. Тому практично за будь-яких умов можна спостерігати тільки досить обмежену кількість інтерференційних смуг.

6. КОНТРОЛЬНІ ЗАПИТАННЯ

1. У чому полягає явище інтерференції? Чому воно є можливим?
2. За якої умови можлива інтерференція двох ідеальних гармонічних світлових хвиль однакової частоти?
3. Чи можлива інтерференція двох монохроматичних світлових хвиль, якщо їхні частоти відрізняються?
4. Якою може бути максимальна та мінімальна інтенсивності світла при накладанні двох когерентних променів, що мають інтенсивності I_0 та $4I_0$?
5. За якої умови інтерференційні смуги від двох когерентних джерел будуть найбільш контрастними?
6. Що називають різницею ходу променів? Визначте різницю ходу променів, які проходять до точки спостереження однакову відстань l: один у повітрі, а інший – в однорідному середовищі з показником заломлення n.
7. Який зв'язок існує між різницею фаз і різницею ходу двох когерентних променів? Некогерентних?
8. Запишіть умови інтерференційних максимумів та мінімумів при двопроменевій інтерференції якщо промені поширюються: а) в різних середовищах; б) в одному середовищі.
9. Чому умови інтерференційних максимумів і мінімумів формулюють не через різницю фаз, а через різницю ходу?
10. Чому при визначенні різниці ходу променів не враховують їхні початкові фази? Як наявність різниці початкових фаз $\delta_0 = \varphi_{02} - \varphi_{01} \neq 0$ не вплинула б на інтерференційні смуги, показані на рис2.5б?
11. Чому інтерференційні смуги спостерігаються тільки при малій відстані між монохроматичними когерентними джерелами?
12. Як зміниться ширина інтерференційної смуги при двопроменевій інтерференції, якщо одночасно змінити відстань між когерентними джерелами та відстань між джерелами й екраном в n разів? Розгляньте всі можливі варіанти.
13. Яку кількість інтерференційних смуг можна було б теоретично спостерігати на екрані при накладанні двох ідеальних когерентних хвиль% Чим би відрізнялися у такому випадку смуги від тих, що визначаються формулами (2.16) та (2.17)?
14. Чому в реальних дослідах спостерігається обмежена кількість інтерференційних смуг? Поясніть зміст поняття часової когерентності.
15. Що таке довжина та час когерентності? Як вони пов’язані між собою?
16. Виразіть граничну кількість інтерференційних смуг, яку можна спостерігати при двопроменевій інтерференції, з довжиною когерентності світла та кутовою відстанню між точковими когерентними джерелами.
17. Поясніть зміст поняття просторової когерентності та її вплив на інтерференційну картину.
18. Дві паралельні світні щілини шириною b розміщені на відстані h одна від одної і випромінюють когерентне світло з довжиною хвилі λ. На якій найменшій відстані від щілин можна поставити екран, аби спостерігати інтерференційні смуги?

5.3. СПОСТЕРЕЖЕННЯ ІНТЕРФЕРЕНЦІЇ СВІТЛА

Нижче розглянуті такі питання:

1. Інтерференційні схеми
2. Інтерференція в тонких пластинах
3. Контрольні запитання

1. ІНТЕРФЕРЕНЦІЙНІ СХЕМИ

Принцип отримання когерентних пучків. У попередній лекції (п. 2.3) говорилося, що реальне світлове випромінювання складається з окремих незалежних «шматочків»-цуґів із довільними початковими фазами. Тому світлові пучки від незалежних джерел є некогерентними і не можуть інтерферувати. Але, як свідчить дослід, інтерференцію світла можна спостерігати, взявши не два, а одне точкове джерело світла. Якщо його випромінювання якось поділити на два рознесені пучки, а потім звести їх у просторі, то на екрані в області перекриття пучків утворюється типова інтерференційна картина. Це легко пояснити. Справді, при вказаному поділі

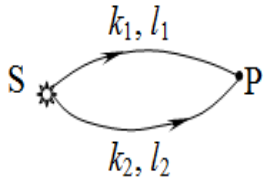


Рис. 3.1

випромінювання промені від джерела S, які потрапляють у точки накладання P, проходять різні шляхи, як схематично показано на рис. 3.1, і створюють світлові коливання

$$E_1 = E_{01}\cos\left(\omega t - k_1 l_1 + \varphi_0\right)$$

і

$$E_2 = E_{02}\cos\left(\omega t - k_2 l_2 + \varphi_0\right).$$

При цьому початкова фаза ***в джерелі*** φ_0, як зазначалося, швидко і невпорядковано змінюється з часом, але будь-якої миті вона ***однакова в обох променях***. Тому різниця фаз

$$\delta = k_2 l_2 - k_1 l_1 = \text{const}$$

не залежить від часу, тобто промені є когерентними. Але слід зауважити, що такий висновок передбачає, що в точку спостереження кожної миті приходять коливання від одного й того самого цуґу. Тому на практиці використовують випромінювання від малої ділянки джерела, яке має достатній ступінь просторової когерентності. Крім того джерело має бути високо монохроматичним для забезпечення достатньої часової когерентності.

Спосіб поділу випромінювання одного джерела на когерентні пучки не має принципового значення. На практиці використовуються різні прийоми, або ***інтерференційні схеми***, які ґрунтуються на відбиванні та заломленні світла. Розглянемо деякі з них.

Метод Юнґа. У цій інтерференційній схемі, спрощено показаній на рис.3.2, світло від джерела S спочатку пропускається крізь одну вузьку щілину і далі спрямовується на непрозорий екран E_1 з двома близько розташованими одна від одної вузькими щілинами S_1 і S_2. Після проходження крізь щілини світло поширюється в різних напрямках так, ніби випромінюється двома щілинними когерентними джерелами S_1 і S_2. Тому на екрані E_2 спостерігається система світлих і темних смуг, характерних для двопроменевої інтерференції.

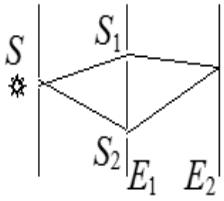


Рис. 3.2

Варто згадати, що для виконання умов часової та просторової когерентності джерело має бути високо монохроматичним, щілини – дуже вузькими, а екран – достатньо віддаленим. За таких умов у схемі Юнґа інтенсивність когерентних пучків і, ясна річ, смуг на екрані є малими. Тому вона, як говорять, є не світлосильною і не має практичного застосування. Але метод Юнґа має велику історичну цінність як перша схема, в якій на досліді було продемонстровано інтерференцію світла.

Біпризма Френеля. Біпризма Френеля являє собою дві однакові призми з дуже малим заломлюючим кутом ϑ, що склеєні малими основами (рис. 3.3) Щілинне джерело світла S розташоване на деякій відстані r від біпризми паралельно до ребра тупого кута. Світло, що падає на біпризму, після заломлення в кожній половині біпризми утворює два розбіжні пучки, обмежені променями $1-1'$ і $2-2'$. Ці пучки є когерентними, оскільки утворені поділом випромінювання від одного джерела S. Вони накладаються в секторі між променями 1–2, який є зоною інтерференції. Тому на екрані E між точками M і N спостерігаються світлі й темні інтерференційні смуги.

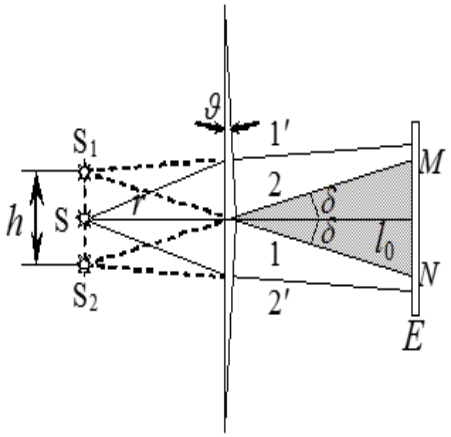


Рис. 3.3

Із закону заломлення випливає, що при малому заломлюючому куті ϑ промені після кожної половини біпризми йдуть так, якби вони виходили із однієї точки – S_1 або S_2, розташованої, як показано на рис. 3.3. Відповідно, спостерігачеві, котрий перебуває в зоні інтерференції, буде здаватися, що світло йде від двох джерел S_1 і S_2. Тому можна говорити, що біпризма Френеля замість реального некогерентного джерела S формує два уявні когерентні джерела S_1 і S_2, які дають на екрані типову картину двопроменевої інтерференції у вигляді світлих і темних смуг паралельних ребру біпризми. Положення смуг визначаються формулами (2.16), а ширина – формулою (2.17), в якій[1] $l = r + l_0$, і $h = 2(n-1)\vartheta r$ (Такий вираз h випливає із закону заломлення при малому заломлюючому куті призми, див. Лекція 5.1, п. 1.5, Приклад 1.9). Отже,

$$\Delta x = \frac{(r + l_0)\lambda}{2(n-1)\vartheta r} \approx \frac{l_0 \lambda}{2(n-1)\vartheta r}. \qquad (3.1)$$

(Тут враховано, що зазвичай біпризму розміщують неподалік від джерела, тобто $r << l_0$).

У попередній лекції (п. 2.2) було показано, що візуально спостерігати інтерференційні смуги на екрані при $l \sim 1$ м можна тільки за умови $h \le 0{,}3$ мм. Тому, прийнявши в (3.1) $r \approx 10$ см, і n =1,5 (скло), отримаємо $\vartheta \le 3 \cdot 10^{-3}$ рад $\approx 10'$. Отже, біпризма Френеля повинна мати дуже малі заломлюючі кути.

Бідзеркало Френеля. В цій схемі (рис. 3.4) для поділу фронту хвилі, що йде від щілинного джерела S, використовують не біпризму, а бідзеркало – два плоскі дзеркала, які утворюють двогранний кут близький до 180° (кут $\vartheta \sim 10'$). Внаслідок відбивання світла джерела S від половинок бідзеркала утворюються два розбіжні когерентні пучки, котрі можна вважати такими, що випромінюються двома джерелами S_1 і S_2 – уявними зображеннями джерела S у бідзеркалі. Неважко збагнути, що на рис. 3.4 всі промені від джерела S_1 йдуть вище променя 1, а від джерела S_2 – нижче променя 2. Тому зона інтерференції локалізована між променями 1 і 2. Для захисту від некогерентних прямих променів джерела S між ним і екраном встановлюють непрозору шторку Шт.

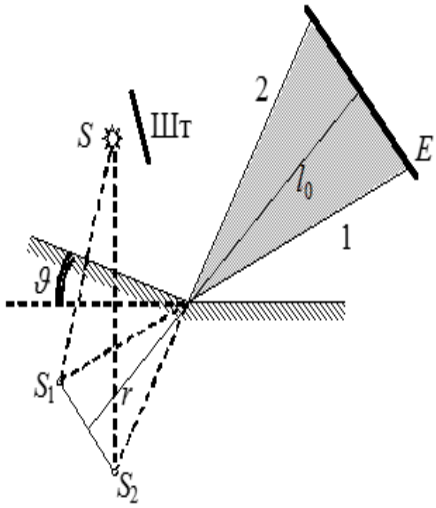


Рис. 3.4

У схемі з бідзеркалом Френеля конфігурація когерентних пучків і взаємне розташування уявних джерел та екрана є аналогічною до біпризми. Тому й ширина інтерференційної смуги у випадку бідзеркала виражається аналогічною формулою:

$$\Delta x = \frac{(r + l_0)\lambda}{2\vartheta r} \approx \frac{l_0 \lambda}{2\vartheta r}, \qquad (3.1a)$$

яка випливає із закону відбивання й відрізняється від (3.1) лише відсутністю величини $(n\text{-}1)$ у знаменнику.

Дзеркало Ллойда. В цій інтерференційній схемі для отримання когерентних пучків використовується одне плоске дзеркало Дз («дзеркало Ллойда»), рис. 3.5. Щілинне джерело S розміщують на відстані $d \ll l$ від дзеркала Дз, а когерентними джерелами є дійсне джерело S і розташоване симетрично до площини дзеркала його уявне зображення S_1. Відтак ширина інтерференційної смуги визначається формулою:

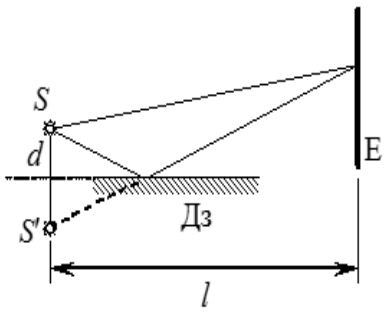


Рис. 3.5

$$\Delta x = \frac{l\lambda}{2d}. \qquad (3.2)$$

2. ІНТЕРФЕРЕНЦІЯ В ТОНКИХ ПЛАСТИНАХ

У розглянутих інтерференційних схемах когерентні хвилі, що утворюються з падаючої, мають однакові амплітуди, але поширюються в різних, хоча й близьких, напрямах, тобто мають різні фронти. Тому такий спосіб утворення когерентних світлових пучків називається ***методом поділу фронту хвилі***. Проте когерентні пучки можна отримати й ***методом поділу амплітуди***. Класичним прикладом цього методу є спостереження інтерференції при відбиванні світла від тонких прозорих пластин (плівок).

Різниця ходу відбитих променів. Нехай промінь світла від джерела *S* падає на поверхню прозорої плоскопаралельної пластини, рис. 3.6.

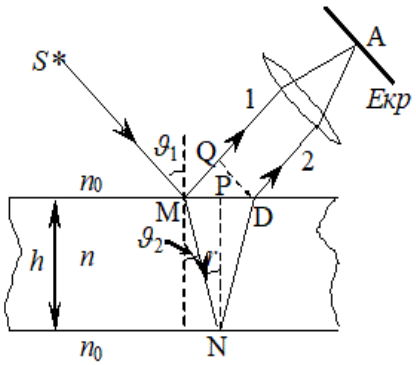


Рис. 3.6

У точці падіння M він поділяється на відбитий MQ та заломлений MN промені, що є когерентними. Далі промінь MN відбивається від нижньої поверхні пластини, потрапляє на верхню і після заломлення на ній виходить із пластини під таким самим кутом, як і промінь MQ. Так із одного падаючого утворюються два паралельні когерентні промені 1 і 2, що мають певну різницю ходу. Тому, якщо їх звести за допомогою збірної лінзи, то на екрані у фокальній площині лінзи буде спостерігатися чітка інтерференційна картина. Таку інтерференцію можна спостерігати й візуально: у цьому випадку лінзою є кришталик, а екраном – сітківка ока.

Результат інтерференції в кожній точці екрана визначається різницею ходу променів 1 і 2. Для її розрахунку позначимо: показники заломлення середовища та пластинки, як n_0 і n, довжини відрізків $MQ = l_0$, MN = ND = *l*, MD = *d*, товщину пластини *h*.

Шляхи променів 1 і 2 від точок Q, D до точки A однакові. Тому їхня різниця ходу визначається різницею оптичних шляхів Δ' між точками M, Q і D, відповідно, та втратою півхвилі на одній із поверхонь На якій саме, залежить від співвідношення між величинами *n* і n_0, (див. Лекція 5.1, п. 3):

$$\Delta = \Delta' - \frac{\lambda}{2}, \qquad (3.3)$$

де

$$\Delta' = 2nl - n_0 l_0. \qquad (3.3а)$$

Величину Δ' не важко обчислити на основі геометричних міркувань і законів відбивання та заломлення світла. З рис.3.6 випливає, що

$$l = \frac{h}{\cos\vartheta_2} = \frac{h}{\sqrt{1-\sin^2\vartheta_2}}, \quad l_0 = d\sin\vartheta_1 = 2h\,\mathrm{tg}\,\vartheta_2\sin\vartheta_1 = \frac{2d\sin\vartheta_2\sin\vartheta_1}{\sqrt{1-\sin^2\vartheta_2}}.$$

Із цих виразів після замін

$$\operatorname{tg}\vartheta_2 = \frac{\sin\vartheta_2}{\sqrt{1-\sin^2\vartheta_2}} \quad \text{і} \quad \sin\vartheta_2 = \frac{n_0}{n}\sin\vartheta_1$$

отримаємо:

$$\Delta' = 2h\sqrt{(n/n_0)^2 - \sin^2\vartheta_1},$$

і

$$\Delta = 2h\sqrt{(n/n_0)^2 - \sin^2\vartheta_1} - \frac{\lambda}{2}. \qquad (3.4)$$

Якщо світло падає на пластину з повітря, то $n_0 = 1$ і $\lambda = \lambda_0$ – довжина світлової хвилі у вакуумі. Тоді

$$\Delta = 2h\sqrt{n^2 - \sin^2\vartheta_1} - \frac{\lambda_0}{2}. \qquad (3.4а)$$

При нормальному падінні $(\vartheta_1 = 0)$:

$$\Delta = 2hn - \frac{\lambda_0}{2}. \qquad (3.4б)$$

Згідно з виразом (3.4*а*) і умовами (2.10), (2.10*а*), максимуми інтерференції будуть спостерігатися, коли

$$2h\sqrt{n^2-\sin^2\vartheta_1}=\left(m+\frac{1}{2}\right)\lambda_0, \qquad (3.5)$$

а мінімуми при

$$2h\sqrt{n^2-\sin^2\vartheta_1}=m\lambda_0, \qquad (3.5a)$$

де $m=0,1,2,3,\ldots$ – порядок інтерференції.

При нормальному падінні світла на пластину ці умови спрощуються і набувають вигляду:

для максимумів

$$2hn=\left(m+\frac{1}{2}\right)\lambda_0 \quad \text{або} \quad 2h=\left(m+\frac{1}{2}\right)\lambda, \qquad (3.6)$$

і для мінімумів

$$2hn=m\lambda_0 \qquad \text{або,} \qquad 2h=m\lambda \qquad (3.6a)$$

де $\lambda=\lambda_0/n$ – довжина світлової хвилі в пластині.

Зауваження. При інтерференції в тонких пластинах через ефект утрати півхвилі на поверхні умови максимумів і мінімумів немов «обмінюються місцями». Про це корисно пам'ятати при розв'язуванні задач.

Перед тим, як розглядати конкретні випадки інтерференції у прозорих пластинах, зауважимо, що таку інтерференцію можна спостерігати тільки в дуже тонких пластинах (плівках). Це випливає з умов (3.6), (3.6*a*) та вимоги часової когерентності (2.21), за якими для максимальної припустимої товщини пластини виходить:

$$\Delta_{max} \approx 2nh_{max}, \qquad h_{max} \approx \frac{\lambda^2}{2n\Delta\lambda}.$$

Прийнявши для оцінки n =1,5 (скло), λ = 0,6 мкм (жовте світло) і $\left(\lambda/\Delta\lambda\right)=100$, отримаємо для граничної товщини пластини h_{max} = 20 мкм.

Ще одне зауваження стосується того, що промені 1 і 2 є не єдині когерентні промені, що утворюються при поділі падаючого. Справді, промінь MN не лише відбивається, а й заломлюється на нижній поверхні пластини, а промінь ND не лише заломлюється, а й відбивається від верхньої поверхні, а потім знов відбивається й заломлюється на нижній, і так далі. Отже, з кожного падаючого променя утворюється по два ряди формально когерентних променів – тих, які відбиваються від пластини та тих, які проходять крізь неї. Тому інтерференція спостерігається як для відбитих, так і для заломлених променів. При цьому заломлений промінь не втрачає півхвилі на поверхнях пластини. Тому у випадках, коли у відбитих променях утворюються максимуми, в заломлених – мінімуми, і навпаки.

Залежно від умов спостереження інтерференції в тонких пластинах розрізняють декілька окремих типів інтерференційних смуг.

Кольори тонких пластин. Із умов (3.6), (3.6*а*) випливає, що розподіл інтенсивностей (інтерференційна картина) при відбиванні світла від тонких пластин залежать від кута падіння та товщини пластини в кожній точці поверхні. Тому за різних умов спостерігається декілька характерних видів інтерференції в тонких пластинах.

Розглянемо для початку випадок, коли з повітря на всю поверхню плоскопаралельної пластини падає плоска монохроматична хвиля, інакше – паралельний світловий пучок, із кутом падіння $\vartheta_1 = const$. У такому разі різниці ходу всіх відбитих когерентних променів будуть однакові. Отже, результат інтерференції в усіх точках сходження променів теж буде однаковим, і інтерференційні смуги утворюватися не будуть. Але ефект інтерференції в даному випадку виявлятиметься в періодичній зміні інтенсивності відбитого світла при поступовій зміні кута падіння. При кутах ϑ_1, що задовольняють умову мінімумів (3.5*а*), пластинка у відбитих променях буде темною, а за умови (3.5) – буде яскраво виблискувати. Особливо ефектно така інтерференція виглядає при опроміненні незабарвленої пластини білим світлом. Біле світло є сумішшю променів усіх кольорів від червоного до фіолетового. Але при кожному заданому куті падіння умова максимуму (3.5) точно виконується тільки для якоїсь однієї довжини хвилі λ. Тому пластина буде відбивати тільки промені, що мають близькі довжини хвиль і, відповідно, однаковий колір. Тож попри опромінення білим світлом, пластина буде виглядати забарвленою у той чи інший колір залежно від кута падіння світла на її поверхню. Тому розглянутий тип інтерференції називають «***кольорами тонких пластин***».

При поступовій зміні кута падіння забарвлення пластинки буде неперервно змінюватися, проходячи через усі кольори райдуги. Колір пластини буде змінюватися також і при зміні її товщини.

Кольорами тонких пластин, зокрема, є добре відоме забарвлення тонких плівок бензину на поверхні калюж. При цьому рухливість і мінливість кольорів пояснюється неконтрольованою невеликою зміною товщини плівки в різних місцях завдяки рухові повітря.

Лінії рівного нахилу. Інший тип інтерференції спостерігається при опроміненні плоскопаралельної пластини світлом від точкового джерела. В цьому випадку при заданих товщині пластини, показнику заломлення та довжині світлової хвилі результат інтерференції відбитих променів визначається тільки величиною кута падіння ϑ_1 світла на пластину. При даному куті падіння всі когерентні промені після накладання утворюють одну інтерференційну смугу – світлу, якщо величина ϑ_1 задовольняє умову (3.5), або темну за умови (3.5*а*). Якщо переміщуватись уздовж пластини в певному напрямку, величина ϑ_1 буде почергово задовольняти умови (3.5) і (3.5*а*) при різних порядках m. Як наслідок, при відбиванні від пластини світла точкового джерела інтерференційна картина матиме вигляд темних та світлих смуг, кожній з яких відповідає задане значення кута ϑ_1, інакше – однаковий нахил падаючих променів. Тому цей тип інтерференції називається «*лініями рівного нахилу*». Форма ліній рівного нахилу залежить від геометрії джерела світла. Наприклад, для точкового джерела це дуги кола, а для лінійного (щілинного) джерела паралельного до пластини – прямі смуги.

Лінії рівної товщини. Цей тип інтерференції спостерігається в тонких пластинах із не паралельними поверхнями, зокрема – в тонкому плоскому клині з малим кутом $\delta = const$ при вершині (кутом сходження поверхонь, рис. 3.7).

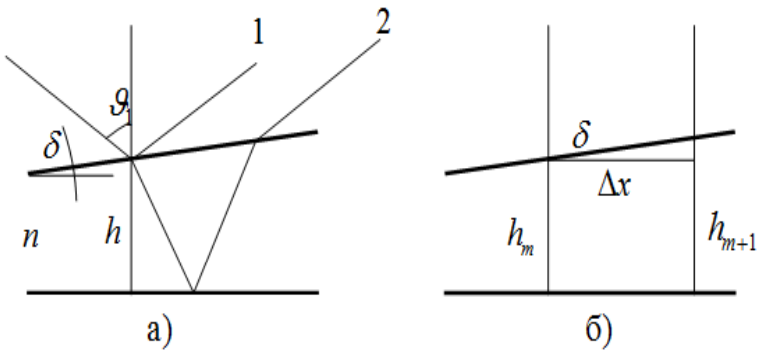


Рис. 3.7

Когерентні промені в клині утворюються так само, як і в плоскопаралельній пластині – внаслідок відбивання від обох поверхонь. Відміна лише в тому, що через непаралельність поверхонь шляхи заломленого променя від точки падіння до нижньої поверхні і після відбивання – до верхньої поверхні клина, строго говорячи, не однакові, а промені 1 і 2 не паралельні. Але через малість кута δ ці відмінності є неістотними. Тому для клина різниця ходу променів і умови утворення інтерференційних смуг з великою точністю визначаються тими самими формулами, що й для плоскопаралельної пластини.

При опроміненні клина паралельними променями кут падіння $\vartheta_1 = const$, і різниця ходу Δ відбитих променів (3.4) визначається тільки товщиною клина h в точці падіння. При переміщенні вздовж поверхні клина величина Δ монотонно змінюється й періодично задовольняє умови мінімумів і максимумів. Тому при опроміненні клина широким паралельним світловим пучком у відбитих променях утворюється система інтерференційних смуг. Такі смуги називаються ***«лініями рівної товщини»***, бо кожна з них відповідає одній і тій самій товщині клина. Наприклад, при нормальному падінні, згідно з умовами (3.6) і (3.6*а*), темна смуга порядку m

утворюється при відбиванні від усіх точок, де клин має фіксовану товщину

$$h_m = m\frac{\lambda_0}{2n} = m\frac{\lambda}{2}, \qquad (3.7)$$

а для світлої смуги

$$h_m = \left(m + \frac{1}{2}\right)\frac{\lambda_0}{2n} = \left(m + \frac{1}{2}\right)\frac{\lambda}{2}, \qquad (3.7а)$$

де $m = 0,1,2, \ldots$, і λ – довжина світлової хвилі у клині.

Із рис. 3.7б видно, що ширина інтерференційної смуги (відстань між сусідніми мінімумами) дорівнює $\Delta x = (h_{m+1} - h_m)/\delta$ і, при нормальному падінні, згідно з умовою (3.7), визначається формулою:

$$\Delta x = \frac{\lambda_0}{2n\delta} = \frac{\lambda}{2\delta}. \qquad (3.8)$$

Із цієї формули випливає, що у плоскому клині ($\vartheta = \text{const}$) величина $\Delta x = \text{const}$, тобто лінії рівної товщини розташовані на однаковій відстані одна від одної.

Вираз (3.8) пояснює також, чому інтерференцію в клині можна візуально спостерігати тільки при малих кутах клина ϑ. Справді, прийнявши $\lambda_0 = 600$ нм (жовте світло), n =1,5 (скло) і $\Delta x_{\min} = 0,2$ мм (роздільна здатність ока), отримаємо:

$$\frac{\lambda_0}{2n\delta} \geq \Delta x_{\min} \quad \delta \leq \frac{\lambda_0}{2n\Delta x_{\min}} = 10^{-3} \text{ рад} \approx 4'.$$

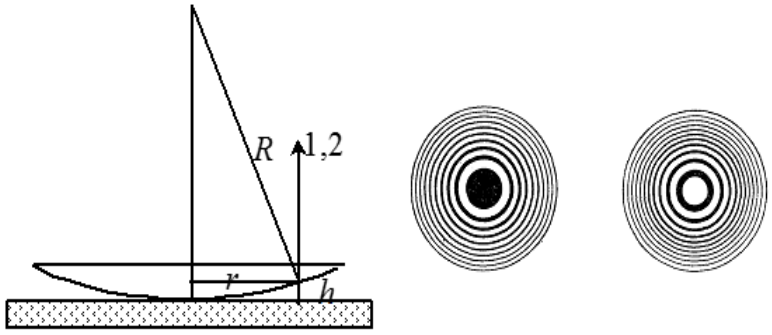
Це дуже мала величина, тож поверхні клина мають бути майже паралельні. (Практично такий повітряний клин можна отримати, поклавши скраю між двома скляними пластинками волосину).

Форма ліній рівної товщини визначається профілем поверхні клина. Наприклад, у плоскому клині лінії рівної товщини мають вид еквідистантних прямих смуг паралельних до ребра клина. Тому будь-які відхилення у формі ліній чи відстані між ними вказує на те, що поверхня не є плоскою. Це використовується в оптичній промисловості для контролю форми поверхонь виробів, які виготовляються. Так само, якщо поверхня має якісь мікродефекти, наприклад подряпини, то це викличе локальне спотворення інтерференційних смуг у відповідному місці. Це дозволяє візуально контролювати чистоту поверхонь оптичних виробів. Слід зазначити, що лінії рівної товщини в місці подряпини зміщуються на одну ширину смуги Δx при глибині подряпини в одну довжину хвилі λ. Зсув інтерференційних смуг нескладно контролювати з точністю $0,1\Delta x$ та навіть більшою, що дозволяє виявляти й усувати нерівності розміром в $0,1\lambda$, тобто всього $\sim 10^{-8}$ м.

Кільця Ньютона. Так називаються інтерференційні смуги, що виникають при опромінюванні плоскоопуклої лінзи покладеної опуклою стороною на товсту плоску прозору пластину, рис. 3.8. Інтерференційна картина створюється променями 1 і 2, відбитими від суміжних поверхонь лінзи та пластини. Ці поверхні утворюють повітряний клин, отож інтерференційні смуги є лініями рівної товщини і при нормальному опроміненні мають форму концентричних світлих і темних кілець.

Інтерференційні смуги в клині можна спостерігати тільки при малій товщині та куті клина. Тому для спостереження кілець Ньютона використовують лінзи з великим радіусом кривизни (з малою оптичною силою), причому чіткі смуги спостерігаються лише в невеликому околі точки дотику лінзи та пластини, яка є центром інтерференційної картини. За таких умов можна вважати, що промені 1 і 2 відбиваються в зворотному до падаючих променів напрямку. Тому їхня різниця ходу визначається виразом (3.4*б*) при n =1:

$$\Delta = 2h - \frac{\lambda}{2}. \qquad (3.9)$$



а) **б)** **в)**

Рис. 3.8

Як видно з рис. 3.8а, товщина повітряного зазору h, радіус кривизни лінзи R і радіус інтерференційного кільця r пов'язані співвідношенням:

$$r^2 = R^2 - (R-h)^2 \quad \Rightarrow \quad r^2 = 2Rh + h^2.$$

Оскільки $h \ll 1$, величина h^2 є неістотною, тому

$$r^2 = 2Rh. \qquad (3.9a)$$

Підставляючи сюди вирази h, які відповідають мінімумам (3.7) і максимумам (3.7 *а*) інтерференції, отримаємо наступні вирази для радіусів кілець Ньютона.

темні: $$r_m = \sqrt{mR\lambda}, \qquad m = 0,1,2,\ldots, \qquad (3.10)$$

світлі: $$r_m = \sqrt{(2m+1)R\lambda/2}, \quad m = 0,1,2,\ldots. \qquad (3.10a)$$

Ці формули можна об'єднати в одну:

$$r_m = \sqrt{kR\lambda/2}, \qquad (3.10б)$$

в якій парні значення k відповідають темним, а непарні – світлим кільцям.

Картина кілець Ньютона у відбитих променях, яка схематично показана на рис. 3.8б, має дві особливості. По-перше, центральне темне кільце, для якого $m = 0$ і $r = 0$, начебто мало б виглядати як темна точка. Але в дійсності спостерігається темний круг помітного радіуса. Це пояснюється тим, що інтерференційне послаблення відбитих променів спостерігається не лише в точці О, а і в прилеглій області, де різниця ходу променів ще мало відрізняється від $\lambda/2$. По-друге, відстань між сусідніми темними кільцями (ширина інтерференційної смуги), що визначається формулою

$$\Delta r_m = r_{m+1} - r_m = \sqrt{R\lambda}\left(\sqrt{m+1} - \sqrt{m}\right),$$

зменшується при збільшенні порядку m. Отож при віддаленні від центра смуги поступово зближуються і врешті перестають розрізнятися. Тому навіть при використанні високо монохроматичного світла можна спостерігати лишень обмежену кількість кілець.

Інтерференційні кільця утворюють і промені, що після подвійного відбивання від поверхонь лінзи та пластини проходять у напрямку падіння світла на лінзу. В цьому випадку кожен із променів двічі відбивається від більш густого середовища і двічі втрачає півхвилі. Тому в центрі утворюється максимум, рис. 3.8в. Але в даному випадку інтерференційні смуги утворюються на тлі потужного світлового потоку, що напряму проходить крізь лінзу та пластину. Це утруднює спостереження кілець Ньютона на просвіт.

На завершення додамо, що відслідковуючи форму та локальні спотворення кілець Ньютона, можна контролювати і коригувати форму та чистоту поверхні лінз у процесі виготовлення.

Просвітлення оптики. Як говорилося раніше (лекція 1, п. 1.3), в оптичних приладах втрати світла на відбивання на прозорих поверхнях можуть бути значними. Наприклад, у складних об'єктивах вони досягають 50% і більше. Тому для зменшення цих утрат застосовують технологію ***просвітлення оптики,*** що ґрунтується на інтерференції світла в тонких плівках. Ідея полягає в тому, що в прозорій плівці внаслідок інтерференції погашаються відбиті від її поверхонь промені із довжиною хвилі λ, що задовольняє умову (3.6*а*). При цьому послаблюються також і промені з усіма прилеглими довжинами хвиль, які задовольняють умову часової когерентності. Тому для зменшення втрат на відбивання на лінзи та інші прозорі деталі наносять тонкі прозорі плівки, показник заломлення й товщину яких підбирають так, аби в робочому діапазоні довжин хвиль відбиті промені були максимально послаблені, а ті, що проходять, відповідно, мали би максимальну можливу інтенсивність.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Поясніть, чому не можна спостерігати інтерференцію світла від двох незалежних джерел.
2. Поясніть принцип отримання когерентних світлових пучків. Чому при встановленні умов інтерференційних максимумів і мінімумів не враховують початкові фази променів, які накладаються?
3. Як зміниться інтерференційна картина на екрані в досліді Юнга (рис. 3.2), якщо перекрити скляною пластинкою: а) обидва отвори в екрані ; б) тільки один отвір?
4. Як буде змінюватись інтерференційна картина на екрані в досліді з біпризмою Френеля (рис. 3.3) при збільшенні відстані між біпризмою та: а) джерелом; б) екраном?
5. На якій відстані від лінії перетину площин дзеркала та екрана в схемі Ллойда розташовані на екрані максимум і мінімум порядку $m = 0$ при довжині світлової хвилі і відстані між монохроматичним джерелом (вузькою світною щілиною) і площиною дзеркала рівній d ?
6. Поясніть природу райдужного забарвлення плівок бензину на поверхні води та мильних пухирів.
7. Коли ми спостерігаємо райдужне забарвлення бензинової плівки на поверхні калюжі, то де вона товща – там, де має кольори в синій частині спектра, чи там де в червоній?
8. Чим визначається ширина інтерференційної смуги при відбиванні монохроматичного світла від плоского клина? Чому кут клина має бути дуже малим?
9. Від чого залежить відстань від ребра тонкого клина до останньої інтерференційної смуги, що спостерігається у відбитих променях?
10. При спостереженні кілець Ньютона лінзу через світлофільтр опромінювали світлом від звичайної електричної лампочки. При цьому для двох різних світлофільтрів з однаковою довжиною хвилі в максимумі смуги пропускання спостерігалася різна кількість кілець. Поясніть вірогідну причину вказаної відмінності.
11. Як і в скільки разів зміняться радіуси кілець Ньютона, якщо зазор між лінзою та пластиною заповнити прозорою рідиною з показником заломлення $n = 1{,}69$?
12. Опишіть, як має змінюватися картина кілець Ньютона при дуже повільному підійманні лінзи без зміни її просторової орієнтації.

5.4. ДИФРАКЦІЯ СВІТЛА

У даній лекції розглядаються такі питання:

1. Явище дифракції. Принцип Гюйгенса-Френеля
2. Дифракція Фраунгофера на щілині
3. Дифракція Фраунгофера на ґратці
4. Дифракція рентгенівських променів на кристалах
5. Контрольні запитання

1. ЯВИЩЕ ДИФРАКЦІЇ. ПРИНЦИП ГЮЙГЕНСА-ФРЕНЕЛЯ

Дослід свідчить, що у в однорідному вільному від перешкод середовищі будь-які хвилі поширюються прямолінійно. Цей факт інколи розглядають навіть як один із законів геометричної оптики. Але коли хвиля натрапляє на якесь непрозоре тіло чи на отвір у непрозорому екрані, вона відхиляється від прямолінійного напрямку поширення й оминає перешкоду та заходить у область геометричної тіні. **Таке**

відхилення хвиль від прямолінійного напрямку поширення у середовищах із різкими неоднорідностями, яке не зумовлене відбиванням чи заломленням, називається дифракцією.

Дифракцію легко спостерігати на поверхні води. Якщо на шляху хвилі трапляється невелика перешкода, що виступає з води, скажімо камінь, то тінь (тиха вода) утворюється за перешкодою тільки у деякому збіжному конусі, а далі хвиля поширюється так, ніби перешкоди й не було. При цьому можна помітити, що повз перешкоду значних розмірів, приміром, великий щит, хвиля проходить практично без відхилення. Така особливість є притаманною й хвилям іншої фізичної природи. Скажімо, дифракцію звукових хвиль ми спостерігаємо, коли, стоячи впритул до стіни будинку, чуємо голос людини, яка знаходиться за рогом. У випадку світла одним з проявів дифракції є те, що після малого отвору в непрозорому екрані світло поширюється по всіх напрямках. Це, зокрема, використовується в досліді Юнґа, про який говорилося в попередній лекції.

Принцип Гюйгенса-Френеля. Зважаючи на електромагнітну природу світла, теоретично будь-яку задачу дифракції можна вирішити за допомогою загальних рівнянь електромагнітного поля (рівнянь Максвелла). Але це пов'язано з великими математичними труднощами, тому на практиці дифракцію та інші задачі оптики розглядають на основі спрощених нестрогих методів аналізу, одним із яких є принцип Гюгенса-Френеля. Згідно з принципом Гюйгенса (див. Лекція 5.1, п. 1.2) всі точки хвильового поля можна розглядати як джерела елементарних вторинних сферичних хвиль. Це дозволяє за допомогою відповідної геометричної побудови прослідкувати еволюцію фронту хвилі, відтак – її поширення.

Розглянемо в такий спосіб поширення плоскої світлової хвилі, що падає по нормалі на непрозорий екран з отвором, при двох різних розмірах отвору (рис 4.1). Для цього виділимо деяку множину точок на поверхні кожного отвору і покажемо маленькими півколами фронти елементарних вторинних хвиль від кожної точки. Обвідна всіх півкіл зображує фронт результуючої хвилі одразу після проходження крізь отвір, а лінії, проведені з кожної точки по нормалі до обвідної, показують світлові промені, що виходять із отвору. Розглядаючи таку *побудову Гюйгенса*, можна зробити два суттєві висновки. По-перше, на краях отвору відбувається викривлення фронту хвилі, що означає відхилення променів в бік геометричної тіні, тобто – дифракцію. І, по-друге, у випадку а) відхиляється набагато менша частка променів, ніж у випадку б). Інакше говорячи, дифракція виражена тим сильніше, чим менший розмір отвору.

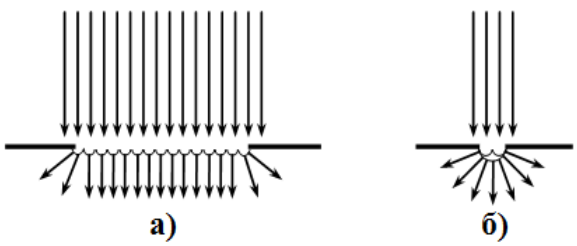


Рис. 4.1

Такий прогноз, на загал, узгоджується з дослідом. Але реальна картина дифракції є складнішою. А саме, на поставленому за отвором екрані спостерігається не просто розмите зображення отвору, а упорядкована система світлих і темних смуг, форма котрих відображає форму отвору. Пояснення цьому дав Френель, який доповнив принцип Гюйгенса уявленням про когерентність та інтерференцію вторинних хвиль. Підставою для такої думки є те, що в усіх точках хвильової поверхні фаза коливань однакова. Об'єднаний ***принцип Гюгенса-Френеля*** спрощено можна сформулювати наступним чином.

> ***Всі точки відкритої хвильової поверхні даної хвилі є джерелами елементарних вторинних когерентних хвиль. При цьому амплітуда (й інтенсивність) коливань, що збуджуються хвилею в будь-якій точці спостереження, визначається результатом інтерференції всіх вторинних хвиль, які приходять у цю точку.***

Відповідно до цієї концепції Френель запропонував і алгоритм розрахунку відносних амплітуд. Для визначення результуючих коливань $E(t)$, які створюються в заданій точці спостереження P відкритою частиною S фронту

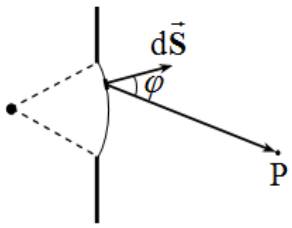


Рис. 4.2

монохроматичної хвилі (рис. 4.2) необхідно додати з урахуванням амплітуд і фаз елементарні коливання $\mathrm{d}E$, що приходять у точку P від усіх елементарних площадок $\mathrm{d}S$ (точкових вторинних джерел) поверхні S: $E = \int \mathrm{d}E$. Оскільки кожне вторинне джерело випромінює елементарну сферичну хвилю, то

$$\mathrm{d}E = \frac{\mathrm{d}a}{r}\cos(\omega t - kr), \qquad (4.1)$$

r – відстань від площадки $\mathrm{d}S$ до точки спостереження, ω і k – відповідно частота та хвильове число, $\mathrm{d}a$ – амплітудний множник. Згідно з Френелем, величина $\mathrm{d}a$ є прямо пропорційною площі ділянки $\mathrm{d}S$ і залежить від її просторової орієнтації, так що $\mathrm{d}a = K(\varphi)a_0\mathrm{d}S$, де коефіцієнт $K(\varphi)$ визначається кутом φ між нормаллю до площадки $\mathrm{d}S$ і напрямом на точку спостереження P так, що при $\varphi = 0$ $K = 1$, а при $\varphi = 90°$ $K = 0$.

Відповідно,

$$\mathrm{d}E = K(\varphi)\frac{a_0}{r}\cos(\omega t - kr)\mathrm{d}S, \qquad (4.1a)$$

і

$$E = \int_S K(\varphi)\frac{a_0}{r}\cos(\omega t - kr)\mathrm{d}S. \qquad (4.2)$$

Цей інтеграл можна розглядати як математичний вираз принципу Гюйгенса-Френеля.

Дифракція Френеля та дифракція Фраунгофера. Через невизначеність коефіцієнта $K(\varphi)$ інтеграл (4.2) на загал не дозволяє обчислювати абсолютні значення напруженості поля та інтенсивності світлової хвилі в різних точках. Але в деяких важливих випадках він дозволяє визначати відносні величини, тобто встановлювати характеристики просторового розподілу інтенсивностей дифрагованої хвилі.

Залежно від умов спостереження вирізняють два види дифракції: ***дифракцію Френеля*** і ***дифракцію Фраунгофера***. У першому випадку джерело світла, перешкода та екран для спостереження розташовані на порівняно невеликій відстані одне від одного. Через це на перешкоду та на екран падають не паралельні промені, тобто неплоскі хвилі. При дифракції Фраунгофера на перешкоду спрямовують і на екрані спостерігають результат дифракції плоских хвиль. Плоским хвилям відповідають паралельні промені, тому цей вид дифракції ще називають ***дифракцією в паралельних променях***.

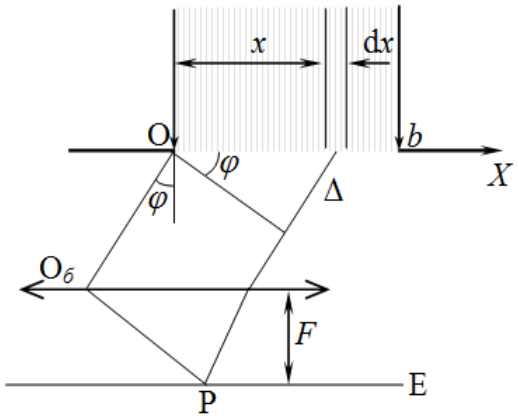


Рис. 4.3

Дифракцію Фраунгофера технічно легше спостерігати, і вона знаходить застосування в оптичних приладах. Окрім того фраунгоферову дифракцію простіше досліджувати теоретично. Тому далі розглянемо тільки цей вид дифракції. Схема спостереження показана на рис. 4.3. Для отримання паралельних променів, які спрямовуються на перешкоду, скажімо вузьку щілину, використовують лазерний пучок, або світловий пучок від збірної лінзи, у фокусі котрої знаходиться точкове джерело. За перешкодою також ставлять збірну лінзу (об'єктив) $O_б$ і в її фокальній площині – екран для спостереження E. За такої умови в будь-яку точку на екрані приходять тільки промені, що після перешкоди падають на лінзу під одним і тим самим кутом дифракції. Тому на екрані спостерігається картина саме фраунгоферової дифракції.

2. ДИФРАКЦІЯ ФРАУНГОФЕРА НА ЩІЛИНІ

Дифракція Фраунгофера на щілині. Розглянемо за допомогою принципу Гюйгенса-Френеля фраунгоферову дифракцію паралельного монохроматичного пучка променів із довжиною хвилі λ, який падає по нормалі на довгу (формально – нескінченну) щілину ширини b. Для цього подумки розіб'ємо щілину (тобто відкриту частину хвильового фронту падаючого пучка) на елементарні поздовжні смужки ширини $\mathrm{d}x$ і запишемо вираз для елементарних коливань $\mathrm{d}E$, які збуджуються на екрані променями, що виходять із однієї смужки під заданим кутом дифракції φ і визначаються загальним виразом (4.1a). При цьому врахуємо, що реально значима дифракція спостерігається тільки при малих кутах φ і при малій ширині щілини b, тож в (4.1a) можна покласти K =1 і $a_0/r = \mathrm{const}$. Крім того очевидно, що площа смужки є прямо пропорційною її ширині: $\mathrm{d}S \sim \mathrm{d}x$. Відтак вираз (4.1$a$) набуває вигляду

$$\mathrm{d}E = A\cos\big(\omega t - \delta(x)\big)\mathrm{d}x,$$

де A = const, $\delta(x)$ – різниця фаз коливань, які збуджуються в точці спостереження на екрані довільною (з координатою x) і першою смужки (з координатою $x = 0$) смужками. Відповідно, для результуючої напруженості маємо:

$$E = A\int_0^b \cos\big(\omega t - \delta(x)\big)\mathrm{d}x,$$

Для обчислення інтеграла треба знайти вираз $\delta(x)$. Це легко зробити, бо, як доводить теорія і підтверджує дослід, лінза не вносить додаткової різниці ходу в промені, що проходять крізь неї. Тому різниця ходу променів, які розглядаються, $\Delta(x) = x\sin\varphi$ (див. рис. 4.3). Отже, за формулою (2.9) з лекції 5.2 величина $\delta(x)$ визначається як

$$\delta(x) = \frac{2\pi\sin\varphi}{\lambda}x.$$

Отже

$$E = A\int_0^b \cos\left(\omega t - \frac{2\pi\sin\varphi}{\lambda}x\right)\mathrm{d}x.$$

Задля наочності введемо нову змінну:

$$\omega t - \frac{2\pi\sin\varphi}{\lambda}x = \theta \quad\Rightarrow\quad \mathrm{d}x = -\frac{\lambda}{2\pi\sin\varphi}\mathrm{d}\theta.$$

Тоді

$$E = A\frac{\lambda}{2\pi\sin\varphi}\big(\sin\theta(0) - \sin\theta(b)\big),$$

де $\theta(0)=\omega t$ і $\theta(b)=\omega t-\dfrac{2\pi b\sin\varphi}{\lambda}$. Підставивши ці значення та скориставшись формулою різниці синусів, отримаємо рівняння результуючих коливань поля в точці спостереження:

$$E=A\frac{\lambda}{\pi\sin\varphi}\sin\frac{\pi b\sin\varphi}{\lambda}\cos\left(\omega t-\alpha_0\right),$$

де початкова фаза $\alpha_0=\pi b\sin\varphi/\lambda$. Амплітуду цих коливань визначається формулою:

$$E_0=Ab\frac{\sin\left(\pi b\sin\varphi/\lambda\right)}{\left(\pi b\sin\varphi/\lambda\right)}.$$

Відповідно, інтенсивність $I\sim E_0^2$ визначається наступною *дифракційною формулою*:

$$I_\varphi=I_0\frac{\sin^2\left(\pi b\sin\varphi/\lambda\right)}{\left(\pi b\sin\varphi/\lambda\right)^2}. \qquad (4.3)$$

Функція (4.3) визначає розподіл інтенсивностей (*дифракційну картину*), яка є зображенням щілини на екрані при фраунгоферовій дифракції. Властивості такої функції добре відомі з математики. При $\varphi=0$ (центр екрана) і чисельник, і знаменник в (4.3) є нульовими, тобто маємо невизначеність $\left(0/0\right)$. Проте вона легко розкривається за правилом Лопіталя і дає $I\left(0\right)=I_0$. А коли чисельник обертається на нуль при $\varphi\neq 0$, маємо мінімуми інтенсивності $I=0$. Це буде за умови $\left(\pi b\sin\varphi/\lambda\right)=\pm m\pi$, або

$$b\sin\varphi=\pm m\lambda,\quad m=1,2,\ldots. \qquad (4.4)$$

Цей вираз називається ***умовою дифракційних мінімумів щілини***.

Між мінімумами розміщені ***побічні максимуми***. Строге дослідження їхнього положення та величини пов'язане з певними незручностями, але наближено можна вважати, що побічні максимуми розміщуються посередині між мінімумами, тобто за умови

$$b\sin\varphi=\pm\left(m+\frac{1}{2}\right)\lambda,\quad m=1,2,\ldots. \qquad (4.5)$$

У такому разі для інтенсивностей побічних максимумів, які близькі до центра екрана ($\varphi<<1$, $\sin\varphi=\varphi$), формула (4.3) дає:

$$I'_m=\frac{4}{\pi^2\left(2m+1\right)^2}I_0. \qquad (4.6)$$

З цього виразу випливає, що побічні максимуми є слабкими самі по собі й поступово деградують. Для інтенсивностей перших трьох побічних максимумів формула (4.6) дає всього $0{,}045I_0$, $0{,}016I_0$ та $0{,}008I_0$, відповідно.

Графік розподілу інтенсивностей в залежності від $\sin\varphi$ показаний на рис. 4.4. У центрі екрана спостерігається інтенсивна світла смуга (центральний дифракційний максимум), оточена з обох боків слабкими вторинними смугами (побічними максимумами), що розділені темними смугами (мінімумами), положення яких визначається умовою (4.4).

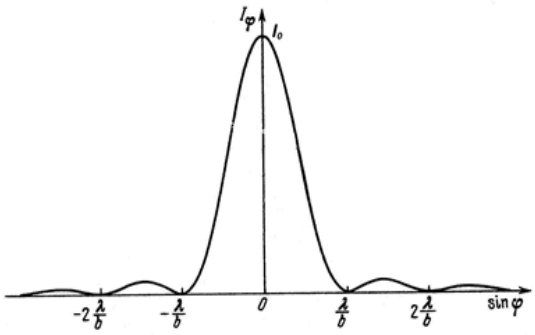


Рис. 4.4

Слід зауважити, що практично описану дифракцію можна спостерігати тільки, коли ширина щілини є малою і лежить у визначених межах. Принцип Гюйгенса-Френеля передбачає, що розміри перешкод є суттєво більші за довжину хвилі, але з умов (4.4), (4.5) видно, що збільшення ширини щілини веде до зменшення кутів дифракції та стягання дифракційних смуг до центра, так що вони врешті решт перестають розрізнятися. Як наслідок, дифракція реально спостерігається при ширині щілини в десятки – сотні довжин світлової хвилі.

Подібна картина у вигляді концентричних світлих і темних кілець спостерігається і для круглого отвору, але розрахунки в цьому випадку є складнішими.

3. ДИФРАКЦІЯ ФРАУНГОФЕРА НА ҐРАТЦІ

Дифракційна ґратка. Через малу інтенсивність смуг дифракція на окремій щілині чи отворі не знаходить практичних застосувань. Але ефект значно підсилюється в *дифракційній ґратці* – впорядкованій системі перешкод для проходження світла.

Найпростіша одновимірна дифракційна ґратка являє собою систему однакових довгих вузьких щілин, які розділені однаковими непрозорими проміжками, як схематично показано в розрізі на рис. 4.5.

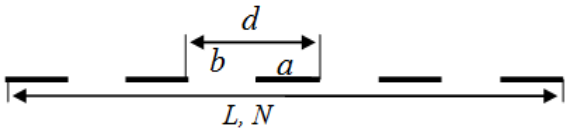


Рис. 4.5

Традиційно такі ґратки виготовляють на спеціальних прецизійних верстатах, наносячи на плоску скляну пластинку алмазною голкою систему штрихів-подряпин, які слугують непрозорими проміжками між неушкодженими ділянками-щілинами. Використовують також ґратки інших типів, зокрема, відбивальні ґратки, в яких штрихи наносяться на непрозору дзеркальну поверхню. Існують і складніші ґратки, як наприклад, двовимірна ґратка, що являє систему малих отворів у вигляді правильної геометричної сітки. Прикладом природніх тривимірних ґраток є монокристали – речовини, атоми яких розташовані в строгому геометричному порядку і відіграють роль непрозорих перешкод для світла.

Оптичні характеристики простої ґратки визначаються трьома параметрами. Це:

- період $d = a + b$ – відстань між сусідніми щілинами, який виражають у мм або мкм. У техніці замість періоду часто використовують обернену величину $f = 1/d\,(1/\text{мм})$ – кількість штрихів на одиницю довжини ґратки;
- ширина щілини b або величина d/b;
- загальна кількість штрихів (щілин) N, яка визначається розміром ґратки: $N = L/d = Lf$.

Дифракційна картина від ґратки. Розглянемо фраунгоферову дифракцію на одновимірній плоскій ґратці із заданими параметрами d, b, N. Схема спостереження та сама, що й для однієї щілини (рис. 4.3): на ґратку нормально падає паралельний пучок монохроматичних променів з довжиною хвилі λ, і дифракційна картина спостерігається на екрані у фокальній площині збірної лінзи $O_б$.

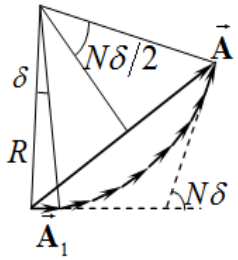


Рис. 4.6

Згідно з принципом Гюйгенса-Френеля інтенсивність світла у кожній точці екрана визначається результатом інтерференції N когерентних пучків від усіх щілин ґратки. При цьому з міркувань симетрії зрозуміло, що амплітуди коливань A_1 в усіх пучках однакові, як також і зсуви фаз δ коливань від сусідніх щілин. Тому підрахунок зручно робити за допомогою векторної діаграми рис. 4.6, зобразивши у вигляді ламаної послідовність N векторів складових коливань і замикаючий вектор, який визначає результуючі коливання в точці спостереження. Позаяк усі складові амплітуди і зсуви фаз однакові, вказана ламана є частиною правильного многокутника, і результуючий вектор $\vec{A}$ є хордою дуги кола з деяким радіусом R і центральним кутом рівним $N\delta$. Тому результуюча амплітуда

$$A = 2R\sin\left(N\delta/2\right).$$

З іншого боку (див. рис.), радіус кола $R = A_1/2\sin(\delta/2)$, тож результуюча амплітуда

$$A = A_1 \frac{\sin(N\delta/2)}{\sin(\delta/2)}. \qquad (4.7)$$

Взявши до уваги, що інтенсивність $I \sim A^2$, із виразу (4.7) для результуючої інтенсивності отримаємо:

$$I = I_1 \frac{\sin^2(N\delta/2)}{\sin^2(\delta/2)}, \qquad (4.8)$$

де I_1 – інтенсивність, яка створюється однією щілиною, N – кількість щілин у ґратці, δ – зсув фаз коливань, що збуджуються двома сусідніми щілинами в даній точці екрана. Величина δ визначається формулою (2.9) через різницю ходу сусідніх променів $\Delta = d\sin\varphi$ (див. рис. 4.7):

$$\delta = \frac{2\pi}{\lambda}\Delta = \frac{2\pi d\sin\varphi}{\lambda}. \qquad (4.9)$$

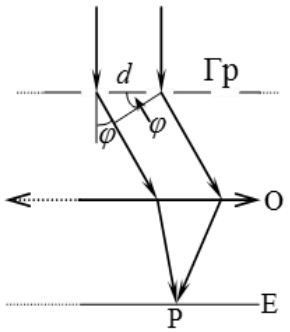


Рис. 4.7

Отже

$$I = I_1 \frac{\sin^2\left((N\pi d\sin\varphi)/\lambda\right)}{\sin^2\left((\pi d\sin\varphi)/\lambda\right)}. \qquad (4.10)$$

Згідно з цією формулою, дифракційна картина на екрані складається з великої кількості смуг (дифракційних максимумів) різної інтенсивності. Зокрема, за умови $\delta = \pm m\cdot 2\pi$ спостерігаються так звані ***головні максимуми***, умовою утворення котрих є:

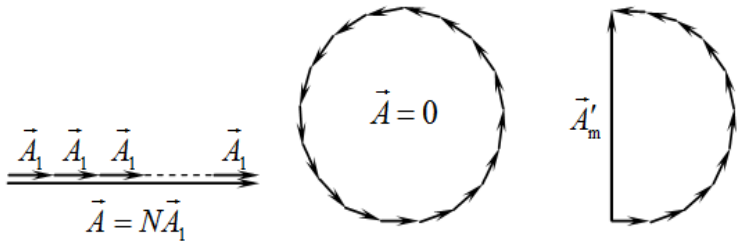
$$d\sin\varphi = \pm m\lambda, \quad m = 0, 1, 2, \ldots, \qquad (4.11)$$

де число m називається ***порядком*** головного максимуму.

Утворення головних максимумів легко пояснити. При $\delta = \pm m\cdot 2\pi$ всі складові вектори на діаграмі мають однаковий напрям і лежать на одній прямій (рис. 4.8а). Тому результуюча амплітуда та інтенсивність мають най-більші можливі значення:

$$A = NA_1 \quad \Rightarrow \quad I = N^2 I_1. \qquad (4.12)$$

У реальній ґратці $N >> 1$, тому головні максимуми є дуже інтенсивними.



а) **б)** **в)**

Рис. 4.8

При виконанні умови (4.11) у дифракційній формулі (4.10) і чисельник і знаменник дорівнюють нулю. Якщо ж на нуль обертається тільки чисельник, то утворюються дифракційні мінімуми (темні смуги) з інтенсивністю $I = 0$. Так буде, коли $\left((N\pi d\sin\varphi)/\lambda\right) = \pm m'\pi$, і складові вектори на діаграмі утворюють замкнену ламану (рис. 4.8б). Відповідно, умовою дифракційних мінімумів ґратки є:

$$d\sin\varphi = \pm\frac{m'}{N}\lambda. \qquad (4.13)$$

При цьому цілі числа m' не можуть бути кратними до кількості щілин ґратки N, бо в такому разі умова (4.13) вироджується в умову головних максимумів (4.11). Для ілюстрації нижче наведені можливі значення m' для перших чотирьох порядків m головного максимуму:

$$\begin{array}{lllll} m = 0, & 1, & 2, & 3, & \ldots \\ m' = 1, 2, \ldots, N-1, & N+1, \ldots, 2N-1, & 2N+1, \ldots, 3N-1, & 3N+1, \ldots & \end{array} \qquad (4.14)$$

звідки видно, що між сусідніми головними максимумами є $N-1$ мінімумів.

Це означає, що між головними максимумами утворюється ще й по $N-2$ так званих *побічних* дифракційних максимумів. На векторній діаграмі при утворенні побічного максимуму складові вектори згортаються в «напівзамкнений» многокутник, як зображено на рис. 4.8в.

Розрахунки показують, що побічні максимуми, як і при дифракції на щілині, мають незначну інтенсивність, яка спадає при віддаленні від головного максимуму. Тому реально вони роздільно не спостерігаються й утворюють лише слабку фонову освітленість поміж головними максимумами. Тож майже вся світлова енергія, що проходить крізь ґратку, зосереджується в області головних максимумів, інтенсивність яких визначається формулою (4.12): $I = N^2 I_1$. При цьому високоякісна оптична дифракційна ґратка може мати період $d \sim 1$ мкм і довжину $L \sim 10$ см. За таких умов кількість щілин сягає величини порядку 10^5, а інтенсивність головних максимумів $I \sim 10^{10} I_1$. Таке колосальне підсилення має просте пояснення. В реальній ґратці кількість побічних максимумів між двома головними практично дорівнює кількості щілин ґратки *N*, тобто є дуже великою. Тому побічні максимуми займають майже всю площу екрана, і головні максимуми виявляються надзвичайно вузькими. Отож майже вся світлова енергія, що потрапляє на екран, концентрується на дуже малій поверхні, чим і пояснюється надзвичайно висока інтенсивність головних максимумів.

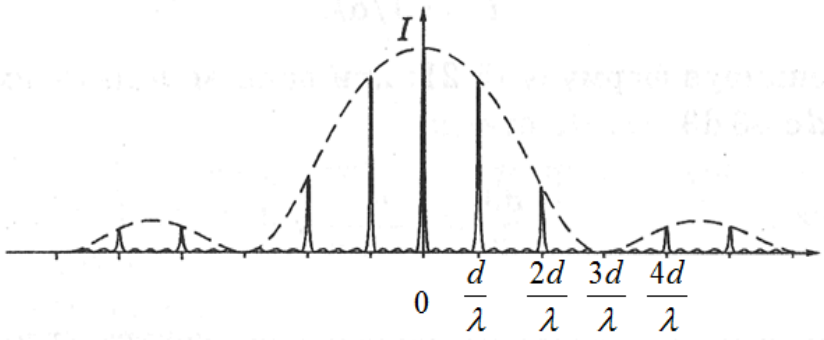


Рис. 4.9

Ще одна особливість ґратки полягає в тому, що інтенсивність I_1 у формулі (4.12), яка створюється однією щілиною, теж залежить від кута дифракції, відповідно до формули (4.3) і рис.4.5. Тому так само залежить від кута дифракції й інтенсивність головних максимумів, як це схематично показано на рис.4.9. Зокрема, при куті дифракції, для якого

$$\sin\varphi = \pm m\frac{\lambda}{b},$$

інтенсивність головного максимуму $I = 0$, навіть якщо цей кут задовольняє умову головного максимуму. Порядки m_0 максимумів, які випадають, залежать від періоду d та ширини щілини ґратки b, згідно з умовами:

$$\begin{aligned} d\sin\varphi &= m_0\lambda \\ b\sin\varphi &= m\lambda \end{aligned} \quad \Rightarrow \quad m_0 = m\frac{d}{b}, \quad m = 1, 2, \ldots$$

Слід зауважити, що позаяк побічні дифракційні максимуми щілини є слабкими, більш або менш інтенсивні головні максимуми спостерігаються лише в області між першими мінімумами щілини при $\sin\varphi = \pm\lambda/b$. Тому максимальний порядок $m_{\max}$ і кількість головних максимумів $N_{\max}$, які можна *реально* спостерігати, складають:

$$m_{\max} = \frac{d}{b} - 1, \quad N_{\max} = 2m_{\max} + 1 = 2\frac{d}{b} - 1.$$

(Зауважимо, що це число не збігається з формальною величиною, котра фігурує в текстах навчальних задач і визначається умовою $\varphi \le 90°$).

Окрім того, як і у випадку щілини, дифракційні смуги у ґратці спостерігаються лише коли її період потрапляє у визначений інтервал значень. Як випливає з умови (4.11), при $d \to \lambda$ кут $\varphi \to 90°$, і дифракційна картина відсутня, бо на екрані формально утворюється лише центральний максимум. Так само не спостерігається дифракція й при дуже великому періоді ґратки, бо кути дифракції стають такими малими, що максимуми зливаються, утворюючи одну освітлену смугу в центрі екрана. Тому реально дифракційна картина спостерігається, коли період ґратки складає десятки – сотні довжин хвилі світла, яким опромінюється ґратка.

Дифракційна ґратка як оптичний прилад. Згідно з умовою (4.11), кути дифракції, а отже й положення головних максимумів на екрані, залежать від довжини хвилі світла, що проходить крізь ґратку. Тому при опроміненні ґратки білим світлом, яке містить усі довжини хвиль в інтервалі $(0{,}4 \div 0{,}75)$ мкм, головні максимуми для різних довжин хвилі утворюються при різних кутах, тож і в різних точках екрана. Відповідно на екрані замість дуже вузьких максимумів спостерігаються ***дифракційні спектри*** – відносно широкі забарвлені смуги із стандартною послідовністю кольорів від фіолетового до червоного. При цьому розподіл інтенсивностей на екрані $I(x)$ буде відображати спектр, тобто залежність інтенсивності від довжини хвилі $I(\lambda)$ у пучку, що падає на ґратку. Наприклад, якщо на ґратку спрямувати світло з лінійчастим спектром, до якого входять лише окремі вузькі інтервали довжин хвиль (*спектральні лінії*), то на екрані буде спостерігатися система вузьких кольорових смуг, яскравість і положення котрих точно відображає інтенсивності та довжини хвиль у падаючому пучку. Виняток становить лише центральний максимум, який лишається вузьким і не забарвленим, оскільки при $m = 0$ і $\varphi = 0$ умова (4.11) виконується для всіх довжин хвилі.

Здатність розкладати падаюче світло в спектр робить ґратку цінним спектральним оптичним приладом, який використовується в науці і техніці для аналізу спектрального складу досліджуваного випромінювання та для отримання інтенсивних пучків монохроматичного світла із заданою довжиною хвилі.

Якість дифракційної ґратки як оптичного приладу визначається двома параметрами – ***дисперсією*** та ***роздільною здатністю***. Дисперсія показує, як сильно ґратка відхиляє промені з різними довжинами хвилі. Розрізняють кутову та лінійну дисперсію.

Кутова дисперсія визначається, як

$$D_\varphi = \frac{\mathrm{d}\varphi}{\mathrm{d}\lambda}\,(\text{рад/нм}), \qquad (4.15)$$

де $\mathrm{d}\varphi$ – кут між напрямками головних максимумів дифракції для двох променів із граничнс малою різницею довжин хвилі $\mathrm{d}\lambda$ (рис. 4.10).

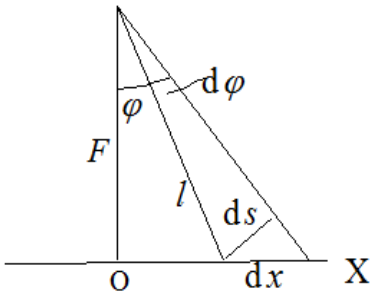


Рис. 4.10

Вираз D_φ можна отримати диференціюванням умови (4.11):

$$d\cos\varphi\cdot\mathrm{d}\varphi = m\,\mathrm{d}\lambda \quad \Rightarrow \quad D_\varphi = \frac{m}{d\cos\varphi}. \qquad (4.16)$$

Лінійна дисперсія визначається як

$$D_l = \frac{\mathrm{d}x}{\mathrm{d}\lambda}\,(\text{мм/нм}), \qquad (4.17)$$

де $\mathrm{d}x = \mathrm{d}s/\cos\varphi$ (див. рис. 4.10).

Якщо на шляху дифрагованого пучка поставити вузьку щілину шириною h, то крізь неї пройде випромінювання із спектральною шириною (інтервалом довжин хвилі) $\Delta\lambda = h/D_l$. Отже, чим більшу дисперсію має ґратка, тим вищий ступінь монохроматичності світла, яке можна отримати за її допомогою.

При невеликих кутах дифракції $dx \approx F\,\mathrm{d}\varphi$, отже

$$D_l = D_\varphi\cdot F = m\frac{F}{d}, \qquad (4.17a)$$

де F – фокусна відстань лінзи, що проектує дифракційний спектр на екран.

Другий параметр – ***роздільна здатність*** – визначає спроможність ґратки розрізняти спектральні лінії, тобто вузькі смуги з близькими довжинами хвиль.

Якщо на ґратку падає пучок, який складається з двох різних ліній, то можливі два крайні випадки. При значній відміні в довжинах хвилі ліній на екрані спостерігаються два окремі максимуми, як показано на рис. 4.11а. А при близьких довжинах хвилі лінії зливаються в одну нерозділену смугу, рис. 4.11б.

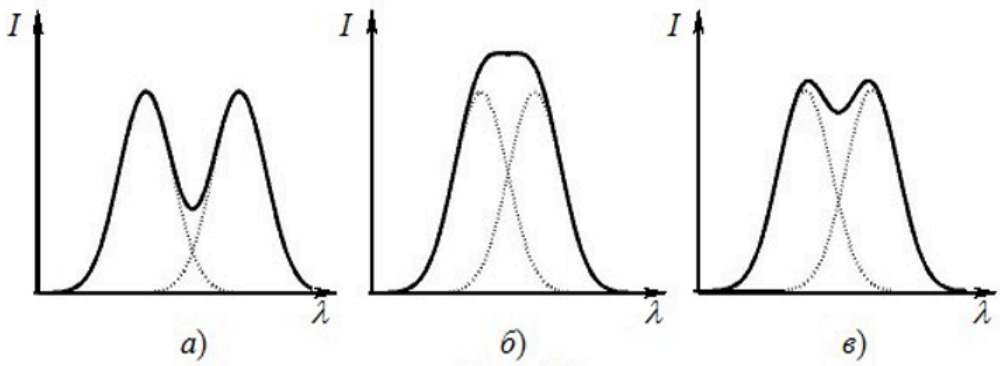


Рис. 4.11

Чи будуть розділені дві лінії, залежить як від довжини хвилі λ і спектральної відстані $\delta\lambda$ між ними, так і від самої ґратки. Тому кожна ґратка має свою *роздільну здатність*, яка визначається виразом

$$R = \frac{\lambda}{\delta\lambda}, \qquad (4.18)$$

де $\delta\lambda$ – найменша різниця довжин хвилі двох спектральних ліній, при якій вони ще є розділеними, а λ – довжина хвилі однієї з них.

Аби роздільна здатність R та інтервал $\delta\lambda$ були визначеними, прийнято наступний ***критерій Релея***:

> ***дві смуги однакової форми та інтенсивності вважаються вже розділеними, коли дифракційний максимум однієї потрапляє на перший мінімум іншої,***

як показано на рис. 4.11в. За таких умов між максимумами утворюється достатньо глибокий провал $I_{\min} \approx 0{,}8I_0$, і положення кожного піку лишається практично не спотвореним.

Згідно з критерієм Релея, головний максимум порядку m для спектральної лінії з довжиною хвилі $\lambda+\delta\lambda$ має співпадати з найближчим до нього дифракційним мінімумом того самого порядку m для лінії з довжиною хвилі λ. Отже, в умові (4.13) $m' = mN+1$ і, згідно з (4.11) і (4.13), маємо:

$$\begin{aligned} d\sin\varphi &= m\left(\lambda+\delta\lambda\right) \\ d\sin\varphi &= \left(m+\frac{1}{N}\right)\lambda \end{aligned} \quad \Rightarrow \quad \delta\lambda = \frac{\lambda}{mN}.$$

Таким чином, роздільна здатність ґратки (4.18) дорівнює

$$R = mN, \tag{4.19}$$

або

$$R = m\frac{L}{d}, \tag{4.19a}$$

де L – довжина ґратки, d – її період.

Слід зазначити, що позаяк ці формули встановлені для найкращих умов спостереження, вони визначать теоретичну межу роздільної здатності даної ґратки, а реальна величина R зазвичай є меншою.

Отримані формули показують, що обидва спектральні параметри ґратки – дисперсія та роздільна здатність – покращуються при збільшенні порядку спектра m. Але можливість використання спектрів високих порядків на практиці обмежується тим, що вони частково або повністю перекриваються. Таке перекривання настає тим раніше, чим ширший інтервал довжин хвиль входить у досліджуване світло. Так, для білого світла часткове перекриття спостерігається вже в спектрі другого порядку (див. п.4.6, Задача 4.5). Тому на практиці зазвичай використовуються спектри другого-третього порядку.

4. ДИФРАКЦІЯ РЕНТГЕНІВСЬКИХ ПРОМЕНІВ НА КРИСТАЛАХ

Дифракція спостерігається не тільки на системах паралельних вузьких щілин (одновимірних ґратках), а й на упорядкованих дво- та тривимірних сукупностях точкових перешкодах. Такими природніми системами є кристали – тіла, в яких атоми розміщуються у правильному геометричному порядку, утворюючи тривимірну ґратку, на якій може спостерігатися дифракція.

Відстань між атомами в кристалах, яка відіграє роль періоду ґратки, становить $d \sim 10^{-10}$ м, тоді як довжина хвилі в оптичному діапазоні $\lambda \approx \left(10^{-7} \div 10^{-6}\right)$ м, отже $\left(\lambda/d\right) \approx \left(10^{3} \div 10^{4}\right)$. Тому в оптичному діапазоні дифракція на кристалах не спостерігається. Але при опроміненні кристалів рентгенівськими променями, що мають значно меншу довжину хвилі, спостерігаються добре виражені дифракційні явища.

Розглянемо дифракцію при відбиванні рентгенівських променів у кристалах за спрощеним методом Вульфа-Бреґґа. В цьому методі кожен шар атомів розглядається як площина, що частково віддзеркалює падаючі на неї промені. Сукупність паралельних атомних площини утворює відбивну дифракційну ґратку з періодом d рівним відстані між такими площинами (рис. 4.12).

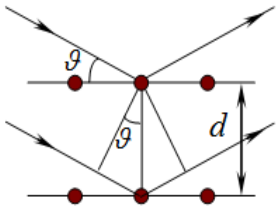


Рис. 4.12

Якщо на кристал спрямувати паралельний пучок рентгенівських променів із заданою довжиною хвилі λ, то при відбиванні від указаної множини атомних площин утвориться велика кількість паралельних когерентних променів, які, інтерферуючи, створять систему дифракційні максимумів. Ці максимуми будуть найінтенсивнішими при різниці ходу променів, відбитих від сусідніх атомних площин, $\Delta = \lambda$. Показник заломлення рентгенівських променів у кристалах майже не відрізняється від одиниці, тому величина Δ залежить тільки від відстані між атомними площинами та кута ϑ між ними та падаючими променями (***кута ковзання***). Як видно з рис. 4.12, $\Delta = 2l = 2d\sin\vartheta$, отже, дифракційні максимуми задовольняють наступну ***умову Вульфа-Бреґґа***:

$$2d\sin\vartheta = k\lambda, \qquad (4.20)$$

де k – порядок максимуму.

Дифракційні максимуми при рентгенівській дифракції на кристалах являють собою не послідовність смуг, а упорядковану сукупність цяток. Це зумовлено тим, що, дифракція відбувається не на лінійній, а на просторовій ґратці. При цьому в кожному кристалі можна виділити багато систем паралельних атомних площин із різною орієнтацією та щільністю розміщення атомів. Тому на фотопластинці або екрані монітора утворюється складна сукупність світних цяток, положення та яскравість яких відображає розташування атомів у кристалі. Тож на основі розроблених методик за параметрами дифракційної картини можна визначити тип кристалічної ґратки та відстані між атомами в даному кристалі. На цьому ґрунтується *рентгеноструктурний аналіз* – дуже інформативний і точний метод дослідження внутрішньої будови кристалів.

Наявність в кристалі мікроскопічних дефектів порушує правильне розташування атомів, що призводить до відповідних локальних спотворень дифракційної картини. Аналізуючи їх, можна виявляти та визначати характеристики мікроскопічних дефектів у досліджуваній речовині. На цьому ґрунтується *рентгенівська дефектоскопія*.

Рентгенівська дифракція використовується і в *рентгенівській спектроскопії*, де за допомогою кристала з відомими параметрами ґратки вимірюють спектральний склад (довжини хвиль) досліджуваного рентгенівського випромінювання так само, як в оптичній спектроскопії вимірюють спектральний склад світла за допомогою оптичної дифракційної ґратки з відомими параметрами.

5. КОНТРОЛЬНІ ЗАПИТАННЯ

1. У чому полягає явище дифракції та за яких умов воно спостерігається?
2. Сформулюйте та запишіть аналітичний вираз принципу Гюйгенса-Френеля. На яких міркуваннях він ґрунтується?
3. Що таке дифракція Фраунгофера та як її забезпечують при використанні звичайних (не лазерних) джерел випромінювання?
4. Чому при дифракції утворюється дифракційна картина (максимуми та мінімуми інтенсивності)?
5. Чому не спостерігаються дифракційні смуги при дуже малій та дуже великій ширині щілини? Яку оптимальну ширину повинна мати щілина для спостереження дифракції?
6. Які відмінності є між центральним і побічними дифракційними максимумами при опроміненні щілини білим світлом.
7. Що таке дифракційна ґратка? Чим дифракція на ґратці відрізняється від дифракції на одній щілині?
8. Зобразіть векторні діаграми для першого та другого побічного максимумів дифракційної ґратки та оцініть за їхньою допомогою відносну інтенсивність цих максимумів I/I_0 (I_0 – інтенсивність головного максимуму).
9. Зобразіть на рисунку приблизний вигляд дифракційної картини для умовної ґратки з $N = 4$ щілин в інтервалі $0 \le \sin\varphi \le (6d/\lambda)$ у двох випадках: 1) при відношенні періоду ґратки до ширини щілини $(d/b) = 3{,}5$ і 2) $(d/b) = 2{,}0$.
10. Відомо, що ґратка розкладає біле світло в спектр. Чи означає це, що промені з різною довжиною хвилі виходять із ґратки під різними кутами?
11. Що таке дисперсія ґратки, та яку властивість ґратки вона визначає? Як пов'язані кутова та лінійна дисперсія ґратки?
12. Що називається роздільною здатністю ґратки? Як формулюється критерій розділеності смуг (критерій Релея)?
13. Якими параметрами ґратки визначається її роздільна здатність?
14. Поясніть метод Вульфа-Бреґґа аналізу дифракції рентгенівських променів на кристалах і запишіть умову Вульфа-Бреґґа.
15. Чому на кристалах не спостерігається дифракція світла?

5.5. ПОЛЯРИЗАЦІЯ СВІТЛА

У лекції розглянуто такі питання:

1. Поляризоване та природнє світло
2. Отримання лінійно поляризованого світла
3. Еліптична та колова поляризація
4. Отримання еліптично та циркулярно поляризованого світла
5. Аналіз поляризованого світла
6. Контрольні запитання

1. Поляризоване та природне світло

Як розглянуті тут, так й інші хвильові явища є однаково властивими як для поздовжніх, так і для поперечних хвиль, таких як світло. Але між поздовжніми та поперечними хвилями є суттєва відміна: для поздовжньої хвилі не можна вказати площину, в якій відбуваються коливання, тоді як у поперечній вона є. У кожній точці ця площина, утворена напрямом поширення хвилі(напрямком хвильового вектора $\vec{k}$) та напрямком коливань світлового вектора $\vec{E}$ в даний момент часу. Тому напрям коливань у поперечній хвилі може бути впорядкованим.

Поперечні хвилі, в яких коливання певним чином упорядковані в просторі, називають поляризованими.

В оптиці поляризаційні ефекти засвідчили поперечність світлових хвиль ще до встановлення їхньої електромагнітної природи і занйшли застосування в науці та техніці.

Конкретним прикладом поляризованої хвилі може бути ідеальна плоска монохроматична світлова хвиля:

$$\vec{E}(\vec{r},t) = \vec{E}_0 \cos(\omega t - \vec{k}\cdot\vec{r}). \qquad (5.1)$$

У цій хвилі $\vec{E}_0 = const$ і $\vec{k} = const$, отже коливання відбуваються вздовж заданого напрямку (лінії вектора $\vec{E}_0$) у фіксовані й площині $(\vec{E}_0, \vec{k})$/ Тому така хвиля називається *лінійно-* або *плоско-* поляризованою. Відповідно, площина, в якій здійснюються коливання світлового вектора, називають ***площиною поляризації***. (*Примітка. Раніше площиною поляризації називали площину коливань магнітного вектора $\vec{H}$, і ця застаріла термінологія інколи ще зустрічається в старих книжках*).

Крім лінійної є ще два види поляризації, при яких площина коливань обертається навколо променя з кутовою швидкістю, рівною коловій частоті хвилі ω. Це ***еліптична*** та ***колова (циркулярна)*** поляризації. У першому випадку проекція кінця світлового вектора на перпендикулярну до променя площину описує еліпс, а в другому – коло. Залежно від напрямку обертання світлового вектора ці види поляризації поділяють на ***праву*** та ***ліву***. Якщо при спостереженні назустріч променю вектор $\vec{E}$ обертається за годинниковою стрілкою, то поляризацію називають правою, а при протилежному напрямі – лівою.

Будь-яка неперервна монохроматична хвиля типу (5.1) «автоматично» є поляризованою. Але випромінювання реальних джерел (окрім лазерів) зазвичай не виявляє жодних ознак поляризації. Причина в тому, що реальне світлове випромінювання складається з окремих «шматків» хвилі – цугів (див. Лекція 5.2, п.2.3), в яких не лише початкова фаза, а й напрям коливань світлового вектора є не прогнозованими. Через це в світловому пучку орієнтація площин поляризації дуже швидко і хаотично змінюються в часі. Таке світло не є поляризованим і називається ***природнім***. Проте лінійно поляризоване світло можна отримати з природного за допомогою спеціальних приладів, які називають поляризаторами.

Поляризатор незалежно від принципу дії та конструкції виділяє світлові коливання, що відбуваються в одній площині – площині поляризатора, і в той чи інший спосіб відсіює коливання, що відбуваються у перпендикулярній площині . Тому при падінні на поляризатор природнього світла на виході утворюється світловий пучок, що поляризований у площині поляризатора.

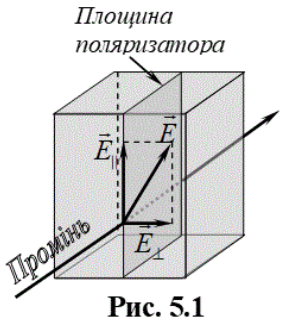


Рис. 5.1

В ідеальному поляризаторі (досконалому приладі, що дає повну поляризацію і зовсім не поглинає світло) інтенсивність поляризованого пучка на виході дорівнює половині інтенсивності пучка природнього світла на вході. Це легко довести. Світловий вектор $\vec{E}_{\text{пр}}$ у падаючому природньому пучку можна розкласти на паралельну та перпендикулярну до площини поляризатора складові, як показано на рис. 5.1:

$$\vec{E}_{\text{пр}} = \vec{E}_{\parallel} + \vec{E}_{\perp}.$$

Поляризатор пропускає тільки паралельну складову світлового вектора, отже у поляризованому промені на виході поляризатора миттєва напруженість поля

$$E = E_{\parallel} = E_{\text{пр}} \cos\alpha,$$

де кут α визначає миттєвий напрям вектора $\vec{E}$ природнього світлі на вхлді і є випадковою швидкозмінною величиною.

Інтенсивність світла I є прямо пропорційною середній величині квадрата світлового вектора: $I \sim \langle E^2 \rangle = E_{\text{пр}}^2 \langle \cos^2\alpha \rangle$. Отже,

$$I = I_{\text{пр}} \langle \cos^2\alpha \rangle.$$

У природному пучку напрям коливань світлового вектора дуже швидко змінюється, і всі значення α є рівноймовірними. Тому $\langle \cos^2\alpha \rangle = \dfrac{1}{2}$, і для інтенсивності на виході поляризатора маємо:

$$I = \frac{1}{2} I_{\text{пр}}. \qquad (5.2)$$

Якщо на поляризатор спрямувати не природній, а лінійно поляризований промінь інтенсивності I_0, площина поляризації якого складає заданий кут $\alpha = \text{const}$ з площиною поляризатора, то на виході отримаємо промінь, поляризований вже в *площині поляризатора*, і з інтенсивністю

$$I = I_0 \cos^2\alpha. \qquad (5.3)$$

Цей вираз називається ***законом Малюса***. Завдяки йому поляризатор можна використовувати не тільки для отримання, а й для аналізу поляризованого світла, тобто для визначення орієнтації площини коливань та інших характеристик поляризованої світлової хвилі.

На практиці доводиться мати справу також із ***частково поляризованим*** світлом, яке можна розглядати як суміш лінійно поляризованого та природнього світла. Наближеність такого світла до поляризованого характеризують ***ступенем поляризації*** P:

$$P = \frac{I}{I_0} = \frac{I}{I + I_{\text{пр}}}, \qquad (5.4)$$

де I – інтенсивність поляризованої компоненти, I_0 – повна інтенсивність частково поляризованого пучка, $I_{\text{пр}}$ – інтенсивність природньої компоненти.

Ступінь поляризації на практиці можна визначити, спрямувавши частково поляризоване світло на поляризатор і обертаючи його навколо променя. При цьому від природньої компоненти на виході матимемо постійну інтенсивність $I_{\text{пр}}/2$, а от поляризована компонента при паралельній орієнтації площини коливань та площини поляризатора повністю пройде, а при перпендикулярній – повністю затримається. Тому в першому випадку на виході з поляризатора повна інтенсивність буде максимальною:

$$I_{max} = I + \frac{1}{2} I_{\text{пр}},$$

а в другому – мінімальною:

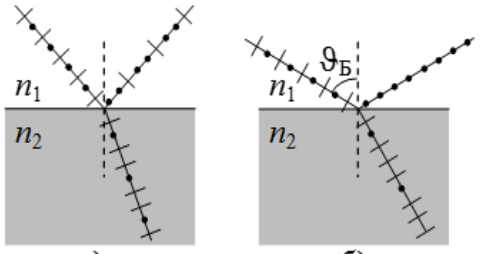
$$I_{min} = \frac{1}{2} I_{\text{пр}}.$$

Тож інтенсивність поляризованої компоненти можна виразити як $I = I_{max} - I_{min}$, а повну інтенсивність частково поляризованого пучка – як $I_0 = I_{max} + I_{min}$. Відтак ступінь поляризації визначається формулою:

$$P = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}. \qquad (5.4a)$$

2. ОТРИМАННЯ ЛІНІЙНО ПОЛЯРИЗОВАНАНОГО СВІТЛА

Поляризація при відбиванні. Із законів електродинаміки випливає, що при падінні світлової хвилі на межу поділу двох діелектриків розподіл енергії між відбитою та заломленою хвилями (див. Лекція 5.1) залежать не лише від кута падіння, а й від напрямку коливань світлового вектора. При цьому при падінні на межу поділу природнього світла в заломленому промені переважають коливання в площині падіння, а у відбитому – в перпендикулярній площині. (Це схематично показано на рис. 5.2а,



а) б)

Рис. 5.2

де коливання в площині падіння показані рисками, а в перпендикулярній площині – крапками). Отже, на загал і відбитий, і заломлений промені є частково поляризованими. Але тоді як для перпендикулярних коливань коефіцієнт відбивання із збільшенням кута падіння веь час зростає, для паралельних він спочатку зменшується і при певному куті падіння $\vartheta_Б$ обертається на нуль, а потім зростає. Тож при такому куті падіння відбитий промінь виявляється повністю поляризованим у площині перпендикулярній до площини падіння, (рис. 5.2б). Тому

> ***кут падіння світла на межу поділу діелектриків, при якому відбиті промені є повністю поляризованими, називається кутом повної поляризації або Брюстера*** $\vartheta_Б$.

При цьому заломлений промінь лишається частково поляризованим, але із максимально можливим ступенем поляризації.

У теорії доводиться, що кут $\vartheta_Б$ залежить від показників заломлення діелектриків і визначається наступною ***умовою або законом Брюстера***:

$$\operatorname{tg}\vartheta_Б = \frac{n_2}{n_1}. \qquad (5.5)$$

Із закону Брюстера та закону заломлення (Лекція 5.1) випливає, що при куті падіння $\vartheta_1 = \vartheta_Б$ кут заломлення задовольняє умову $\sin\vartheta_2 = \cos\vartheta_1$, $\vartheta_1 + \vartheta_2 = 90°$. Отже, при падінні променя на межу поділу середовищ під кутом Брюстера відбитий та заломлений промені є взаємно перпендикулярні, як показано на рис. 5.2б.

Таким чином, діелектрична пластинка, на яку спрямовано під кутом Брюстера пучок природнього світла, у відбитих променях працює як поляризатор. Але в такому поляризаторі інтенсивність поляризованого пучка недостатньо висока через малий коефіцієнт відбивання світла від діелектриків. До того ж напрям поляризованих променів при відбиванні різко змінюється, що створює великі незручності при конструюванні і застосуванні поляризаторів на відбитих променях. Тому на практиці використовують заломлені промені, котрі проходять крізь пластинку без відхилення. Правда, такі промені, хоча й інтенсивні, але поляризовані лише частково і мають невеликий ступінь поляризації. Проте цю ваду долають, спрямовуючи світло під кутом Брюстера не на одну пластинку, а на *стопу* – стос покладених одна на одну прозорих пластин. У такому разі при проходженні крізь кожну пластину заломлений пучок весь час збагачується на промені з коливаннями в площині падіння. Тож при достатній кількості пластин на виході стопи можна отримати інтенсивний заломлений пучок із великим ступенем поляризації.

Поляризація при двозаломленні. Стопа забезпечує досить високий ступінь поляризації, але на практиці зазвичай використовують поляризаційні пристрої, що ґрунтуються на явищі двозаломлення і мають кращі технічні характеристики. Явище двозаломлення полягає в роздвоєнні світлових променів у анізотропних кристалах. А саме, при падінні променя на поверхню такого кристала утворюється два поляризовані заломлені промені (рис. 5.3), котрі поширюються в кристалі з різними швидкостями. Ці промені називаються ***звичайним*** і ***незвичайним***. Звичайний промінь *о* задовольняє всі закони геометричної оптики, а незвичайний *е* – ні. Зокрема, він відхиляється навіть при нормальному падінні світла на поверхню кристала і не лежить у площині падіння.

У кожному анізотропному кристалі є напрям поширенні світла, вздовж якого двозаломлення не відбувається Такий напрям називають ***оптичною віссю*** кристала, а будь-яка площина, що проходить через оптичну вісь, називається ***головною площиною*** або ***головним перерізом*** кристала. Зазвичай використовують ту головну площину, в якій лежить промінь. При цьому незвичайний промінь *е* виявляється поляризованим у головній площині, а звичайний *о* – у площині перпендикулярній до головної

Звичайний і незвичайний промені на загал мають різні напрямки, за винятком поширення перпендикулярно до оптичної осі. У цьому випадку розділення променів у просторі не відбувається, але вони мають різні швидкості та показники заломлення $n_o \neq n_e$.

Кут розходження звичайного та незвичайного променів у кристалі є невеликим. Тому в кристалічних поляризаторах за допомогою відповідних технічних прийомів звичайний та незвичайний промені відокремлюють один від одного й використовують якийсь один. Наприклад, в поляризаторі, який називається призмою Ніколя або просто ***ніколем***, для відокремлення променів використовують явище повного

внутрішнього відбивання. Для цього ніколь виготовляють із двох однакових призм з ісландського шпату, які склеюють діагональними гранями за допомогою прозорої речовини – канадського бальзаму. Показник заломлення бальзаму n має проміжне значення між показниками заломлення шпату для звичайного n_o та незвичайного n_e променів, так що $n_e < n < n_o$. Враховуючи, що для звичайного променя бальзам є менш густим середовищем, призми вирізають так, що звичайний промінь повністю відбивається від склейки і далі поглинається зафарбованою бічною гранню призми. Відтак із ніколя виходять тільки промені поляризовані в головній площині призми.

Слід зазначити, що ***незвичайний*** промінь є таким тільки всередині кристала, а за межами кристала промені *e* і *o* відрізняються лише напрямком поляризації.

Зауважимо також, що є і кристали, котрі мають дві оптичні осі. Але в загальному курсі фізики вони не розглядаються.

Дихроїзм. Для отримання поляризованого світла використовують також *дихроїзм* – здатність деяких речовин пропускати світлові коливання одного напрямку і поглинати коливання перпендикулярного напрямку. Наприклад, кристалічна пластинка з турмаліну завтовшки 1 мм практично повністю поглинає промені із світловим вектором перпендикулярним до головної площини і пропускає тільки світло, поляризоване у головній площині. Таку саму властивість мають дешеві і широко вживані ***поляроїди*** – штучно виготовлені дихроїчні колоїдні плівки.

3. ЕЛІПТИЧНА ТА КОЛОВА ПОЛЯРИЗАЦІЯ

Накладання лінійно поляризованих хвиль. Як було сказано раніше, крім лінійної існують ще два види поляризації – еліптична та колова, при яких світловий вектор $\vec{E}$ обертається навколо променя. Покажемо, що ці види поляризації можна розглядати як результат суперпозиції плоско поляризованих хвиль.

Нехай дві когерентні плоскі хвилі з амплітудами E_{01} та E_{02} і частотою ω, які поляризовані у взаємно перпендикулярних площинах XOZ і YOZ (рис. 5.4), поширюються «до нас», тобто в додатному напрямку осі OZ. Тоді в якійсь площині перпендикулярній осі OZ хвилі створюють коливання світлового вектора, котрі задамо рівняннями:

$$E_x = E_{01}\cos\omega t \quad \text{і} \quad E_y = E_{02}\cos(\omega t + \delta), \tag{5.6}$$

де δ – не залежно від часу різниця фаз.

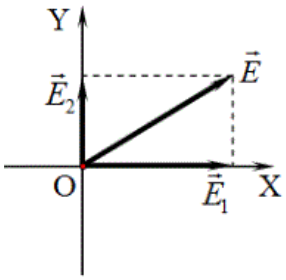


Рис. 5.4

«Координати» E_x та E_y визначають модуль і напрям результуючого вектора $\vec{E} = \vec{E}_1 + \vec{E}_2$ в будь-який момент часу. Тому, виключивши з рівнянь (5.6) параметр t (час), можна отримати траєкторію руху кінця вектора $\vec{E}$. Для цього друге рівняння (5.6) перепишемо як

$$\frac{E_y}{E_{02}} = (\cos\omega t\cos\delta - \sin\omega t\sin\delta) \quad \Rightarrow \quad \frac{E_y}{E_{02}} = \left(\cos\omega t\cos\delta - \sqrt{1-\cos^2\omega t}\cdot\sin\delta\right)$$

і підставимо вираз $\cos\omega t = E_x/E_{01}$ з першого рівняння. Тоді дістанемо:

$$\frac{E_y}{E_{02}} = \frac{E_x}{E_{01}}\cos\delta - \sin\delta\sqrt{1-\left(\frac{E_x}{E_{01}}\right)^2}.$$

Після піднесення до квадрату і нескладних перетворень отримаємо:

$$\frac{E_x^2}{E_{01}^2} + \frac{E_y^2}{E_{02}^2} - 2\frac{E_xE_y}{E_{01}E_{02}}\cos\delta = \sin^2\delta. \tag{5.7}$$

Це рівняння є загальним рівнянням еліпса. Отже, при накладанні вказаних лінійно поляризованих хвиль в загальному випадку утворюється результуюча хвиля, в якій світловий вектор обертається навколо променя і змінюється так, що проекція його кінця описує еліпс, тобто виникає еліптично поляризоване світло.

Конкретні параметри поляризації залежать від амплітуд E_{01}, E_{02} і різниці фаз δ хвиль, що накладаються. Зокрема, залежно від знаку $\sin\delta$ поляризація може бути правою або ліво. У цьому легко переконатися, знайшовши з (5.6) похідну по часу:

$$(\mathrm{d}E_x/\mathrm{d}t) = -\omega E_{01}\sin\omega t, \quad (\mathrm{d}E_y/\mathrm{d}t) = -\omega E_{02}\sin(\omega t + \delta).$$

При $t = 0$ $(\mathrm{d}E_x/\mathrm{d}t) = 0$ і $(\mathrm{d}E_y/\mathrm{d}t) = -\omega E_{02}\sin\delta$. Тож при спостереженні проти ходу променя у випадку $0 < \delta < \pi$ вектор $\vec{E}$ обертається за годинниковою стрілкою і світло є право поляризованим. Відповідно, при $-\pi < \delta < 0$ поляризація ліва.

Окремі випадки. Величина різниці фаз та співвідношення амплітуд визначають орієнтацію еліпса (5.7) і його ексцентриситет («сплющеність»). При цьому є декілька важливих окремих випадків.

1) $\delta = \pm\pi/2$. У цьому випадку рівняння (5.7) надуває вигляду

$$\frac{E_x^2}{E_{01}^2} + \frac{E_y^2}{E_{02}^2} = 1,$$

тобто осі еліпса орієнтовані вздовж осей координат. Якщо $E_{01} > E_{02}$, то еліпс сплющений уздовж осі OY, а при $E_{01} < E_{02}$ – уздовж осі OX.

2) За умови $E_{01} = E_{02} = E_0$ рівняння (5.7) перетворюється на рівняння кола

$$E_x^2 + E_y^2 = E_0^2.$$

Отже якщо накладаються когерентні плоско поляризовані хвилі із зсувом фаз у чверть періоду та однаковою амплітудою, то можна отримати хвилю, що поляризована по колу в той або інший бік.

3) $\delta = 0,\ \pm\,\pi$. У цьому випадку з рівняння (5.7) випливає, що

$$E_y = \pm\frac{E_{02}}{E_{01}}E_x,$$

і еліпс вироджується у відрізок прямої. Тож кінець вектора $\vec{E}$ здійснює коливання вздовж фіксованого напрямку і першій і третій (при $\delta = 0$) або другій та четвертій чвертях (при $\delta = \pm\pi$. Напрям коливань визначається амплітудами E_{01} та E_{02}, зокрема, при $E_{01} = E_{02}$ він складає кут 45° з осями координат. Отже, вибираючи потрібне співвідношення E_{01}/E_{02}, можна отримати бажану орієнтацію площини поляризації результуючої хвилі.

4. ОТРИМАННЯ ЕЛІПТИЧНО ТА ЦИРКУЛЯРНО ПОЛЯРИЗОВАНОГО СВІТЛА

На практиці поляризоване по еліпсу та колу світло можна отримати, пропускаючи лінійно поляризоване світло крізь двозаломлюючу кристалічну пластинку.

Хід променів у кристалічній пластинці. Якщо з одновісного кристала вирізати паралельно до оптичної осі OO' пластинку товщиною h (рис. 5.5а) і спрямувати на неї по нормалі промінь лінійно поляризованого монохроматичного світла, то звичайний та незвичайний промені не будуть розділятися, але поширюватимуться з різними швидкостями і матимуть різні показники заломлення n_o і n_e. Тому вони пройдуть у пластинці різні оптичні шляхи $L_o = n_o h$ та $L_e = n_e h$ і на виході матимуть різницю ходу



Рис. 5.5 *а)*

$$\Delta = (n_e - n_o)h$$

та за умовами когерентності (див. Лекцію 5.2) – незалежну від часу різницю фаз

$$\delta = \frac{2\pi}{\lambda}\Delta = \frac{2\pi}{\lambda}(n_e - n_o)h. \qquad (5.8)$$

У такому разі з пластинки вийде поляризоване світло, що описується загальним рівнянням (5.7). Характеристики поляризації залежать від знаку та величини δ, які при заданій товщині пластинки визначаються показниками заломлення променів *o* і *e*. Тому всі кристали поділяють на два типи – *додатні*, якщо $n_o > n_e$, та *від'ємні*, якщо $n_o < n_e$.

Таким чином, вибираючи необхідну товщину пластинки заданого типу, можна отримати бажаний характер поляризації. Розглянемо деякі важливі окремі випадки.

Пластинка у чверть хвилі (***пластинка*** $\lambda/4$). У пункті 5.3. було показано, що при накладанні хвиль із різницею фаз $\delta = \pm\pi/2$ утворюється світло поляризоване або по еліпсу з приведеними осями, або по колу. Цю умову формально можна записати як $\delta = m\pi/2$, де $m = 1, 3, 5, \ldots$. Підставивши її в (5.8) отримаємо необхідну тлвщину пластинки:

$$m\frac{\pi}{2} = \frac{2\pi}{\lambda}|n_e - n_o|h \quad \Rightarrow \quad h = \frac{m}{|n_e - n_o|}\cdot\frac{\lambda}{4}, \;\; m = 1, 3, 5, \ldots \qquad (5.9)$$

Таку пластинку умовно називають *пластиною у чверть хвилі* (*пластинкою* $\lambda/4$.

Уздовж якої з координатних осей буде сплющений еліпс залежить від співвідношення амплітуд звичайного та незвичайного променів E_{01}/E_{02}, що визначається кутом φ між площиною поляризації падаючого променя та оптичною віссю пластинки (рис. 5.5*б*).

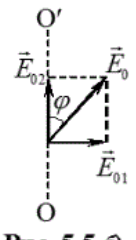


Рис. 5.5 *б)*

Напрям поляризації залежить від типу пластинки та її товщини й змінюється на протилежний при переході у формулі (5.9) від одного значення m до наступного. Наприклад, у пластинці з додатнього кристала при $m = 1$ поляризація буде правою, а при $m = 3$ – лівою.

В окремому випадку $\varphi = 45^\circ$ амплітуди звичайного та незвичайного променів будуть однаковими, тому з пластинки вийде світло поляризоване по колу. Отримання такого світла є одним з основних призначень пластинки у чверть хвилі.

Слід відзначити, що пластинка $\lambda/4$ «працює» і в зворотньому напрямі. Якщо спрямувати на неї циркулярно поляризований промінь, то він розділиться на два лінійно поляризовані у взаємно перпендикулярних напрямках промені *o* та *e*, зсунуті по фазі на $\pi/2$. Всередині пластинки вони отримують додаткову різницю фаз $\pm\pi/2$, відтак на виході матимемо лінійно поляризоване світло, площина поляризації якого складає кут 45° з оптичною віссю пластинки.

Пластинка у півхвилі (***пластинка*** $\lambda/4$). Це так само, як говорилося раніше, вирізана з кристала пластинка з товщиною

$$h = \frac{m}{n_e - n_o} \cdot \frac{\lambda}{2}, \quad m = 1, 3, 5, \ldots \tag{5.10}$$

Різниця фаз звичайного та незвичайного променів на виході з пластинки тепер складає $\delta = m\pi$. Тож світло, що виходить, лишається лінійно поляризованим, але площина поляризації в загальному випадку повертається на кут 2φ (подвоєний кут між площиною поляризації падаючого променя та оптичною віссю пластинки) і розміщується симетрично відносно оптичної осі (рис. 5.6*б*). Тому, спрямувавши на пластинку $\lambda/4$ світло, що лінійно поляризоване під кутом $\varphi = 45^\circ$ до оптичної осі, можна «повернути» площину поляризації на 90°, як це буває необхідно на практиці.

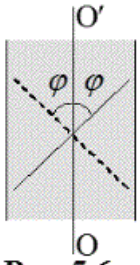


Рис. 5.6

5. АНАЛІЗ ПОЛЯРИЗОВАНОГО СВІТЛА

На практиці часом виникає необхідність аналізувати світло, що використовується, тобто з'ясовувати, яким воно є: поляризованим чи природнім, а якщо поляризованим, то як саме. Розглянемо, як це можна зробити в різних конкретних випадках.

Плоско поляризоване світло. Переконатися, що світло є лінійно поляризованим і визначити орієнтацію площини поляризації можна за допомогою поляризатора з відомим положенням його площини пропускання. Для цього пучок досліджуваного світла спрямовують на поляризатор і, обертаючи його навколо променів, вимірюють інтенсивність світла на виході. Якщо світло плоско поляризоване, то при положенні площини поляризатора, що співпадає з площиною поляризації променя, інтенсивність буде максимальною, а при повороті поляризатора на 90° зменшуватиметься практично до нуля, відповідно до закону Малюса (5.3).

Природнє та циркулярно поляризоване світло. Інтенсивність на виході при обертанні поляризатора навколо променя лишається сталою не лише для природного світла, а й для світла поляризованого по колу. Але ці випадки можна розрізнити, якщо перед поляризатором на шляху падаючого променя встановити пластинку в чверть хвилі. В такому разі, якщо досліджуване світло поляризоване по колу, то пластинка перетворить його на плоскополяризоване, і це буде легко встановити, обертаючи поляризатор навколо променя. Природнє ж світло, пройшовши крізь пластинку $\lambda/4$, так і лишиться природнім. Тому при обертанні поляризатора навколо променя інтенсивність на виході змінюватися не буде.

Еліптично-поляризоване та частково-поляризоване світло. Якщо спрямувати на поляризатор еліптично-поляризоване або частково-поляризоване світло, то при обертанні навколо променя кутова залежність інтенсивності на виході буде виглядати однаково. А саме, при деякому положенні площини поляризатора вона буде максимальною I_{max}, а при повороті на 90° стане мінімальною (але не нульовою) I_{min}. Визначити, яким саме є дане світло, можна так само, як у випадку циркулярно поляризованого та природнього світла. Якщо пластинка $\lambda/4$, встановлена на шляху падаючого променя, перетворює його на плоскополяризований, то досліджуване світло є еліптично-поляризованим. Натомість у випадку частково поляризованого світла пластинка не вплине на залежність інтенсивності на виході від кута повороту поляризатора навколо променя. При цьому у випадку еліптично поляризованого світла положення максимального пропускання поляризатора вкаже орієнтацію більшої осі еліпса, а інтенсивності I_{max} та I_{min} визначать відношення його осей $E_{01}/E_{02} = \sqrt{I_{max}/I_{min}}$. У випадку частково поляризованого світла аналогічно визначається положення площини поляризації поляризованої компоненти та, за формулою (5.4а) – ступінь поляризації досліджуваного частково поляризованого світла.

Напрям поляризації. Пластинка $\lambda/4$ дозволяє встановити також правою чи лівою є поляризація даного еліптично- або циркулярно-поляризованого світла. Для цього на ній вказано так званий β-***напрям*** - напрям коливань у промені (*o* чи *e*), що поширюється з більшою швидкістю, тож має менший показник заломлення. Розглянемо на конкретному прикладі, як це робиться.

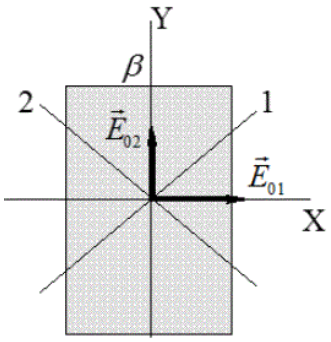


Рис. 5.7

Нехай на чвертьхвильову пластинку падає в напрямку «до нас» циркулярно поляризований промінь. Виберемо для зручності координатні осі, як показано на рис. 5.7. Тоді промінь поширюється в додатньому напрямку осі OZ і є суперпозицією двох променів $\vec{E}_1$ і $\vec{E}_2$ однакової амплітуди, які поляризовані, як розглядалось у пункті 5.3: $\vec{E}_1$ у площині XOZ, а $\vec{E}_2$ в площині YOZ. При цьому, якщо поляризація права (світловий вектор $\vec{E} = \vec{E}_1 + \vec{E}_2$ обертається для нас за годинниковою стрілкою), то складова $\vec{E}_2$ у падаючому промені випереджає за фазою складову $\vec{E}_1$ на $\pi/2$. Але при поширенні в пластинці $\lambda/4$ промінь $\vec{E}_2$ через більшу швидкість поширення і менший показник заломлення відстане від променя $\vec{E}_1$ на $\pi/2$ (див. формулу (5.8)). Тому кінцева різниця фаз буде нульовою, і з пластинки вийде промінь поляризований лінійно під кутом $\varphi_{\text{пр}} = 45^\circ$ до осі OX (вздовж лінії 1 на рис. 5.7). При лівій поляризації складова $\vec{E}_2$ у падаючому промені за фазою відстає від $\vec{E}_1$. Тож на виході різниця фаз буде рівна π, і промінь на виході з пластинки буде поляризований під кутом $\varphi_{\text{лв}} = 135^\circ$ до осі OX (вздовж лінії 2 на рис. 5.7). Величину кута φ, отже й напрям колової поляризації падаючого на пластинку $\lambda/4$ променя, визначають за допомогою поляризатора встановленого за пластинкою на шляху променя, що пройшов.

Напрям поляризації еліптично-поляризованого світла визначається так само. Але в цьому випадку $\varphi_{\text{пр}} \neq 45^\circ$ і $\varphi_{\text{лв}} \neq 135^\circ$.

6. КОНТРОЛЬНІ ЗАПИТАННЯ

1. Які хвилі звуться поляризованими? Чи можуть бути поляризованими звукові хвилі?
2. Назвіть та опишіть існуючі види поляризації. Який між ними є зв'язок?
3. У чому полягає закон Малюса?
4. Що таке поляризатор? На яких фізичних явищах ґрунтується робота поляризаторів?
5. Яку частку інтенсивності I/I_0 пропускає ідеальний поляризатор при падінні на нього: а) природнього, б) плоско поляризованого та в) циркулярно поляризованого світла?
6. Що таке ступінь поляризації світла та як його можна визначити?
7. Що таке кут Брюстера та чому він дорівнює?
8. Зобразіть на одному графіку приблизну залежність ступеня поляризації P відбитого та заломленого променів від кута падіння ϑ природнього світла з повітря на поверхню прозорого діелектрика.
9. Зобразіть на одному графіку приблизну залежність коефіцієнта відбивання плоско поляризованого світла від кута падіння на поверхню прозорого діелектрика, якщо площина поляризації падаючого променя: а) є перпендикулярною до площини падіння і б) співпадає з площиною падіння.
10. Покажіть приблизний графік залежності коефіцієнта відбивання R від кута падіння ϑ при падінні природнього світла з повітря на пластину прозорого діелектрика.
11. У чому полягає та коли спостерігається явище двозаломлення світла? Чим відрізняються звичайний і незвичайний промені та що у них є спільного?
12. Що таке оптична вісь та головна площина кристала? Поясніть, чи можуть не лежати в головній площині двозаломлюючого кристала: а) падаючий промінь; б) заломлені промені?
13. Чи завжди в двозаломлюючому кристалі існують звичайний і незвичайний промені?
14. Опишіть поширення світла в двозаломлюючому кристалі, коли падаючий промінь напрямлений: а) вздовж оптичної осі кристала і б) перпендикулярно до неї.
15. Поясніть принцип отримання світла поляризованого по еліпсу або по колу. Які вимоги мають задовольняти лінійно поляризовані промені, аби з них можна було отримати циркулярно поляризований промінь?
16. Як можна визначити положення площини поляризації лінійно поляризованого світла?
17. Як можна розрізнити природнє та циркулярно поляризоване світло? Еліптично та частково поляризоване світло?
18. Як можна встановити напрям обертання світлового вектора в еліптично-поляризованому промені, або в промені, що поляризований по колу?

5.6. ПОШИРЕННЯ СВІТЛА В РЕЧОВИНІ

У цій лекції розглядаються наступні питання:

1. Дисперсія світла

2. Групова швидкість

3. Поглинання та розсіювання світла

1. ДИСПЕРСІЯ СВІТЛА

У Лекції 5.1 (п. 1.1) було показано, що строго монохроматична хвиля поширюється із швидкістю, яка співпадає з фазовою швидкістю

$$v = \frac{\omega}{k}.$$

При цьому частота монохроматичної хвилі ω визначається тільки джерелом хвилі, а швидкість v – тільки властивостями середовища, в якому вона поширюється. Отже, на перший погляд фазова швидкість хвилі не повинна залежати від її частоти. Але насправді властивості середовища, які визначають швидкість поширення хвилі, залежать від її частоти. Тому між фазовою швидкістю та частотою хвилі існує опосередкований функціональний зв'язок: $v = v(\omega)$.

Залежність фазової швидкості хвилі від її частоти називається дисперсією.

У випадку світлових хвиль на практиці замість швидкості v використовують показник заломлення $n = (c/v)$, а замість частоти – довжину хвилі λ. Відтак, явище *дисперсії світла* розглядають як залежність показника заломлення (оптичної густини) середовища від довжини світлової хвилі: $n = n(\lambda)$. Кількісною мірою цієї залежності є величина $\mathrm{d}n/\mathrm{d}\lambda$, яка називається *дисперсією речовини*.

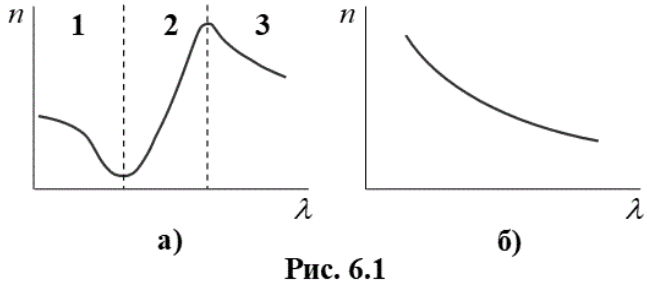


Рис. 6.1

Закон дисперсії світла $n(\lambda)$ є складним і не може бути виражений точними формулами. Для ізотропних прозорих речовин характер залежності $n(\lambda)$ якісно показаний на рис. 6.1а кривою, що складається з трьох характерних ділянок. Ділянки 1 і 3, на яких $(\mathrm{d}n/\mathrm{d}\lambda) < 0$ називаються областями *нормальної дисперсії*, а ділянка 2, де $(\mathrm{d}n/\mathrm{d}\lambda) > 0$ – областю *аномальної дисперсії*. В безбарвних речовинах ділянка аномальної дисперсії потрапляє в ультрафіолетову область спектра. Тому у видимому діапазоні спостерігається тільки ділянка 3, як показано на рис. 6.1б. *Примітка*. Кількісно зміни n є набагато меншими, ніж показано на рисунку для виразності. Для більшості речовин вони складають всього біля 1%.

Завдяки дисперсії світлові промені з різними довжинами хвиль по-різному заломлюються на межі поділу середовищ з різними показниками заломлення – фіолетові сильніше, а червоні слабше. Саме це дозволило Ньютону, пропускаючи вузький пучок сонячного світла крізь призму, розкласти його у спектр і довести, що біле світло є складним і містить промені різних кольорів.

Елементарне пояснення дисперсії світла можна отримати на основі електромагнітної теорії, згідно з якою показник заломлення прозорої речовини визначається її діелектричною проникністю:

$$n = \sqrt{\varepsilon}.$$

Величина ε, згідно з класичною теорією діелектриків, залежить від зміщення електронів в атомах під дією електричного поля. У випадку світла це зміщення визначається амплітудою A вимушених коливань електронів під дією електричного поля світлової хвилі $E = E_0 \cos \omega t$. При цьому теорія доводить, що

$$\varepsilon = 1 + \eta A. \qquad (6.1)$$

де константа η залежить від індивідуальних властивостей діелектрика.

У класичній теорії вважається, що електрон зв'язаний з ядром атома квазіпружно, так, немов прикріплений пружинкою. Тому в речовині без поглинання електрони атомів розглядаються як гармонічні осцилятори з певною частотою ω_0. Отже, згідно з теорією вимушених коливань, для амплітуди A за відсутності загасання можна записати

$$A \sim \frac{1}{\omega_0^2 - \omega^2}$$

і подати вираз (6.1) у вигляді:

$$\varepsilon(\omega) = 1 + \frac{a}{\omega_0^2 - \omega^2}, \qquad (6.2)$$

де a – коефіцієнт пропорційності, який, згідно з розрахунками, визначається зарядом і масою електрона та концентрацією електронів у даному діелектрику.

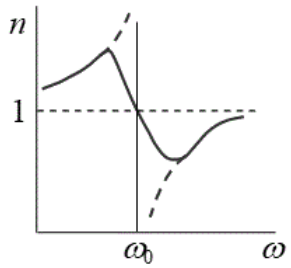


Рис. 6.2

Записана функція $\varepsilon(\omega)$ при $\omega = \omega_0$ має розрив $\varepsilon \to \pm\infty$, показаний штриховими відрізками на рис. 6.2. Але цей розрив є фіктивним. Він зумовлений тим, що у виразі (6.2) не враховано загасання коливань, яке завжди існує. Тому насправді залежність $\varepsilon(\omega)$ виглядає, як суцільна крива на рис. 6.2. Подібною є й залежність $n(\omega) = \sqrt{\varepsilon(\omega)}$, що відповідає вигляду дисперсійної кривої $n(\lambda)$ на рис. 6.1. При цьому ділянка аномальної дисперсії потрапляє в окіл власної (резонансної) частоти ω_0, де амплітуда вимушених коливань електронів у речовині різко зростає за рахунок енергії світлової хвилі. (В атомі зв'язок різних груп електронів з ядром не однаковий. Тому може існувати декілька резонансних частот і відрізків типу рис. 6.1 на дисперсійній кривій). Тому області аномальної дисперсії відповідають смугам поглинання світла у речовині. Цей теоретичний висновок також підтверджується на досліді.

2. ГРУПОВА ШВИДКІСТЬ

Заслуговує на увагу ще одна особливість дисперсійної кривої рис. 6.1а. А саме, на ділянці 1 нормальної дисперсії показник заломлення $n < 1$, і $v > c$. Виходить, що фазова швидкість світла в речовині може бути більшою, ніж у вакуумі. Це може видатися парадоксальним, оскільки в теорії відносності встановлено, що величина $c = 3 \cdot 10^8$ м/с є граничною, тобто максимальною можливою швидкістю. Але фазова швидкість не є фізичною величиною по суті. Вона визначає не рух фізичних об'єктів (частинок речовини або поля), а переміщення математичних хвильових поверхонь, тобто точок із заданим значенням формальної характеристики хвилі – фази φ. Тому можливість ситуації, коли $v > c$, не суперечить теорії відносності. Але така можливість вочевидь указує на те, що фазова швидкість v не визначає реальну швидкість хвилі, тобто швидкість поширення коливань і перенесення енергії хвилі. Виняток становлять лишень електромагнітні хвилі у вакуумі, де вони при будь-якій частоті поширюються із швидкістю $v = c$.

Реальна швидкість поширення хвилі називається *груповою швидкістю* і в середовищі відрізняється від фазової швидкості. Ця відміна зумовлена тим, що в реальній хвилі завжди присутні коливання з різними частотами, які в середовищі з дисперсією мають різні фазові швидкості. Аби переконатись у цьому, спочатку розглянемо гранично спрощений приклад поширення «реальної» хвилі $E = E_1 + E_2$, що складається всього з двох гармонічних хвиль з близькими частотами ω та $\omega + \Delta\omega$ та хвильовими числами k і $k + \Delta k$, так що $\Delta\omega \ll \omega$ і $\Delta k \ll k$. Ці хвилі описуються рівняннями:

$$E_1 = E_0 \cos(\omega t - kx),$$

$$E_2 = E_0 \cos((\omega + \Delta\omega)t - (k + \Delta k)x).$$

Додавши ці рівняння матимемо:

$$E = 2E_0 \cos\left(\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x\right) \cos\left(\left(\omega + \frac{\Delta\omega}{2}\right)t - \left(k + \frac{\Delta k}{2}\right)x\right).$$

Через те, що $\Delta\omega \ll \omega$ і $\Delta k \ll k$, в останньому множнику можна знехтувати величинами $\Delta\omega/2$ і $\Delta k/2$ і згорнуто записати:

$$E(x,t) = A(x,t)\cos(\omega t - kx), \tag{6.3}$$

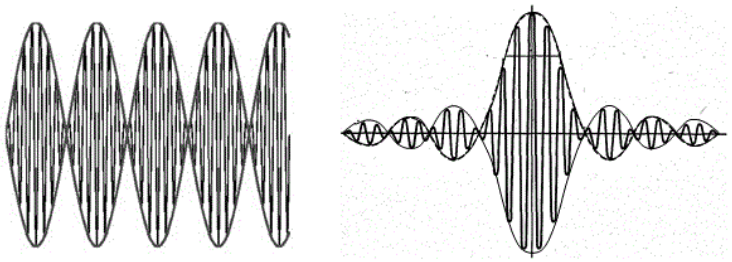
де функція

$$A(x,t) = 2E_0 \cos\left(\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x\right) \tag{6.3a}$$

змінюється дуже повільною порівняно з $\cos(\omega t - kx)$. Тому рівняння (6.3) можна трактувати як рівняння хвилі з частотою ω, хвильовим числом k і фазовою швидкістю $v = (\omega/k)$, в якій амплітуда $A = |A(x,t)|$ є різною в різні моменти часу і в різних точках простору. Зокрема, в будь-який заданий момент часу $t = \tau$ значення амплітуди розподіляються вздовж напрямку поширення хвилі періодично за законом

$$A(x) = 2E_0 \left|\cos\left(\frac{\Delta k}{2}x - \alpha\right)\right|, \quad \alpha = (\Delta\omega\tau/2). \tag{6.4}$$

Відповідно, існує множина точок із максимальною амплітудою $A_0 = 2E_0$, розміщених на відстані X одна від одної, і така сама множина точок, в яких $A = 0$, тобто коливання відсутні. Позаяк $\Delta k \ll k$, відстані між такими точками $X = (2\pi/\Delta k)$ є набагато більші за довжину хвилі $\lambda = 2\pi/k$. Отже, складена хвиля (6.3) являє собою послідовність окремих згустків з великої кількості коливань, як показано на рис. 6.3а.



а) б)

Рис. 6.3

З плином часу координати точок із заданим, приміром максимальним, значенням амплітуди змінюються, тож хвиля рухається. Швидкість цього руху можна знайти поклавши в рівнянні (6.3а) $A = 2E_0$. Тоді:

$$\left|\cos\left(\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x\right)\right| = 1 \quad \Rightarrow \quad x = \pm\frac{2n\pi}{\Delta k} + \frac{\Delta\omega}{\Delta k}t, \tag{6.5}$$

або

$$x = x_0 + ut,$$

де x_0 – початкова координата якоїсь вибраної точки з максимальною амплітудою коливань, а u – швидкість її руху.

Отже, утворена з двох гармонічних складових «реальна» хвиля являє собою послідовність згустків коливань, які рухаються зі швидкістю

$$u = \frac{\Delta\omega}{\Delta k}. \tag{6.6}$$

За відсутності дисперсії ця величина співпадає з фазовою швидкістю v. Справді, якщо фазова швидкість не залежить від частоти та довжини хвилі, то

$$x = \frac{\omega}{k} = const \quad \Rightarrow \quad \Delta\omega = v\Delta k \quad \Rightarrow \quad u = \frac{\Delta\omega}{\Delta k} = v.$$

Але в середовищах з дисперсією, а такими є всі реальні середовища, швидкість поширення хвилі u відрізняється від фазової швидкості: $u \neq v$.

У дійсності в будь-якій реальній хвилі присутні коливання не з двома, а з усіма можливими частотами в деякому інтервалі $\omega \pm \Delta\omega/2$ і хвильовими числами в інтервалі $k \pm k/2$ в околі центральної частоти ω та центрального хвильового числа k. У такому разі розрахунки показують, що за умови $\Delta\omega \ll \omega$ і $\Delta k \ll k$ при накладанні всіх складових розподіл амплітуд у результуючій хвилі має вигляд рис. 6.3б, тобто утворюється практично один згусток («пакет») коливань. Тому

хвильове утворення, що має малу спектральну ширину, називається хвильовим пакетом або групою хвиль.

Розрахунок показує, що для групи хвиль (хвильового пакета) формула (6.6) трансформується у наступний вираз швидкості руху пакета, або ***групової швидкості*** *\(u\):*

$$u = \frac{\mathrm{d}\omega}{\mathrm{d}k}. \tag{6.7}$$

Групова та фазова швидкості пов'язані між собою. Відповідне співвідношення легко встановити, взявши до уваги, що $\omega = vk$ і в середовищі з дисперсією фазова швидкість є функцією довжини хвилі, тож і хвильового числа k: $v = v(k)$. Тоді з виразу (6.7) отримаємо:

$$u = \frac{\mathrm{d}(vk)}{\mathrm{d}k} \quad \Rightarrow \quad u = v + k\frac{\mathrm{d}v}{\mathrm{d}k}. \tag{6.7а}$$

Це співвідношення можна переписати і через довжину хвилі , зробивши заміни:

$$k = \frac{2\pi}{\lambda} \quad \text{і} \quad \frac{\mathrm{d}v}{\mathrm{d}k} = \frac{\mathrm{d}v}{\mathrm{d}\lambda} \cdot \frac{\mathrm{d}\lambda}{\mathrm{d}k} = -\lambda\frac{\mathrm{d}v}{\mathrm{d}\lambda}.$$

Тоді

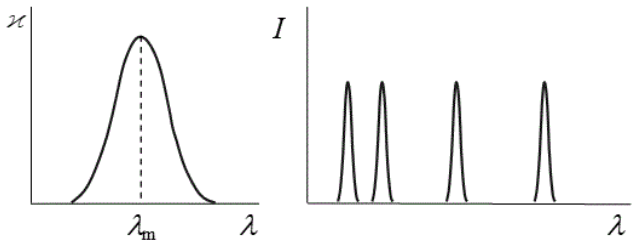
$$u = v - \lambda\frac{\mathrm{d}v}{\mathrm{d}\lambda}. \tag{6.7б}$$

Цей вираз називається ***формулою (співвідношенням) Релея***.

Із виразів (6.7*а*) і (6.7*б*) випливає, що й у реальному хвильовому пакеті за відсутності дисперсії групова швидкість співпадає з фазовою, а в середовищі з дисперсією – відрізняється від неї. При цьому конкретне співвідношення між величинами u і v залежить від значення похідної $\mathrm{d}v/\mathrm{d}k$ або $\mathrm{d}v/\mathrm{d}\lambda$ у виразах (6.7*а*) і (6.7*б*), тобто від закону дисперсії. Зокрема, в областях нормальної дисперсії $(\mathrm{d}v/\mathrm{d}\lambda) > 0$, тому групова швидкість $u < v$ і завжди виявляється меншою за c. Але в області аномальної дисперсії, де $\mathrm{d}v/\mathrm{d}\lambda < 0$, маємо $u > v$. Через це на ділянці, де $n < 1$, виходить $u > c$, що неможливо. Цей парадокс має наступне пояснення. Закон дисперсії в реальних середовищах не буває лінійним. З цієї причини похідна $\mathrm{d}n/\mathrm{d}\lambda$ і величина u мають різні значення при різних λ. Тому під груповою швидкістю (швидкістю пакета) розуміють величину (6.7*а*) чи (6.7*б*), яка відповідає центральним значенням k та λ і визначає рух центра пакета – точки з максимальною амплітудою коливань. Крім того, позаяк окремі гармонічні складові мають різні фазові швидкості, пакет поступово розпливається і в решті решт зникає. Така деградація (*«редукція»*) пакета відбувається тим швидше, чим більша дисперсія. Тому, взагалі, про хвильовий пакет і групову швидкість можна говорити лише при малій дисперсії, коли пакет існує достатньо довго. Але аномальна дисперсія спостерігається у вузькому інтервалі довжин хвилі і тому є дуже сильною. За таких умов поняття хвильового пакета та групової швидкості, тож і формули (6.7*а*) і (6.7*б*), втрачають зміст.

3. ПОГЛИНАННЯ ТА РОЗСІЮВАННЯ СВІТЛА

Поглинання світла. Уявлення про прозоре в усьому оптичному діапазоні середовище є абстракцією. Насправді кожна прозора речовина поглинає світло з певними довжинами хвиль, які в рідинах і твердих тілах утворюють досить широкі смуги поглинання (рис. 6.4а). Якщо смуги поглинання розташовані в ультрафіолетовій або інфрачервоній областях спектра, речовина пропускає всі довжини хвиль видимого світла і виглядає безбарвною. Але коли смуги поглинання потрапляють у певну частину видимого діапазону, при розгляданні на просвіт речовина виявляється забарвленою у доповнюючий колір. Приміром, якщо смуга поглинання даної речовини лежить у зеленій області спектра, то на просвіт вона виглядає пурпуровою (суміш червоного і синього кольорів). Окреме місце посідають одноатомні гази та пари атомарних речовин при невисокому тиску. Їхнє поглинання характеризується великою кількістю надзвичайно вузьких ($\sim 10^{-12}$ м) смуг, які називаються спектральними лініями, як умовно показано на рис.6.4б. Довжини хвиль та інтенсивності спектральних ліній є строго індивідуальними для кожного хімічного елемента, що дозволяє його ідентифікувати по спектрах. Так, наприклад, по спектру поглинання світла в атмосфері Сонця (сонячній короні) був відкритий хімічний елемент гелій («сонячний»).



а) **б)**

Рис. 6.4

При поглинанні частина світлової енергії переходить у внутрішню енергію речовини, тому при поширенні в речовині світловий пучок поступово послаблюється. Закон, за яким відбувається це послаблення, не важко встановити. Нехай в однорідному середовищі поширюється плоска світлова хвиля (паралельний пучок променів). Тоді спад інтенсивності $-\mathrm{d}I$ зумовлюється лише поглинанням і в елементарному шарі речовини є прямо пропорційним самій інтенсивності I та товщині шару $\mathrm{d}x$:

$$-\mathrm{d}I = \varkappa \mathrm{d}x,$$

де $\varkappa$ – коефіцієнт залежний від властивостей речовини.

Це рівняння легко інтегрується методом поділу змінних:

$$\frac{\mathrm{d}I}{I} = -\varkappa \mathrm{d}x \quad \Rightarrow \quad \int\limits_{I_0}^{I} \frac{\mathrm{d}I}{I} = -\varkappa x \quad \Rightarrow \quad I = I_0 e^{-\varkappa x}, \qquad (6.8)$$

де I_0 – інтенсивність у точці $x = 0$, а I – інтенсивність на відстані x у напрямку поширення світла.

Вираз (6.8), який називається ***законом Бугера***, свідчить, що інтенсивність світла в однорідному середовищі спадає експоненціально. Швидкість спадання визначається величиною

$$\varkappa = \frac{\mathrm{d}I}{I\mathrm{d}x},$$

що називається ***коефіцієнтом поглинання*** **і** показує, яка частка енергії поглинається даною речовиною в шарі одиничної товщини.

Зауважимо також, що у формулі (6.8) відстань x можна відраховувати від будь-якої точки x_0 всередині речовини. Зокрема, якщо на пластинку товщини d з коефіцієнтом поглинання $\varkappa$ падає плоска хвиля з інтенсивністю $I_)$, то на виході інтенсивність буде $I = I_0 e^{-\varkappa d}$.

При поширенні в речовині не паралельних променів, скажімо променів від точкового чи лінійного джерела, закон Бугера теж виконується, але у виразі (6.8) замість інтенсивностей фігурують потоки світлової енергії:

$$\Phi = \Phi_0 e^{-\varkappa x}. \qquad (6.8a)$$

Це пояснюється тим, що в сферичній чи циліндричній хвилі незалежно від поглинання інтенсивність спадає з відстанню від джерела через розподіл енергії випромінювання по все більшій поверхні.

Розсіювання світла. Як уже говорилося, згідно з класичною хвильовою оптикою, при поширенні світла в речовині за рахунок енергії світлової хвилі збуджуються вимушені коливання електронів у атомах. Частина цієї енергії поглинається, а решта повертається у вигляді вторинного випромінювання атомів. Але, хоча це випромінювання відбувається в усіх напрямах, в оптично однорідному середовищі ніякого «розпорошення» світлової енергії не спостерігається. Це пояснюється тим, що вторинні промені всіх напрямів, окрім напрямку падаючого променя, взаємно погашаються внаслідок інтерференції.

Інша картина спостерігається в середовищах, які містять оптичні неоднорідності. Такими, зокрема, є мутні середовища – рідини або гази із зваженими в них мікроскопічними частинками, як от задимлене повітря або вода, «підфарбована» невеликою кількістю молока. При проходженні світла в мутному середовищі на чужорідних частинках відбувається дифракція. Але через неупорядкованість розташування та хаотичний тепловий рух цих частинок дифракційні максимуми та мінімуми не утворюються, й інтенсивність дифрагованих променів майже рівномірно розподіляється по всіх напрямках. Така нерегулярна дифракція називається розсіюванням світла. Саме завдяки розсіюванню стають видимими світлові промені в мутному середовищі, як наприклад, сонячні промені в запиленому приміщенні.

Розсіювання світла спостерігається і в чистих рідинах і газах. Це пояснюється тепловим рухом молекул, тому в даному випадку говорять про молекулярне розсіювання. Через хаотичність теплового руху молекул в окремих точках виникають флуктуації – спонтанні відхилення концентрації молекул та інших характеристик середовища від середньостатистичних значень. Відповідно, створюються мікроскопічні неоднорідності, в яких показник заломлення відрізняється від середньої для всього середовища величини. На цих неоднорідностях і відбувається молекулярне розсіювання світла. Правда, цей ефект є незначним, і ми не бачимо розсіювання світла в оточуючому чистому повітрі. Але у товщі всієї атмосфери воно є помітним, через що ми й бачимо небо, а не чорноту Космосу.

Інтенсивність розсіяного світла залежить від його довжини хвилі λ, причому по-різному при різних розмірах неоднорідностей середовища. Відтак виділяють декілька видів розсіювання, одним з яких є *релеєвське* розсіювання, котре спостерігається при розмірах неоднорідностей $\sim 0,1\lambda$ і менше. За таких умов виконується закон Релея, згідно з яким інтенсивність розсіяного світла є обернено пропорційною четвертому степеню довжини хвилі:

$$I \sim \frac{1}{\lambda^4}.$$

Згідно із законом Релея, випромінювання із синьо-фіолетової частини спектра розсіюється сильніше, ніж із оранжево-червоної. Тому у випадку білого світла розсіяні промені є збагаченими на сині кольори, а промені, що проходять прямо, – на червоні. Цим, зокрема, пояснюється блакитний колір неба, який ми спостерігаємо, коли в око попадають розсіяні сонячні промені. Так само при сході та при заході Сонце видається червоним, бо в око потрапляють прямі промені.

Якщо розміри неоднорідностей суттєво перевищують довжину світлових хвиль, то інтенсивність розсіяного світла практично не залежить від λ. Цим пояснюється білий колір хмар, які складаються з крапельок пари достатньо великих порівняно з λ розмірів.

4. КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що таке дисперсія хвиль та у чому полягає її причина?
2. Покажіть хід променя білого світла крізь трикутну призму. Чим спектр, який дає призма, відрізняється від дифракційного спектра?
3. Зобразіть типовий вигляд дисперсійної кривої $n(\lambda)$ в широкому інтервалі довжин хвилі. Охарактеризуйте фазову швидкість світла в кожній з трьох характерних областей дисперсії.
4. Поясніть якісно механізм дисперсії світла (електромагнітних хвиль).
5. Що називається групою хвиль (хвильовим пакетом) та груповою швидкістю? Яким виразом визначається групова швидкість?
6. Чому групова швидкість відрізняється від фазової? Якими співвідношеннями вони пов'язані?
7. Як пояснити, чому на певній ділянці області аномальної дисперсії групова швидкість світла виходить більшою за граничну швидкість $c = 3 \cdot 10^8$ м/с?
8. Запишіть закон поглинання світла в однорідному середовищі. Якщо при проходженні паралельного пучка світла крізь шар якоїсь речовини на першій третині товщини інтенсивність пучка зменшується на η %, то на скільки відсотків змінюється інтенсивність на останній третині товщини шару?
9. Чому вузький промінь світла в кімнаті з чистим повітрям безпосередньо бачити на можна, а в накуреній кімнаті його добре видно?
10. Як пояснити блакитний колір неба та білий колір хмар, який ми спостерігаємо в сонячний день?
11. Як пояснити червоний колір сонця при сході та при заході Сонця?

ЕЛЕМЕНТИ КВАНТОВОЇ ФІЗИКИ

Зміст

ВСТУП

6.1. СВІТЛОВІ КВАНТИ
1. СВІТЛОВІ КВАНТИ
2. ФОТОЕФЕКТ
3. ГАЛЬМІВНЕ РЕНТГЕНІВСЬКЕ ВИПРОМІНЮВАННЯ
4. КОНТРОЛЬНІ ЗАПИТАННЯ

6.2. ФОТОНИ
1. ВЛАСТИВОСТІ ФОТОНІВ
2. ТИСК СВІТЛА
3. ЕФЕКТ КОМПТОНА
4. КОНТРОЛЬНІ ЗАПИТАННЯ

6.3. ХВИЛЬОВІ ВЛАСТИВОСТІ ЧАСТИНОК
1. ПОСТУЛАТИ БОРА
2. ХВИЛІ де БРОЙЛЯ
3. ХВИЛЬОВА ФУНКЦІЯ
4. ПРИНЦИП НЕВИЗНАЧЕНОСТІ
5. ПРИКЛАДИ ЗАСТОСУВАННЯ ПРИЦИПУ НЕВИЗНАЧЕНОСТІ
6. КОНТРОЛЬНІ ЗАПИТАННЯ

6.4. РІВНЯННЯ ШРЬОДІНҐЕРА
1. РІВНЯННЯ ШРЬОДІНҐЕРА
2. ЧАСТИНКА В ОДНОВИМІРНОМУ ПОТЕНЦІАЛЬНОМУ ЯЩИКУ
3. ЧАСТИНКА У ТРИВИМІРНОМУ ПОТЕНЦІАЛЬНОМУ ЯЩИКУ
4. ЛІНІЙНИЙ ГАРМОНІЧНИЙ ОСЦИЛЯТОР
5. ТУНЕЛЬНИЙ ЕФЕКТ
6. КОНТРОЛЬНІ ЗАПИТАННЯ

6.5. АТОМ ГІДРОҐЕНУ
1. КВАНТОВІ ЧИСЛА
2. ЕНЕРГЕТИЧНИЙ ТА ОПТИЧНИЙ СПЕКТРИ АТОМА ГІДРОҐГЕНУ
3. КВАНТУВАННЯ МОМЕНТУ ІМПУЛЬСУ
4. ВИРОДЖЕННЯ ЕНЕРГЕТИЧНИХ РІВНІВ
5. РОЗПОДІЛ ЕЛЕКТРОННОЇ ГУСТИНИ В АТОМІ ГІДРОҐЕНУ
6. ОРБІТАЛЬНИЙ МАГНІТНИЙ МОМЕНТ ЕЛЕКТРОНА
7. КОНТРОЛЬНІ ЗАПИТАННЯ

6.6. СПІН ЕЛЕКТРОНА
1. Атом гідроґену у зовнішньому магнітному полі
2. Неповнота рівняння Шрьодінґера
3. Спін електрона
4. Принцип Паулі. Оболонкова структура атомів
5. Характеристичні рентгенівські спектри
6. Контрольні запитання

ВСТУП

Розглянуті в попередньому розділі та інші оптичні явища, пов’язані з поширенням світла, знайшли пояснення на основі класичної електромагнітної теорії, яка трактує світло як електромагнітні хвилі. Але такі уявлення про світло є неповними. Це стало очевидним насамкінець 19-го і початок 20-го століть, після того, як на основі хвильової теорії не змогли пояснити низку явищ пов’язаних з процесами випромінювання та поглинання електромагнітної енергії атомами речовини. Так само на основі законів класичної фізики не вдалося пояснити незаперечно встановлену на досліді ядерну будову атома. Причина такої недостатності полягає в тому, що властивості мікроскопічних частинок - складових атома, і процеси на атомному і субатомному рівні підпорядковані законам не класичної, а ***квантової*** фізики. В наступних лекціях розглядаються деякі основні положення цієї науки.

6.1. Світлові кванти

6.2. Фотони

6.3. Хвилі властивості частинок

6.4. Рівняння Шрьодінгера

6.5. Атом Гідрогену

6.6. Спін електрона

6.1. СВІТЛОВІ КВАНТИ

У цій лекції представлені такі питання:

1. Світлові кванти
2. Фотоефект
3. Гальмівне рентгенівське випромінювання
4. Контрольні запитання

1. СВІТЛОВІ КВАНТИ

Протягом усієї історії становлення і розвитку оптики конкурували два погляди на природу світла – корпускулярний та хвильовий. Згідно з першим, якого, зокрема, дотримувався Ньютон, світло вважалося мікроскопічними частинками, що рухаються і взаємодіють з тілами за законами механіки. У другому підході воно розглядалося як хвильовий процес у гіпотетичному середовищі – ефірі, що заповнює увесь світовий простір і проникає в усі тіла.

Проте наприкінці вісімнадцятого – початку дев'ятнадцятого століття в оптиці поступово утвердилася хвильова концепція. А після відкриття Максвеллом електромагнітних хвиль і їх експериментального спостереження та вивчення властивостей Герцом здалося, що хвильова електромагнітна концепція світла перемогла остаточно. Зокрема, вона стала теоретичним підґрунтям *фізичної хвильової оптики*, яка дала вичерпне пояснення явищ пов'язаних із поширенням світла, таких як прямолінійність променів в однорідному середовищі, інтерференція, дифракція, поляризація та ін. Та хвильова концепція відображає не всі властивості електромагнітного випромінювання, зокрема, світла. Виявилося, що вона є принципово неспроможною пояснити ряд оптичних явищ, які безпосередньо пов'язані з процесами випускання електромагнітного випромінювання та його взаємодії з речовиною.

В цьому ряду осібно стоїть проблема теплового випромінювання, тобто світлового випромінювання, що випускається нагрітими тілами внаслідок теплового руху та взаємодії атомів. Попри всі зусилля, на основі хвильової теорій та законів термодинаміки теоретично пояснити всі властивості теплового випромінювання не вдалося. При цьому деякі результати виявилися взагалі позбавленими змісту. Наприклад, з класичної теорії випливає формально незаперечний висновок про те, що будь-яке тіло при будь-якій відмінній від абсолютного нуля температурі має випромінювати в простір нескінченно велику енергію. Це вказує на неспроможність самої хвильової концепції в теорії теплового випромінювання.

Вихід із тупикової ситуації знайшов один із творців сучасної фізики **Макс Планк, який в 1900 р. висловив гіпотезу** про те, що

> ***електромагнітне випромінювання випускається тілами не неперервно, а окремими порціями – квантами, ерегія котрих визначається частотою випромінювання згідно з формулою***

$$\varepsilon = h\nu, \tag{1.1}$$

Формула (1.1) називається формулою Планка, а величина $6,626 \cdot 10^{-34}$ Дж·с (джоуль-секунда) називається ***сталою Планка*** і є основною константою квантової фізики.

В теорії часто замість лінійної використовують колову частоту $\omega = 2\pi\nu$ і записують формулу Планка у вигляді

$$\varepsilon = \hbar\omega \tag{1.1a}$$

де константа $\hbar = (h/2\pi) = 1,054 \cdot 10^{-34}$ Дж·с теж називається ***сталою Планка*** (використовують також термін «квант дії»).

Енергію кванта можна виразити і через довжину хвилі випромінювання у вакуумі:

$$\varepsilon = \frac{2\pi\hbar c}{\lambda} = \frac{hc}{\lambda}. \qquad (1.1б)$$

Уявлення про кванти дозволило Планку отримати теоретичну формулу, з якої випливають усі відомі властивості теплового випромінювання. Згодом з'ясувалося, що квантовий механізм випромінювання дозволяє пояснити й інші явища, побудувати теорію яких класична фізика виявилася нездатною.

У цьому зв'язку варто зауважити, що уявлення про дискретність випромінювання використовується і в деяких задачах класичної оптики, де світловий потік трактується як послідовність цугів – окремих «шматків» хвилі малої тривалості. Але революційність гіпотези Планка полягає в тому, що енергія порції випромінювання – кванта визначається тільки частотою, незалежно від його тривалості та амплітуди поля, як у хвильовій теорії.

2. ФОТОЕФЕКТ

Одним із таких «непідвладних» класичній фізиці оптичних явищ є фотоелектричний ефект, коротко – фотоефект. Фотоефект полягає у вивільненні електронів речовини під дією світла – електромагнітного випромінювання оптичного діапазону. Розрізняють зовнішній і внутрішній фотоефект. Зовнішній фотоефект спостерігається в металах і полягає у вильоті електронів під дією світла за межі зразка. (*Примітка. Слід зауважити, що електрони вилітають з речовини також під дією рентгенівських і гама-променів, але в цьому випадку говорять не про фотоефект, а про екзоелектронну емісію).*

При внутрішньому фотоефекті, що спостерігається в діелектриках і напівпровідниках, електрони при опроміненні залишаються всередині речовини, але звільняються від зв'язку з ядрами і стають додатковими носіями струму, що зменшує електричний опір зразка. Тому внутрішній фотоефект називають іще *фоторезистивним* ефектом.

На явищі фотоефекту працюють різні фотоелектричні прилади: фотоелементи, фотоопори, фотодіоди, фототранзистори та ін.

У цій лекції розглядається зовнішній фотоефект, властивості якого найбільш прозоро відображають фізику явища.

Вольт-амперна характеристика фотоелемента. Основні властивості зовнішнього фотоефекту можна встановити, досліджуючи вольт-амперні характеристики вакуумного фотоелемента – двохелектродної електронної лампи, в якій передбачена можливість опромінювання катода світлом.

Вольт-амперною характеристикою (ВАХ) називається залежність струму від напруги в даному електричному приладі. Загальна схема вимірювання ВАХ показана на рис. 1.1. На фотоелемент подається напруга, величину та полярність якої можна змінювати за допомогою потенціометра П. Катод К із досліджуваного металу та анод А розміщені у вакуумному балоні Б з віконцем В, яке виготовляють із матеріалу, що майже без поглинання пропускає світло в широкому діапазоні довжин хвилі. Це дозволяє опромінювати катод світлом із заданою інтенсивністю та частотою. Електрони, звільнені при освітленні катода (*фотоелектрони*), під дією електричного поля рухаються до анода А, створюючи *фотострум*. Напруга між катодом і анодом і струм через фотоелемент реєструються приладами.

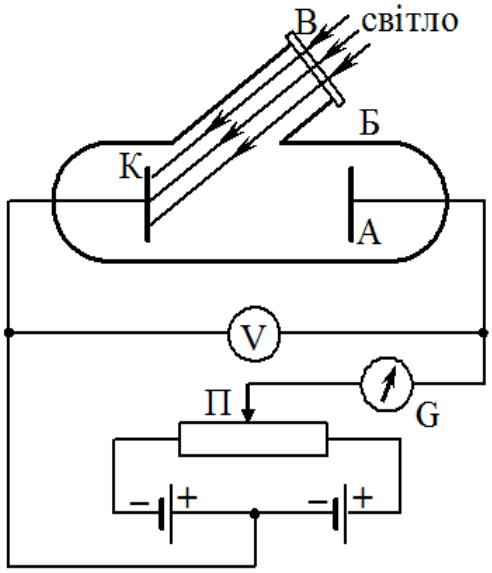


Рис.1.1

Типова ВАХ фотоелемента показана на рис. 1.2, на якому область $U>0$ відповідає прямій напрузі, коли анод має знак «+» і катод «–», а в області $U<0$ напруга є зворотною (на аноді «–» і на катоді «+»).

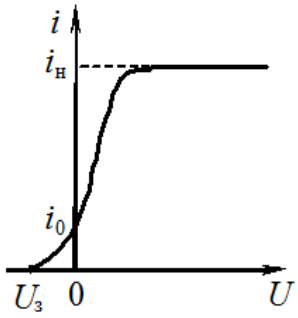


Рис. 1.2

Як видно з рис. 1.2, для ВАХ фотоелемента є характерними наступні особливості:

– Наявність деякого струму i_0 при $U=0$, тобто за відсутності між анодом і катодом електричного поля;

– Насичення: при збільшенні прямої напруги фотострум зростає тільки до певної величини $i_{\text{н}}$, яка називається ***струмом насичення***, і далі лишається практично незмінним;

– Для припинення струму на фотоелемент необхідно подати зворотню напругу, найменша величина якої $U_з$ називається ***затримуючою*** або ***запірною*** напругою фотоелемента.

Аби зрозуміти ці особливості ВАХ найперше візьмемо до уваги, що фотострум визначається величиною заряду, котрий переноситься фотоелектронами: $i = ne$, де n – кількість електронів, які потрапляють на анод за одиницю часу, e – елементарний заряд. Оскільки електрони вириваються з катода не електричним полем, а світлом, їхня кількість не залежить від напруги на фотоелементі. Це пояснює причину насичення і визначає величину струму насичення:

$$i_H = n_0 e, \qquad (1.2)$$

де n_0 – кількість електронів, яка вириває́ться з катода світлом за 1 с.

Таким чином,

вимірюючи струм насичення, ми можемо визначити потік фотоелектронів із металу при опроміненні його світлом.

Наявність струму i_0 говорить про те, що фотоелектрони не просто виходять на поверхню катода, а вилітають із деякою початковою швидкістю, так що можуть потрапити на анод «своїм ходом». Зазвичай $i_0 < i_i$, тож не всі вільні електрони, що вилетіли з катода, потрапляють на анод. Для того є дві причини. Перша полягає в тому, що фотоелектрони вилітають по всіх напрямках, тож лише частка їх рухається в напрямку анода і може потрапити на нього. Друга причина пов'язана з тим, що через виліт електронів поверхня катода під час опромінення світлом заряджається позитивно і притягає до себе електрони, що вилетіли. Тому під час опромінення катода одні електрони вилітають з нього, а інші повертаються назад. В результаті навколо позитивно зарядженого катода утворюється шар негативного заряду («електронна хмарина») та локальне електричне поле подібне до поля конденсатора. Це поле перешкоджає виходу електронів із катода та їхньому вільному рухові в напрямку анода. Тому за відсутності прискорюючої напруги на анод потрапляють лише ті фотоелектрони, що мають кінетичну енергію достатню для подолання цього поля.

При подаванні на фотоелемент прямої напруги для фотоелектронів створюється прискорююче електричне поле, що збільшує кількість електронів, які потрапляють на анод. Тому при підвищенні напруги фотострум поступово зростає аж до насичення, що настає, коли на анод потрапляють усі вирвані з катода електрони.

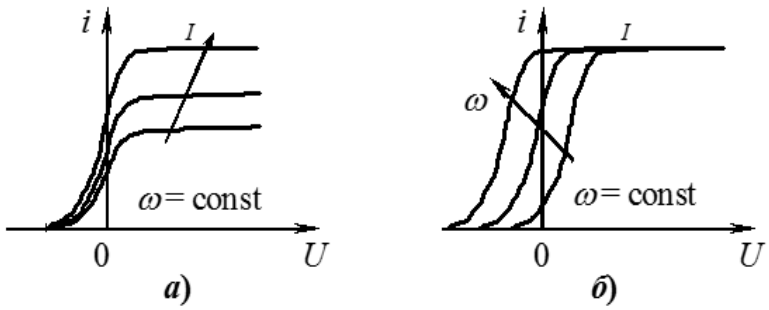
Аналогічно пояснюється й існування запірної напруги. При зворотній полярності напруги між анодом і катодом створюється гальмівне електричне поле, що створює додаткову перешкоду потраплянню фотоелектронів на анод. Тому при збільшенні зворотньої напруги струм поступово зменшується і зовсім зникає, коли жоден фотоелектрон буде не здатний потрапити на анод. Це станеться, коли найшвидші фотоелектрони на шляху від катода до анода витрачатимуть усю свою кінетичну енергію на роботу проти гальмівного поля. Отже, запірна напруга $U_з$ задовольняє умову:

$$eU_з = \frac{mv^2}{2}, \qquad (1.3)$$

де v – максимальна швидкість, з якою електрони виходять з металу. Таким чином,

вимірюючи запірну напругу, можна визначити максимальну швидкість і кінетичну енергію фотоелектронів.

Закономірності зовнішнього фотоефекту. Головні результати дослідження вольт-амперних характеристик фотоелементів за різних умов опромінювання якісно показані на рис. 1.3 і зводяться до наступного.



а) *б*)

Рис. 1.3

При опроміненні катода світлом однієї частоти, але різної інтенсивності (рис. 1.3*а*), запірна напруга фотоелемента лишається незмінною, тоді як струм насичення змінюється прямо пропорційно інтенсивності світла, що падає на катод:

$$i_{\text{і}} \sim I. \qquad (1.4)$$

У цьому полягає закон Столєтова, на якому ґрунтується електронна фотометрія, тобто техніка вимірювання світлових потоків за допомогою фотоефекту.

Але з цих результатів з урахуванням виразів (1.2) і (1.3) випливають і дві важливі фізичні закономірності явища фотоефекту.

А саме:

1. Кількість електронів, які вириваються з металу за одну секунду (потік фотоелектронів), є прямо пропорційною інтенсивності опромінюючого світла:

$$N \sim I. \qquad (1.4a)$$

2. Кінетична енергія, з якою електрони виходять з металу при заданій частоті опромінюючого світла, не залежить від його інтенсивності.

Запірна напруга фотоелемента залежить від частоти опромінюючого світла, як видно з рис. 1.3*б*, на якому показані ВАХ фотоелемента, виміряні при різних частотах і зведені для більшої наочності до однакової величини струму насичення. Ця залежність виявляється лінійною (рис. 1.4), з чого випливає ще одна закономірність фотоефекту:

3. Кінетична енергія, з якою електрони вилітають з металу, визначається тільки частотою опромінюючого світла і лінійно залежить від неї.

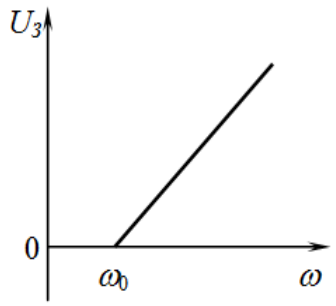


Рис. 1.4

Графік рис. 1.4 перетинає вісь абсцис у деякій точці $\omega_0 > 0$, що відповідає величині $U_з = 0$. При його продовженні далі ми потрапляємо в область позбавлених змісту значень $U_з < 0$, що означає неможливість фотоефекту. Дослід підтверджує такий висновок: при опроміненні фотоелемента світлом із частотою $\omega < \omega_0$ фотострум не з'являється ні за яких умов. Таким чином, існує певна порогова частота світла ω_0, при якій воно ще здатне вивирати електрони. Дослідження показали, що величина ω_0 є індивідуальною для кожного металу. Сказане складає ще одну закономірність фотоефекту:

4. Для кожного металу існує своя «червона межа» фотоефекту – найменша частота ω_0 (або найбільша довжина хвилі λ_0) опромінюючого світла, при якій ще можливий фотоефект.

Таким чином виривати електрони з металу може тільки світло, яке задовольняє умову:

$$\omega \geq \omega_0 \quad \text{або} \quad \lambda \leq \lambda_0. \qquad (1.5)$$

Варто зазначити, що для більшості матеріалів довжина хвилі λ_0 відповідає ультрафіолетовій або синьо-фіолетовій області спектра. А термін «червона» межа пояснюється тим, що найбільшим довжинам хвилі видимого світла відповідає червоний колір.

Ще одна важлива властивість фотоефекту полягає в тому, що

5. Фотоефект є безінерційним.

На це вказує той дослідний факт, що струм у колі фотоелемента з'являється практично одночасно із початком опромінення катода світлом.

Квантова теорія фотоефекту. Рівняння Ейнштейна. Перед розглядом квантової теорії фотоефекту зауважимо, що й класична хвильова теорія не заперечує фотоефекту як такого і пояснює його виходом електронів з металу під дією електромагнітного поля світлової хвилі. Але теоретичні прогнози хвильової теорії щодо конкретних властивостей фотоефекту знаходяться в різкому протиріччі з експериментальними фактами.

Як приклад, проаналізуємо на основі хвильового механізму кінетичну енергію фотоелектронів і, відповідно, запірну напругу фотоелемента.

В хвильовій моделі під дією світла вільні електрони металу починають здійснювати вимушені коливання, амплітуда яких поступово зростає до стаціонарної величини, що є прямо пропорційною амплітуді напруженості електричного поля світлової хвилі. При цьому, кожен електрон отримує від світла певну енергію W прямо пропорційну квадратові амплітуди вимушених коливань A. Своєю чергою величина A є прямо пропорційною амплітуді напруженості електричного поля, квадрат якої визначає інтенсивність світлової хвилі I. Відтак можна записати:

$$W = \alpha I,$$

де α – деякий коефіцієнт пропорційності.

За рахунок енергії коливань електрон може вийти за межі металу. Але при цьому він має витратити певну енергію, що зветься роботою виходу A і йде на подолання сил, які утримують електрони всередині металу. Тому кінетична енергія електрона K на виході з металу мала би дорівнювати

$$K = \alpha I - A.$$

З цього гіпотетичного виразу, який можна умовно назвати «класичним рівнянням фотоефекту», випливають два логічно незаперечні, проте помилкові прогнози:

1). Кінетична енергія фотоелектронів і, відповідно до (1.3), запірна напруга фотоелемента мали би лінійно залежати від інтенсивності опромінюючого світла, в той час як насправді ці величини зовсім не залежать від неї.

2). Для кожного металу мала б існувати порогова інтенсивність $I_0 = (A/\alpha)$, тобто найменша інтенсивність світлового потоку, при якій можливий вихід електронів із металу.

В дійсності нічого подібного немає. Якщо світло задовольняє умову (1.5), то воно вириває електрони при будь-якій, навіть гранично малій, інтенсивності. І, навпаки, якщо $\omega < \omega_0$, то навіть при опроміненні потужним світловим пучком жоден електрон не виходить з металу під дією світла.

Подібні суперечності виникають і при спробі на основі хвильових уявлень пояснити й інші закономірності фотоефекту. Приміром, розрахунки показують, що в моделі вимушених коливань навіть при інтенсивному опромінюванні для накопичення електроном достатньої для виходу з металу енергії знадобився би значний час, тоді як дослід свідчить про безінерційність фотоефекту.

Правильне пояснення фотоефекту дав А. Ейнштейн. Згідно із запропонованим ним *квантовим* механізмом, електрон вилітає з металу за рахунок енергії ε отриманої при поглинанні кванта опромінюючого світла. Частина цієї енергії витрачається електроном на роботу виходу з металу A, а решта складає його кінетичну енергію K. Отже,

$$\varepsilon = A + K. \qquad (1.6)$$

При фотоефекті енергія квантів $\varepsilon = \hbar\omega$ складає порядку $(1-10)$ еВ, і фотоелектрони є нерелятивістськими. Тому вираз (1.6) можна розгорнуто подати як

$$\hbar\omega = A + \frac{mv^2}{2}. \qquad (1.7)$$

Це рівняння називається ***рівнянням Ейнштейна для фотоефекту***.

Одразу зробимо уточнення. Робота виходу A за означенням дорівнює *найменшій* необхідній для виходу електрона з металу енергії. Тому величина v в рівнянні (1.7) означає *найбільшу* можливу швидкість фотоелектронів. Таку швидкість мають електрони, що виходять з гранично тонкого поверхневого шару металу. Електрони, що поглинають світло в глибших шарах, на шляху до поверхні стикаються з іншими електронами та іонами ґратки металу, відтак втрачають деяку енергію ще до виходу. Тож фотоелектрони виходять з металу з усіма можливими швидкостями від 0 до v. Цим пояснюється поступовість зменшення фотоструму в області зворотних напруг.

Рівняння Ейнштейна, яке по суті є виразом закону збереження енергії для процесу взаємодії кванта світла з електроном металу, пояснює всі закономірності зовнішнього фотоефекту.

По-перше, з рівняння (1.7) безпосередньо випливає, що кінетична енергія фотоелектронів і, згідно з (1.3), запірна напруга фотоелемента лінійно залежать від частоти опромінюючого світла. По-друге, очевидно, що фотоефект можливий тільки за умови $\hbar\omega \geq A$, отже, існує найменша частота світла (червона межа), при якій воно ще здатне вивирати електрони з металу. Ця частота дорівнює

$$\omega_0 = \frac{A}{\hbar}, \tag{1.8}$$

а довжина хвилі

$$\lambda_0 = \frac{2\pi\hbar c}{A} = \frac{hc}{A}. \tag{1.8a}$$

Таким чином, рівняння (1.7) не тільки доводить існування червоної межі фотоефекту, а й висвітлює її природу.

Інші закономірності фотоефекту формально в рівнянні Ейнштейна не відображені, але вони випливають із самого квантового механізму, на якому воно ґрунтується.

Наприклад, безінерційність фотоефекту пояснюється тим, що при поглинанні кванта електрон отримує необхідну для виходу енергію одномоментно, а не накопичує її поступово, як мало би бути згідно з хвильовою теорією.

Так само природньо пояснюється в квантовій теорії пряма пропорційна залежність потоку фотоелектронів та струму насичення фотоелемента від інтенсивності опромінюючого світла (закон Столєтова). Справді, згадаймо, що інтенсивність світла – то є енергія випромінювання, котра падає перпендикулярно на одиничну площадку за одиницю часу. З квантової точки зору вона визначається, як

$$I = \frac{n_{кв}\varepsilon}{S}, \tag{1.9}$$

де S –площа поперечного перерізу світлового пучка, що падає на катод, ε – енергія кванта, і $n_{кв}$ – кількість квантів, які падають на катод за 1 с. Отже,

інтенсивність опромінюючого світла є прямо пропорційною кількості квантів, що падають на катод фотоелемента за одиницю часу.

Ясна річ, не кожен падаючий на катод світловий квант поглинається вільним електроном, так само не кожен електрон, який поглинув квант, виходить з катода. Але очевидно, що існує визначена імовірність виходу електронів при опроміненні даного металу заданим світлом. Вона називається ***квантовим виходом*** фотоефекту η і визначається відношенням кількості фотоелектронів до кількості квантів, які падають на метал:

$$\eta = \frac{n_0}{n_{кв}}. \qquad (1.10)$$

Отже,

кількість фотоелектронів, які вилітають з катода, є прямо пропорційною кількості квантів, які падають на нього:

$$n_0 = \eta n_{кв}. \qquad (1.10a)$$

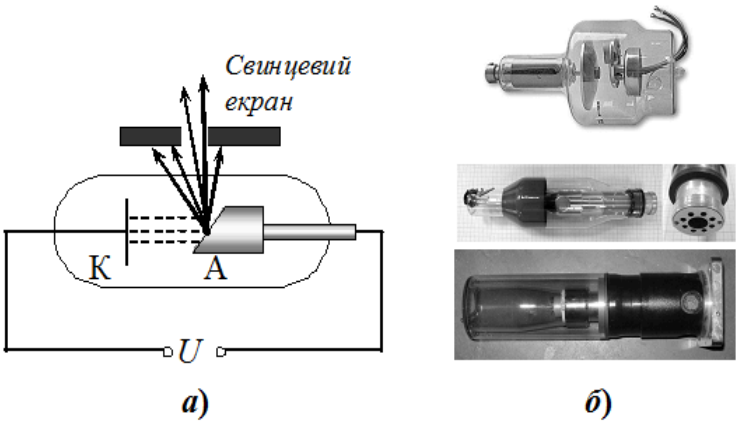
Звідси, враховуючи вирази (1.9) і (1.2), отримуємо співвідношення (1.4*a*) і (1.4), які виражають закон Столєтова:

$$n_0 = \frac{\eta S}{\varepsilon} I \;\Rightarrow\; n_0 \sim I \text{ і } i_H \sim I.$$

Прикметно, що хвильова теорія і тут дає зовсім інший прогноз. За хвильовим механізмом кожне вимушене коливання електрона потенційно є спробою його виходу з металу. Тому кількість електронів, які виходять з металу за 1 с, мала би бути прямо пропорційною частоті опромінюючого світла, що не відповідає дійсності.

3. ГАЛЬМІВНЕ РЕНТГЕНІВСЬКЕ ВИПРОМІНЮВАННЯ

Згідно з класичною електродинамікою Максвелла, всяка заряджена частинка, котра має прискорення, випускає електромагнітне випромінювання. Цим пояснюється так зване ***гальмівне рентгенівське випромінювання***, що спостерігається при бомбардуванні металевої пластинки швидкими електронами. На практиці таке випромінювання отримують за допомогою рентгенівської трубки, яка схематично зображена на рис. 1.5*а*. На рис. 5*б* показані фотографії деяких промислових рентгенівських трубок.



а) *б*)

Рис. 1.5

Рентгенівська трубка по суті є потужним вакуумним діодом, на який подається висока напруга U (десятки й сотні кіловольт). Тому електрони, що емітуються розжареним катодом K, на шляху до анода A отримують велику кінетичну енергію

$$K = eU.$$

При зіткненнях із анодом і різкому гальмуванні кінетична енергія електронного пучка майже повністю переходить у внутрішню енергію анода. (Тому для запобігання сильному нагріванню і розплавленню анода він примусово охолоджується). Але приблизно 1% кінетичної енергії електронів трансформується в електромагнітне випромінювання рентгенівського діапазону, так що для кожного зіткнення можна записати:

$$eU = W + Q, \qquad (1.11)$$

де W – енергія випромінювання, Q – теплова енергія, віддана атомам анода.

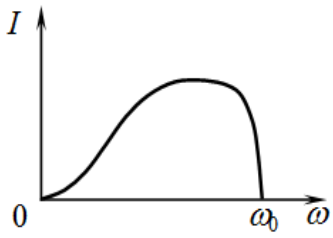


Рис. 1.6

Дослід показує, що розподіл енергії гальмівного випромінювання по частотах (спектр) є неперервним (рис. 1.6), але має короткохвильову межу ω_0, так що в спектрі присутнє лише випромінювання з частотами $\omega < \omega_0$. При цьому частота ω_0 не залежить від матеріалу анода трубки і є прямо пропорційною напрузі на ній:

$$\omega_0 = aU, \quad a = const.$$

Закони класичної фізики не накладають ніяких обмежень на частоту чи довжину хвилі гальмівного випромінювання зарядів і не спроможні пояснити існування короткохвильової межі ω_0. Але в квантовій теорії з цим не виникає жодних проблем. При зіткненні з анодом і гальмуванні електрона випромінюється квант енергії $\varepsilon = \hbar\omega$, тож енергетичний баланс (1.11) має вигляд:

$$eU = \hbar\omega + Q, \tag{1.12}$$

Енергія кванта не може бути більшою, ніж кінетична енергія електрона, за рахунок якої він утворюється:

$$\hbar\omega \le eU.$$

Отже, існує гранична частота випромінювання

$$\omega_0 = \frac{e}{\hbar}U. \tag{1.13}$$

Відповідно, мінімальна довжина хвилі в гальмівному спектрі дорівнює

$$\lambda_0 = \frac{2\pi\hbar c}{eU} = \frac{hc}{eU}. \tag{1.13a}$$

Таким чином, квантова теорія пояснює як саме існування короткохвильової межі в спектрі гальмівного випромінювання, так і незалежність від матеріалу анода та пряму пропорційність напрузі на трубці її частоти.

Спектральна крива рис. 1.6 при підході до граничної частоти обривається дуже круто, тому в експерименті частоту ω_0 можна надійно виміряти і з великою точністю визначити величину $\hbar$:

$$\hbar = \frac{eU}{\omega_0}.$$

Такий спосіб є найточнішим з існуючих експериментальних методів визначення сталої Планка.

Наостанку слід сказати, що при підвищенні напруги на трубці на тлі суцільного гальмівного спектра починають з'являтися дуже вузькі смуги (спектральні лінії), частоти яких є індивідуальними для кожного хімічного елемента. Таке рентгенівське випромінювання, що називається ***характеристичним***, пов'язане з особливостями внутрішньої будови атомів. (Детальніше про це див. Лекція 6.6).

4. КОНТРОЛЬНІ ЗАПИТАННЯ

1. У чому полягає квантова гіпотеза Планка? Чим визначається енергія кванта?

2. Що таке фотоефект? Чому він відноситься до квантових явищ? Які ще квантові оптичні явища Ви знаєте?

3. Зобразіть вольт-амперну характеристику (ВАХ) вакуумного фотоелемента. Чим пояснюється насичення фотоструму та його наявність при зворотній напрузі на фотоелементі?

4. Що таке запірна напруга $U_з$? Чому при напрузі на фотоелементі $U > U_з$ фотострум не одразу приймає максимальне значення $I_н$, а змінюється поступово, в міру зростання напруги?

5. Як залежать параметри ВАХ фотоелемента ($I_н$, $U_з$) від частоти та інтенсивності опромінюючого світла?

6. Укажіть зв'язок між параметрами ВАХ фотоелемента ($I_н$, $U_з$) та фізичними характеристиками фотоелектронів.

7. Сформулюйте закономірності зовнішнього фотоефекту.

8. Доведіть, що властивості фотоефекту неможливо пояснити на основі хвильових уявлень про світло.

9. Поясніть квантовий механізм фотоефекту та запишіть рівняння Ейнштейна.

10. Поясніть на основі квантового механізму всі закономірності фотоефекту.

11. Запишіть рівняння Ейнштейна (1.7) :

– через запірну напругу фотоелемента $U_з$;

– через частоту опромінюючого світла ω (та ν) і червоної межі фотоефекту ω_0 (та ν_0);

– через довжину хвилі опромінюючого світла λ та червоної межі фотоефекту λ_0;

12. Зобразіть ВАХ фотоелемента при опроміненні його світлом червоної межі;

13. Поміркуйте, чи буде змінюватися струм насичення фотоелемента, якщо змінювати частоту опромінюючого світла при незмінній інтенсивності;

14. Згідно з рівнянням Ейнштейна (1.7) кінетична енергія фотоелектронів лінійно залежить від частоти опромінюючого світла. Чи означає це, що при збільшенні частоти світла вдвічі запірна напруга фотоелемента теж зросте в два рази?

15. Поясніть природу гальмівного рентгенівського випромінювання;

16. Покажіть на рисунку залежність інтенсивності гальмівного випромінювання від його частоти (і довжини хвилі) та напруги на рентгенівській трубці;

17. Чим пояснюється існування короткохвильової межі у гальмівному рентгенівському спектрі та від чого і як залежить її частота (і довжина хвилі)?

6.2. ФОТОНИ

У цій лекції розглянуті такі питання:

1. Властивості фотонів
2. Тиск світла
3. Ефект Компотна
4. Контрольні запитання

1. ВЛАСТИВОСТІ ФОТОНІВ

Енергія та імпульс фотона. Результати експериментального Дослідження залежності запірної напруги від частоти світла (Лекція 6.1, п. 1.1, рис. 1.4) для фотоелементів з катодами із різних металів, указують на те, що при фотоефекті електрон поглинає всю енергію кванта світла, що випускається джерелом. А це означає, що енергія кванта і при випромінюванні джерелом, і при поширенні до катода, і при поглинанні електроном лишається зосередженою в мікроскопічній області простору. Тобто,

світлові кванти є специфічними мікроскопічними частинками – носіями електромагнітної енергії.

Уявлення про те, що випромінювання переноситься частинками, ввів Ейнштейн. Пізніше ці частинки по аналогії з носіями електрики – електронами назвали *фотонами*. Таким чином, світло являє собою не сукупність дискретних «порцій» електромагнітної хвилі, а потік елементарних частинок – фотонів.

Фотон, як будь-яка частинка, характеризується визначеною енергією та імпульсом. Але він є дуже специфічними об'єктом, не схожим на звичну частинку.

Перш за все, відповідно до формули Планка енергія фотона визначається не швидкістю його руху, а частотою випромінювання, котре ним переноситься:

$$\varepsilon = \hbar\omega, \tag{2.1}$$

або

$$\varepsilon = h\nu. \tag{2.1а}$$

Енергію фотона можна обчислити й через довжину світлової хвилі *у вакуумі* за формулою:

$$\varepsilon = \frac{2\pi\hbar c}{\lambda} = \frac{hc}{\lambda}. \tag{2.1б}$$

Окрім того фотон не має власної маси (маси спокою m_0). До такого висновку приводить формула енергії релятивістської частинки

$$\varepsilon = \frac{m_0 c^2}{\sqrt{1-(v/c)^2}} \tag{2.2}$$

та той факт, що швидкість світла у вакуумі, тож і швидкість фотонів, $v = c$. Через це енергія фотона має визначену величину тільки при $m_0 = 0$. Отже, фотон є *безмасовою частинкою.*

Імпульс фотона як безмасової частинки, згідно з теорією відносності, визначається тільки його енергією:

$$p = \frac{\varepsilon}{c}. \tag{2.3}$$

Відповідно до виразів (2.1), (1.2*а*) та (2.1*б*), його можна обчислювати за формулами:

$$p = \frac{\hbar\omega}{c} = \frac{h\nu}{c}, \tag{2.4}$$

або

$$p = \frac{2\pi\hbar}{\lambda} = \frac{h}{\lambda}. \tag{2.5}$$

Але в теорії імпульс фотона найчастіше виражають через хвильовий вектор випромінювання, носієм якого він є:

$$p = \hbar k \quad \Rightarrow \quad \vec{p} = \hbar\vec{k}, \quad k = \frac{2\pi}{\lambda}. \tag{2.6}$$

Таким чином, енергія та імпульс «частинки світла» – фотона визначаються спектральним складом випромінювання, а не швидкістю та масою, як для вільної частинки речовини, приміром, для електрона.

Дослід Боте. Існування фотонів було підтверджене і прямими дослідами. Одним із таких є дослід Боте, схема котрого наведена на рис. 2.1.

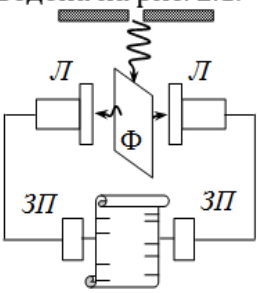


Рис. 2.1

На тонку металеву пластинку (фольгу) Ф в торець спрямовувався вузький пучок рентгенівського випромінювання, через що вона сама починала випускати рентгенівські промені (це явище називається рентгенівською люмінесценцією). По обидва боки від фольги на однаковій відстані розміщувалися лічильники – прилади, при потраплянні в які радіації на виході створюється електричний сигнал. Для реєстрації випромінювання лічильники приєднувалися до записуючих пристроїв *ЗП*, які робили відмітки на рухомій стрічці.

Згідно з гіпотезою Планка світло випромінюється дискретними порціями-квантами, але через велику щільність потоку в пучку воно за звичайних умов сприймається як неперервне. Тому в досліді фольга опромінювалася дуже слабким пучком, й інтенсивність випромінювання, що потрапляло в лічильники, була дуже малою. За таких умов кількість квантів, які потрапляли в лічильники кожної миті теж була малою. Тому від лічильників на записуючі пристрої приходили поодинокі імпульси напруги, котрі впевнено реєструвалися на стрічці.

Можливі результати такого досліду легко спрогнозувати апріорі. Якби кванти були просто порціями хвильової енергії (короткими хвильовими імпульсами), то лічильники мали би спрацьовувати одночасно, позаяк хвильове збудження поширюється від джерела по всіх напрямках з однаковою швидкістю $c = 3 \cdot 10^8$ м/с. У такому разі відмітки на стрічці були би строго синхронними. Але насправді лічильники в досліді спрацьовували ніяк не узгоджено, і мітки на стрічці не збігалися в часі, за винятком поодиноких випадкових подій. Отже, в кожному випадку фольга випромінює або в один, або в інший бік. А це незаперечно свідчить про те, що випромінювання випускається і поширюється у вигляді частинок – фотонів.

Розсіювання фотонів. Із виразу (2.2) випливає ще одна незвична властивість фотона: при будь-якій швидкості $v \neq c$ виходить $\varepsilon = 0$. Це означає, що рух із граничною швидкістю – це єдиний стан, в якому може перебувати фотон. Тому зіткнення фотонів з іншими частинками (розсіювання) теж є дуже специфічним. Для звичайної кульки зміна швидкості при зіткненні з іншим тілом залежить від співвідношення мас і може бути різною. Тільки в граничному випадку зіткнення із нерухомим тілом нескінченної маси (стіною) кулька може або повністю втратити швидкість при абсолютно непружному ударі, або відскочити без зміни швидкості, якщо удар абсолютно пружний. А ось фотон так поводиться завжди, тобто

для фотона при зустрічі з будь-якою частинкою є тільки дві можливості – або непружно зіткнутись і зникнути (поглинутись), або відбитися без зміни величини швидкості.

Яким саме буде зіткнення залежить від величини енергії та імпульсу фотона і від будови та внутрішніх властивостей частинки, з якою він стикається. При цьому виявляється, що при зіткненні з вільними елементарними частинками фотони за будь-яких умов не можуть поглинатися, бо такий процес є несумісним із законами збереження енергії та імпульсу. Аби в цьому переконатися, розглянемо гіпотетичне поглинання (непружне зіткнення) фотона з енергією ε і імпульсом $\vec{p}$ вільним нерухомим електроном з енергією спокою E_0 (Звичайно, це треба робити на основі релятивістської механіки (Розділ 1, Лекція 1.9)). При поглинанні фотона електрон має перебрати на себе всю його енергію та імпульс. Відтак енергія E та імпульс $\vec{p}_e$ електрона після
зіткнення повинні дорівнювати:

$$\begin{cases} E = \varepsilon + E_0, \\ p_e = p. \end{cases} \qquad (2.7)$$

Звідси, взявши до уваги вирази енергії фотона $\varepsilon = pc$ і електрона $E = \sqrt{(p_e c)^2 + E_0^2}$, після перетворень дістанемо:

$$\varepsilon + E_0 = \sqrt{\varepsilon^2 + E_0^2} \quad \Rightarrow \quad (\varepsilon + E_0)^2 = \varepsilon^2 + E_0^2 \quad ??$$

Отримана рівність є очевидно неможливою, що свідчить про несумісність рівнянь (2.7). Це означає, що в розглянутому гіпотетичному процесі поглинання фотона електроном не можливе одночасне збереження енергії та імпульсу, отже, не можливий і сам такий процес. Складається так, що вільний електрон просто не здатен перебрати на себе всю енергію та імпульс фотона.

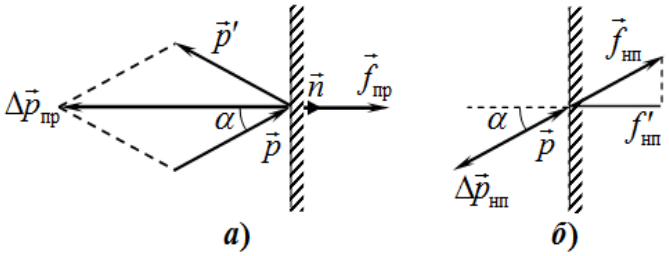
У такому разі постає питання, чому, попри сказане, при фотоефекті вільні електрони металу таки поглинають фотони. Відповідь у тому, що в металі електрони є вільними умовно, лише в сенсі спроможності рухатися по всьому об'єму зразка і створювати електричний струм. Але через взаємодію з іонами ґратки кожен електрон провідності насправді є зв'язаними з усією кристалічною ґраткою металу, і мірою зв'язку є робота виходу електрона. За таких умов поглинання фотона стає можливим, оскільки «зайві» енергія та імпульс передаються всьому металу, збільшуючи його внутрішню енергію. Так само можливе поглинання фотона електроном атома діелектрика, за умови, що енергія фотона достатня для переведення атома у збуджений стан, тобто для зміни його внутрішньої енергії.

2. ТИСК СВІТЛА

Як було доведено в електромагнітній теорії Максвелла і підтверджено експериментально в дослідах Лебедєва, світло створює тиск на тіла, на які падає. У фотонній теорії тиск світла пояснюється зіткненнями фотонів із поверхнею тіла, що опромінюється, подібно до того, як тиск газу в посудині є наслідком зіткнень його молекул із стінками посудини.

Розрахуємо тиск світла на поверхню із коефіцієнтом відбивання ρ.

Зіткнення частинок із стінкою. При зіткненні частинки з нерухомою поверхнею



а) *б*)

Рис. 2.2

(стінкою) між ними виникає сила взаємодії $f = |\Delta\vec{p}|/\Delta t$, де $|\Delta\vec{p}|$ – модуль зміни вектора імпульсу частинки за час взаємодії (час зіткнення) Δt. Тож можна сказати, що при зіткненні на стінку діє середня сила рівна імпульсу, що передається стінці за одиницю часу. При цьому величина переданого імпульсу, окрім іншого, залежить від характеру зіткнення. У випадку пружного зіткнення частинка симетрично відбивається без зміни модуля імпульсу ($\Delta p = p' - p = 0$), як показано на рис. 2.2*а*. Але вектор імпульсу змінюється, і, як видно з рисунка, модуль його зміни складає:

$$|\Delta\vec{p}|_{пр} = |\vec{p}' - \vec{p}| = 2p\cos\alpha,$$

де α – кут падіння частинки на стінку. При непружному зіткненні $\vec{p}' = 0$ (рис.2.2*б*), і

$$\Delta\vec{p}_{нп} = -\vec{p} \quad \Rightarrow \quad |\Delta\vec{p}|_{нп} = p.$$

Тож залежно від виду зіткнення на стінку діє різна за величиною сила. При пружному зіткненні вона рівна $f_{пр} = (2p/\Delta t)\cos\alpha$ і напрямлена по нормалі до стінки, а при непружному сила $f_{нп} = (p/\Delta t)$ і напрямлена під кутом α до нормалі. Але тиск створює лише нормальна складова $f' = f\cos\alpha$. Таким чином, при зіткненнях із стінкою частинки створюють короткочасні мікроскопічні сили тиску

$$f_{пр} = \frac{2p}{\Delta t}\cos\alpha, \quad \text{або} \quad f'_{нп} = \frac{p}{\Delta t}\cos\alpha, \qquad (2.8)$$

залежно від характеру зіткнення.

Тиск світла. Зазвичай щільність фотонів у пучку є дуже високою, тому вони вдаряють у стінку так часто та близько один від одного, що створюють не окремі мікроскопічні «уколи», а постійну силу тиску F рівномірно розподілену по поверхні стінки. Для підрахунку цієї сили припустимо, що на ділянку стінки з коефіцієнтом відбивання ρ і площею S за час Δt падає ΔN фотонів. Частина з них $\Delta N_{\text{і д}} = \rho \Delta N$ стикається зі стінкою пружно (відбивається), а решта $\Delta N_{\text{нп}} = \Delta N - \Delta N_{\text{пр}} = (1-\rho)\Delta N$ – непружно (поглинається). Тому, відповідно до виразів (2.8), сумарна сила тиску $F = \Delta N_{\text{пр}} f_{\text{пр}} + \Delta N_{\text{нп}} f'_{\text{нп}}$ визначається, як

$$F = \frac{n\varepsilon}{c}(1+\rho)\cos\alpha, \qquad (2.9)$$

де $n = (\Delta N/\Delta t)$ – потік фотонів, тобто кількість фотонів, які падають на вибрану ділянку стінки за одиницю часу, і ε – енергія одного фотона. Поділивши цей вираз на площу S виділеної ділянки стінки,
отримаємо величину тиску світла:

$$P = \frac{\Phi}{cS}(1+\rho)\cos\alpha,$$

де $\Phi = n\varepsilon$ – потік світлової енергії, тобто, енергія, що падає на поверхню S за одиницю часу. Величину Φ можна виразити через інтенсивність світла $I = (\Phi/S_\perp)$, яка за означенням дорівнює потокові енергії, що переноситься через одиницю площі поперечного перерізу пучка. Отже, зробивши у виразі тиску заміну $\Phi = IS_\perp$ і врахувавши, що $(S_\perp/S) = \cos\alpha$ (див. рис. 2.3), отримаємо остаточно:

$$P = \frac{IS_\perp}{cS}(1+\rho)\cos\alpha \quad \Rightarrow \quad P = \frac{I}{c}(1+\rho)\cos^2\alpha. \qquad (2.10)$$

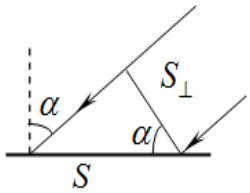


Рис. 2.3

Для крайніх випадків ідеально дзеркальної $(\rho = 1)$ та абсолютно чорної $(\rho = 0)$ поверхні виходить, відповідно:

$$P_{\text{дз}} = \frac{2I}{c}\cos^2\alpha \qquad \text{і} \qquad P_{\text{ч}} = \frac{I}{c}\cos^2\alpha. \qquad (2.10a)$$

При нормальному падінні світла $(\alpha = 0)$:

$$P = (1+\rho)\frac{I}{c}, \quad P_{\text{дз}} = \frac{2I}{c} \qquad \text{і} \qquad P_{\text{ч}} = \frac{I}{c}. \qquad (2.10б)$$

Слід зазначити, що при звичайних інтенсивностях тиск світла є дуже малим. Приміром, тиск сонячного випромінювання на земній орбіті складає всього $5 \cdot 10^{-6}$ Па. Проте, російський фізик Лебедєв ще на початку XX сторіччя, подолавши низку складних технічних проблем, зміг виміряти тиск світла в лабораторних умовах.

Попри малість, тиск світла спричинює видимі ефекти. Зокрема, тиск світла є однією з причин відхилення хвостів комет при проходженні повз Сонце в протилежний від нього бік. На тиску світла також ґрунтуються інженерні проекти «сонячного паруса» як засобу приведення в рух космічних пристроїв.

3. ЕФЕКТ КОМПТОНА

Яскравим доказом фотонної структури електромагнітного випромінювання є ефект Комптона – квантове явище, яке відкрив і теоретично пояснив американський фізик А. Комптон і яке полягає в зміні довжини хвилі рентгенівського випромінювання при його розсіюванні на електронах.

Експеримент. Схема досліду Комптона показана на рис. 2.3. Вузький пучок рентгенівських променів із заданою довжиною хвилі, який формується двома малими отворами (діафрагмами), падає на тонку пластинку досліджуваної речовини (мішень) *М*. Спектральний склад розсіяного випромінювання досліджується за допомогою рентгенівського спектрографа, який працює на принципі рентгенівської дифракції і складається з кристала *Кр* і детектора *Д*. Спектрограф дозволяє вимірювати розподіл інтенсивності в розсіяному пучку по довжинах хвилі $I(\lambda)$ при будь-якому заданому куті розсіювання ϑ

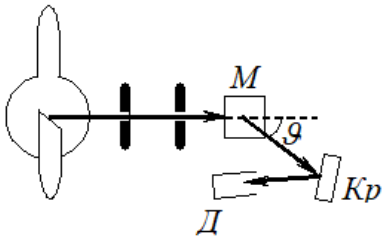


Рис. 2.4

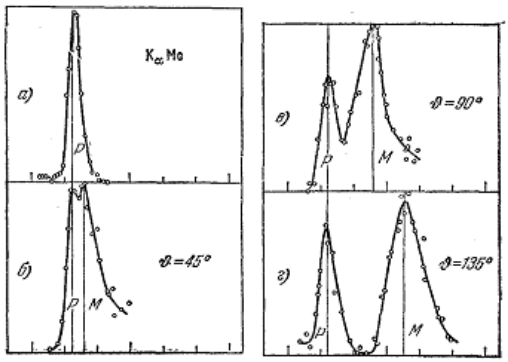


Рис. 2.5

На рис. 2.5 показані результати такого досліду на графіті з використанням K_α-лінії характеристичного рентгенівського випромінювання молібдену. Крива *а*) знята при куті ϑ = 0° і показує довжину хвилі λ випромінювання, що падає на мішень. Криві *б*) – *г*) відповідають розсіюванню під указаними кутами ϑ.

Із наведених кривих видно, що в розсіяному пучку з'являється випромінювання із зміщеною довжиною хвилі $\lambda' > \lambda$ і тим більшою, чим більший кут розсіювання ϑ. Аналіз таких вимірів при різних кутах розсіювання і на різних мішенях показав, що зміна довжини хвилі – *комптонівський зсув* – $\Delta\lambda = \lambda' - \lambda$ не залежить від речовини мішені й визначається тільки кутом розсіювання, згідно з формулою:

$$\Delta\lambda = \lambda' - \lambda = \lambda_c(1 - \cos\vartheta) = 2\lambda_c \sin^2\frac{\vartheta}{2}, \qquad (2.11)$$

де λ_c = 2,43·10^{-12} м = 2,43·пм – універсальна константа, що називається *комптонівською довжиною хвилі електрона*.

Теорія. Насамперед звернімо увагу на те, що на основі хвильових уявлень про випромінювання пояснити ефект Комптона неможливо. Адже в класичній теорії причиною розсіювання є вторинне випромінювання, зумовлене вимушеними коливаннями електронів в атомах мішені під дією електромагнітного поля падаючого випромінювання. Частота вимушених коливань збігається з частотою падаючої хвилі, тож усупереч досліду частота та довжина хвилі падаючого і розсіяного випромінювання не повинні відрізнятися.

Але в квантовій теорії ефект Комптона легко пояснюється. В рамках фотонної концепції розсіювання випромінювання відбувається внаслідок зіткнень одних частинок – фотонів – з іншими частинками – електронами атомів мішені. У випадку рентгенівського випромінювання енергія фотонів складає $\left(10^4-10^5\right)$ еВ, тоді як енергія зв'язку з ядром для валентних електронів не перевищує 25 еВ. Отже, валентні електрони в будь-якій речовині для рентгенівських фотонів є практично вільними елементарними частинками. Тому при зіткненні рентгенівського фотона з валентним електроном поглинання неможливе, (див. п. 2.1), точніше – вкрай мало імовірне. Відбуваються тільки пружні зіткнення, при яких фотон передає частину своєї енергії та імпульсу електрону. Відтак енергія розсіяного фотона стає меншою, а довжина хвилі – більшою, ніж у падаючого. Саме в цьому й полягає ефект Комптона.

Зіткнення між частинками підпорядковані законам збереження енергії та імпульсу. Це дозволяє розрахувати зміну довжини хвилі випромінювання внаслідок указаного пружного розсіяння фотонів на валентних електронах мішені, які можна наближено вважати нерухомими, позаяк їхня швидкість $v \ll c$.

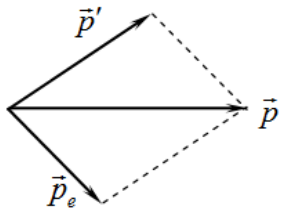


Рис. 2.6

Нехай фотон з енергією ε та імпульсом $\vec{p}$ налітає на нерухомий валентний електрон мішені, який має імпульс $p=0$ і енергію $E_0=mc^2$ (енергія спокою). Після зіткнення фотон відлітає (розсіюється) під деяким кутом ϑ з новими значеннями енергії ε' та імпульсу $\vec{p}'$ (рис. 2.6), а електрон відлітає в інший бік з деяким імпульсом $\vec{p}_e$ та енергією $E=\sqrt{p^2{}_e c^2+m^2c^4}$. При цьому за законами збереження імпульсу та енергії можна записати:

$$\begin{cases}\vec{p}=\vec{p}'+\vec{p}_e, \\ \varepsilon+mc^2=\varepsilon'+\sqrt{\left(p_e c\right)^2+m^2c^4}\end{cases}. \qquad (2.12)$$

Виключивши з цієї системи імпульс електрона, можна виразити імпульс розсіяного фотона p', тож і довжину хвилі розсіяного випромінювання λ' через характеристики фотонів, що налітають. Для цього перегрупуємо
рівняння (2.12) і піднесемо їх до квадрата:

$$\begin{cases}\left(\vec{p}-\vec{p}'\right)^2=p_e^2, \\ \left(\varepsilon-\varepsilon'+mc^2\right)^2=\left(p_e c\right)^2+m^2c^4;\end{cases} \Rightarrow \begin{cases}p^2+p'^2-2pp'\cos\vartheta=p_e^2, \\ \left(\varepsilon-\varepsilon'\right)^2+2\left(\varepsilon-\varepsilon'\right)mc^2=p_e^2c^2.\end{cases}$$

По тому, зробивши заміни $\varepsilon=pc$ і $\varepsilon'=p'c$ і виключивши з рівнянь величину p_e, після спрощень отримаємо:

$$mc\left(p-p'\right)=pp'\left(1-\cos\vartheta\right). \qquad (2.13)$$

Після підстановок $p = h/\lambda$ і $p' = h/\lambda'$ маємо:

$$mc\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) = \frac{(1-\cos\vartheta)}{\lambda\lambda'} \quad \Rightarrow \quad \lambda' - \lambda = \lambda_c(1-\cos\vartheta),$$

де

$$\lambda_c = \frac{h}{mc} = \frac{2\pi\hbar}{mc}. \qquad (2.14)$$

Підстановка констант дає $\lambda_c = 2{,}43\cdot10^{-12}$ м $= 2{,}43$ пм, що збігається із значенням комптонівської довжини хвилі електрона, знайденим у дослідах. Отже розглянутий механізм пружного розсіювання фотонів на вільних електронах вичерпно пояснює природу ефекту Комптона.

Унаслідок зіткнення з рентгенівським фотонами, електрони отримують значну енергію та імпульс і вилітають із мішені. Такі електрони називаються ***електронами віддачі***. Їхню енергію та величину і напрям імпульсу при заданому куті розсіяння також можна теоретично розрахувати з рівнянь (2.12) і порівняти з результатами відповідних вимірів. Проведені спеціальні експериментальні дослідження і в цьому відношенні повністю підтвердили правильність розглянутої фотонної теорії ефекту Комптона.

Зміна довжини хвилі, в принципі, має відбуватися при пружному розсіюванні будь-яких фотонів на будь-яких вільних частинках. Але відомо, що для світла ефект Комптона не спостерігається. Причина полягає в малій величині комптонівського зсуву, максимальна величина якого становить всього $2\lambda_c \approx 5\cdot10^{-12}$м (при $\vartheta = 180°$ у формулі (2.11)). Для рентгенівських променів ця величина складає порядку (0,5 – 5,0) % від довжини хвилі, тому комптонівський зсув можна зареєструвати і виміряти. А ось для світла ($\lambda \sim 10^{-7}$ м) максимальний комптонівський зсув складає величину $\sim 0{,}001\%$, яка знаходиться поза можливостями оптичних спектрометрів.

При розгляді результатів дослідів Комптона (рис. 2.5) може виникнути ще одне дражливе запитання: чому в спектрі *розсіяного* випромінювання спостерігається й компонента з незміщеною довжиною хвилі? Відповідь така. Фотони можуть стикатися і пружно розсіюватися не лише на зовнішніх практично вільних електронах, а й на електронах внутрішніх оболонок атомів мішені. Позаяк такі електрони дуже сильно зв'язані з ядром, розсіянні від них фотон взаємодіє фактично з усім атомом – частинкою в $(10^4 \div 10^6)$ разів масивнішою, ніж електрон. За таких умов комптонівська довжина хвилі (2.14) є меншою за розміри атомного ядра, тож говорити про ефект Комптона позбавлено сенсу.

Наостанку відмітимо, що квантова теорія ефекту Комптона так вичерпно і переконливо пояснила механізм і всі властивості явища, що після цього існування фотонів було визнане як незаперечний факт.

4. КОНТРОЛЬНІ ЗАПИТАННЯ

1. Які факти свідчать про те, що світло являє собою не просто послідовність дискретних порцій електромагнітної (хвильової) енергії, а потік частинок – фотонів.
2. Згідно з формулою (2.1*б*), енергію фотона можна визначити через довжину хвилі випромінювання. Чи означає це, що, позаяк довжина хвилі при переході світла з вакууму в речовину зменшується, то енергія фотона збільшується?
3. Звідки випливає, що фотон є безмасовою частинкою?
4. Чи може змінюватися величина швидкості фотона?
5. Чим відрізняються зіткнення фотонів із іншими частинками від зіткнень між частинками речовини, наприклад, між атомами
6. Як квантова теорія пояснює існування тиску світла? Від чого він залежить? На яку поверхню світло тисне сильніше – на дзеркальну, чи вкриту сажею?
7. У чому полягає ефект Комптона та якою формулою він виражається? Чому дорівнює максимальний можливий комптонівський зсув довжини хвилі в розсіяному пучку?
8. Якою є причина збільшення довжини хвилі при розсіюванні рентгенівських променів?
9. Поясніть, чому з точки зору класичної фізики ефект Комптона є неможливим.
10. Поясніть, чому в розсіяному рентгенівському пучку є й випромінювання з незміщеною довжиною хвилі.
11. Поясніть, чому ефект Комптона не спостерігається для світла.
12. Відомо, що в ефекті Комптона із збільшенням маси атома речовини мішені інтенсивність зміщеної компоненти у розсіяному пучку зменшується, а незміщеної – зростає. Поміркуйте, з чим це може бути пов’язано.

6.3. ХВИЛЬОВІ ВЛАСТИВОСТІ ЧАСТИНОК

У цій лекції мова йде про таке:

1. Постулати Бора
2. Хвилі де Бройля
3. Хвильова функція
4. Принцип невизначеності
5. Приклади застосування принципу невизначеності
6. Контрольні запитання

1. ПОСТУЛАТИ БОРА

Неможливість класичної теорії будови атома. Так склалося, що в роки становлення квантової теорії випромінювання були зроблені не менш революційні відкриття і в царині мікроскопічної будови речовини. Відкриття електрона (Дж. Дж. Томсон, 1898 р.) та атомного ядра (Ернест Резерфорд, 1913 р.) незаперечно довели, що атом не є неподільним (а саме так перекладається грецьке «атомос») і складається з малого масивного ($\sim 10^{-4}$ розмірів і 99,95% маси атома) позитивно зарядженого ядра та відповідної кількості легких негативно заряджених електронів, які рухаються навколо ядра під дією кулонівської сили притягання.

Ядерна модель атома нагадує сонячну систему, в якій «маленькі і легенькі» планети рухаються по орбітах навколо масивного «ядра» – Сонця під дією сили гравітаційного притягання. Але, попри позірну схожість, атом принципово відрізняється від планетарної системи тим, що «планети» – електрони мають електричний заряд. Тому за законами класичної електродинаміки через наявність доцентрового прискорення електрони при обертанні навколо ядра мали би неперервно випускати електромагнітне випромінювання з частотою рівною частоті обертання. При цьому вони через утрату енергії мали би поступово наближатися до ядра і врешті решт упасти на нього, утворивши нейтральну частинку із розмірами ядра. Причому, за розрахунками такий колапс мав би статися практично миттєво (див. п. 3.6, Задача 3.2). Крім того через неперервне збільшення частоти обертання електронів при наближенні до ядра в спектрі випромінювання атомів мала би спостерігатися широка неперервна смуга частот.

В дійсності немає нічого схожого. Атоми є стабільними і без зовнішнього впливу необмежено довго зберігають свої індивідуальні властивості. Виняток становлять тільки радіоактивні елементи. Але спонтанна зміна атомів цих елементів теж зумовлена не падінням електронів на ядро, а процесами в самому ядрі. Спектри випромінювання атомів теж не є неперервними: вони містять індивідуальні для кожного хімічного елемента дискретні частоти.

Постулати Бора. Оскільки ядерна будова атома неспростовно випливає з дослідів Резерфорда, маємо констатувати, що класична фізика не здатна пояснити не лише властивості, а й саме існування атомів. Це означає, що в атомах діють закони, відмінні від законів класичної фізики. Перший крок на шляху до їхнього установлення зробив Нільс Бор, який сформулював наступні положення (*постулати Бора*):

1. ***Атом може перебувати тільки в деяких стаціонарних («квантових») станах, енергії яких утворюють дискретний набір*** $E_1, E_2, E_3, \ldots$. ***У стаціонарних станах електрони атома не випромінюють енергію.***
2. ***Перехід атома із стаціонарного стану з енергією*** E_m ***в інший з меншою енергією*** E_n ***відбувається стрибком і супроводжується випромінюванням одного кванта світла з енергією***

$$h\nu \equiv \hbar\omega = E_m - E_n. \qquad (3.1)$$

Це співвідношення, яке спирається на квантову гіпотезу Планка, називається ***правилом частот Бора***.

Обидва постулати Бора суперечать законам класичної фізики, декларуючи, по-перше, що попри наявність прискорення, електрон в стаціонарному стані атома не випускає електромагнітного випромінювання і, по-друге, що частота випромінювання атома не має безпосереднього відношення до частоти обертання електрона навколо ядра.

На основі вказаних постулатів і законів класичної механіки Бор розрахував розміри найпростішого з атомів – атома Гідрогену та частоти в його спектрі випромінювання і отримав блискучий збіг теоретичних результатів з експериментальними даними (див. п. 3.6, Задача 3.3). Це показало, що постулати Бора правильно відображають існуючий стан речей. Але спроба застосувати борівську теорію до інших атомів не дала суттєвих результатів.

2. ХВИЛІ де БРОЙЛЯ

Гіпотеза де Бройля. Глибинні причини неспроможності класичної фізики та обмеженості досягнень напівкласичної борівської теорії атома відкрив французький фізик Луї де Бройль. Він висловив думку про те, що подібно до того, як випромінювання має властивості частинок, частинки речовини мають властивості хвиль. Відтак він сформулював *гіпотезу де Бройля*, згідно з якою

із кожною вільною частинкою, що має енергію E та імпульс $\vec{p}$, пов'язана певна хвиля з частотою

$$\omega = \frac{E}{\hbar} \tag{3.2}$$

і довжиною хвилі

$$\lambda = \frac{2\pi\hbar}{p} \equiv \frac{h}{p}. \tag{3.3}$$

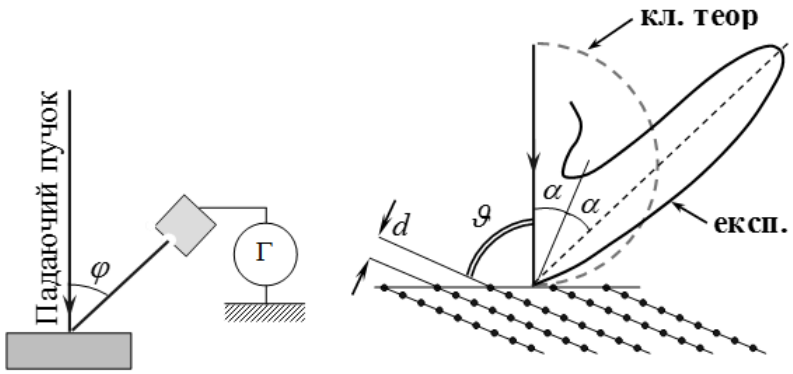
Такі асоційовані з частинками хвилі називають хвилями де Бройля, а вираз (3.3) – співвідношенням де Бройля. Із співвідношення (3.3) випливає зв'язок між хвильовим вектором $\vec{k}$ $\left(k = (2\pi/\lambda)\right)$ дебройлівської хвилі та вектором імпульсу частинки:

$$\vec{k} = \frac{\vec{p}}{\hbar}. \tag{3.4}$$

Вирази (3.2), (3.3) і (3.4) за формою і змістом збігаються з відповідними співвідношеннями для фотонів (Лекція 2, формули (2.1), (2.4), (2.6)). Тож гіпотеза де Бройля по суті стверджує, що

частинки матерії будь-якої природи виявляють і корпускулярні, і хвильові властивості, між якими існує універсальний зв'язок.

Дифракція частинок. Гіпотеза де Бройля майже одразу отримала визнання й лягла в основу квантової механіки – теорії, що адекватно відображає некласичну поведінку мікрочастинок. Цьому сприяло досить швидке її експериментальне підтвердження Девіссоном і Джермером у дослідах по розсіюванню електронів. Схема цих дослідів була така. Вузький пучок електронів, які проходили прискорююче електричне поле із заданою напругою U і набували кінетичну енергію $K = eU$, спрямовувався по нормалі на відполіровану плоску поверхню досліджуваної пластинки-мішені (рис. 3.1*а*) і вимірювався розподіл відбитих електронів по кутах розсіювання φ. Для уловлювання та реєстрації розсіяних електронів використовувався металевий порожнистий циліндр із малим отвором, заземлений через гальванометр Г. Кількість розсіяних під заданим кутом φ електронів $N(\varphi)$ визначалася за величиною струму через гальванометр.



а) *б*)

Рис. 3.1

З класичної точки зору розсіювання електронів на пластинці зумовлене зіткненням одних мікроскопічних кульок – електронів з іншими мікроскопічними кульками – атомами поверхневого шару мішені. За таких умов максимальна кількість частинок відбивається у зворотньому напрямку ($\varphi = 0°$) і монотонно зменшується до нуля при $\varphi = 90°$, як показано в полярних координатах на рис. 3.1*б* штриховою кривою. Форма такої кривої відображує закон розподілу, а відстань від початку координат до якоїсь її точки визначає кількість електронів розсіяних під заданим кутом φ.

За класичною теорією закон розсіювання не має залежати від кінетичної енергії та імпульсу частинок. Але при опроміненні електронами монокристалів Девіссон і Джермер спостерігали на кривих розсіювання піки (суцільна лінія на рис 3.1*б*), положення яких залежить від речовини мішені та енергії електронів. Зокрема, в одному з дослідів при опроміненні нікелю електронами з енергією K = 54 еВ було отримано значення $\varphi = 50°$. Указані піки нагадують дифракційні максимуми рентгенівських променів при відбиванні від кристалів, які задовольняють умову Вульфа-Бреґґа:

$$2d\sin\vartheta = \lambda, \qquad (3.5)$$

де d – відстань між атомними площинами, від яких відбиваються рентгенівські промені, ϑ – кут ковзання, тобто кут між цими площинами та напрямком променів (див. рис. 3.1*б*), λ – довжина хвилі рентгенівських променів. Це наштовхнуло Девіссона і Джермера на думку, що пік розсіяння електронів має аналогічне походження й зумовлений ***дифракцією дебройлівських хвиль*** електронів на кристалічній ґратці мішені. Для перевірки такого

припущення методом рентгеноструктурного аналізу була визначена орієнтація найщільніше заповнених атомами площин у мішені та відстань d між ними. З'ясувалося, що $d = 9{,}1 \cdot 10^{-11}$ м, і кут ковзання $\vartheta = 65°$. Останнє означає, що кут падіння пучка на атомні площини $\alpha = 90° - \vartheta = 25° = \varphi/2$, тож максимум розсіяння спостерігається в напрямку дзеркального відбивання електронів. У такому разі, згідно з умовою (3.5), довжина хвилі, що дифрагує, складає $\lambda = 2 \cdot 9{,}1 \cdot 10^{-11} \cdot \sin 65^\circ = 1{,}67 \cdot 10^{-10}$ м. Цю величину можна порівняти з дебройлівською довжиною хвилі $\lambda_{\text{Д}}$ (формула (3.3)) електронів, які падають на мішень, урахувавши, що їхній імпульс $p = \sqrt{2mK} = \sqrt{2meU}$. При $U = 54$ В, як було в досліді, обчислення дають $\lambda_{\text{Б}} = 1{,}65$ нм. Майже точний збіг величин $\lambda_{\text{Б}}$ і λ доводить дифракційну природу піків розсіяння в дослідах Девіссона і Джермера, отже, – правильність гіпотези де Бройля. Варто відмітити, що навіть невеличка різниця між розрахунковим значенням $\lambda_{\text{Д}}$ і виміряним λ знайшла пояснення: вона зумовлена деякою зміною дебройлівської довжини хвилі електронів у мішені порівняно з вакуумом, тобто – ефектом заломлення.

Пізніше інші дослідники спостерігали дифракцію нейтронів та легких атомів і молекул. Це засвідчило, що хвильові властивості притаманні всім мікрочастинкам, а не лише електронам. Більш за те, хвильові властивості частинок знайшли практичне застосування в таких галузях експериментальної та прикладної фізики, як електронна мікроскопія, електронографія та нейтронографія, електронна та нейтронна дефектоскопія, тощо.

3. ХВИЛЬОВА ФУНКЦІЯ

Швидкість хвиль де Бройля. Дебройлівська хвиля є невід'ємною приналежністю частинки і не віддільна від неї. Тому швидкість дебройлівської хвилі має збігатись із швидкістю руху частинки. Спочатку визначимо фазову швидкість $v_ф = \omega/k$ хвилі де Бройля, котра, згідно з (3.2) і (3.4), визначається енергією E та імпульсом p частинки як

$$v_ф = \frac{E}{p}.$$

Для елементарної частинки реальним є рух із релятивістською швидкістю, тому її енергія пов'язана з імпульсом співвідношенням

$$E = c\sqrt{p^2 + m^2c^2},$$

де m – власна маса (маса спокою) частинки. Відповідно, для фазової швидкості виходить:

$$v_ф = \frac{c\sqrt{p^2 + m^2c^2}}{p} = c\sqrt{1 + \left(mc/p\right)^2} \quad \Rightarrow \quad v_ф > c.$$

Таким чином, фазова швидкість хвиль де Бройля є більшою за граничну швидкість c. Саме по собі це не є чимось неможливим (див. Розділ 5, Лекція 5.6), бо реальною швидкістю поширення хвилі є групова швидкість $u = \mathrm{d}\omega/\mathrm{d}k$. З урахуванням співвідношень (3.2) і (3.4) для дебройлівської хвилі вона виражається як

$$u = \frac{\mathrm{d}}{\mathrm{d}p}\left(c\sqrt{p^2 + m^2c^2}\right) = \frac{c}{\sqrt{1 + \left(mc/p\right)^2}}. \qquad (3.6)$$

Підставивши сюди вираз релятивістського імпульсу $p = mv\Big/\sqrt{1 - \left(v/c\right)^2}$, після нескладних перетворень отримаємо $u = v$. Таким чином, швидкість руху частинки збігається із груповою швидкістю її дебройлівської хвилі.

Хвильова функція. Отриманий результат дозволяє трактувати асоційовану хвилю як зв'язаний із частинкою хвильовий пакет, тобто «згусток коливань» (див. Розділ 5, Лекція 5.6). Отже, інформацію про місцеположення (вважатимемо, що воно визначається однією координатою x) і рух мікрочастинки можна отримати, не розглядаючи її як матеріальну точку, а досліджуючи асоційовану з нею хвилю.

Саме такий підхід використовується у квантовій (хвильовій) механіці. Відповідно, основним поняттям квантової механіки є хвильова функція частинки $\Psi(x,t)$ – функція координат і часу, що визначає параметри асоційованої з частинкою хвилі в будь-якій точці простору й у будь-який момент часу. Тому важливо розуміти природу дебройлівських хвиль і зміст хвильової функції. І перше, що спадає на думку — це, що сама частинка і є дебройлівським пакетом, тобто мікроскопічним згустком хвиль. Але таке просте уявлення не витримує критики хоча б тому, що елементарна частинка є неподільною, а хвиля на межі дотику двох середовищ ділиться, частково відбиваючись і частково проходячи з одного середовища в інше. До того ж з'ясувалося, що хвильова функція є суттєво комплексною, причому і дійсна, і уявна її складові містять інформацію про фізичний стан частинки. Це змушує визнати, що дебройлівська хвиля як така не може бути якимось фізичним процесом у самій частинці чи в прилеглому до неї просторі.

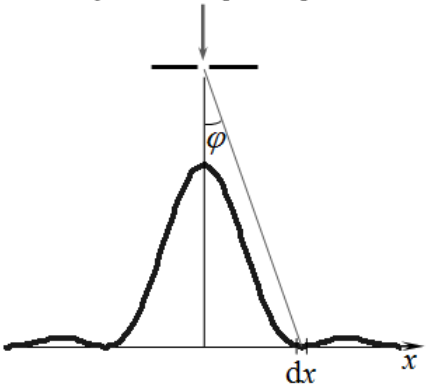


Рис. 3.2

Загально прийнята інтерпретація хвильової функції належить М Борну. Для кращого її розуміння проаналізуємо досліди по дифракції електронів на щілині. Пучок електронів із заданою дебройлівською довжиною хвилі (монохроматичний пучок) спрямовується нормально на вузьку щілину і вимірюється розподіл електронів на розміщеному за щілиною екрані-детекторі (рис. 3.2). Таким детектором може бути чутлива фотопластинка, на якій в місці потрапляння електронів виникає почорніння, або люмінесцентний екран із чутливим фотоприймачем, який дозволяє реєструвати короткі спалахи світла в місцях потрапляння на екран електронів. Досліди показали, що при інтенсивному опроміненні щілини електрони утворюють на екрані систему максимумів і мінімумів інтенсивності $I(x)$, положення яких збігається з положенням дифракційних смуг дебройлівських хвиль електронів.

Інтенсивність пучка в заданій точці екрана $I(x)$ – то є кількість електронів $\mathrm{d}N_x$, які потрапляють у задану елементарну смужку на екрані за одиницю часу, віднесена до її ширини dx:

$$I(x)=\frac{\mathrm{d}N_x}{\mathrm{d}x}.$$

З іншого боку, позаяк величина $I(x)$ визначається законом дифракції дебройлівських хвиль, вона має бути прямо пропорційною квадратові їхньої амплітуди і визначатися за правилами алгебри комплексних величин як

$$I(x)\sim\left|\Psi(x)\right|^2\equiv\Psi(x)\Psi^*(x),$$

де $\Psi^*(x)$ – комплексно спряжена функція.

Отже, кількість електронів, що потрапляють у дане місце на екрані є прямо пропорційною квадрату модуля хвильової функції у цьому місці:

$$\mathrm{d}N_x\sim\left|\Psi(x)\right|^2\mathrm{d}x. \qquad (3.7)$$

Апріорі можна припустити, що причиною "хвильової" поведінки електронів у пучку є якась специфічна взаємодія між ними. Для перевірки такого припущення описаний дослід було повторено при гранично слабкому потоці електронів, коли електрони проходили крізь щілину по одному з інтервалом часу в десятки тисяч разів більшим, ніж час їхнього прольоту через всю установку. Тож взаємодія між електронами виключалась. Такий дослід дав наступні результати:

1) Кожен окремий електрон потрапляє в певну точку екрана, а не створює мікроскопічну дифракційну картину. Отже, при взаємодії з екраном електрон виявляє себе не як хвиля, а як корпускула (точковий об'єкт).

2) Окремі електрони потрапляють у різні точки екрана не передбачувано, так що неможливо знати, в яку саме точку потрапить наступний електрон.

3) В різні області екрана електрони потрапляють не однаково часто, а так що після тривалого опромінювання щілини утворюється така сама дифракційна картина, як і при короткочасному опроміненні інтенсивним пучком.

З наведених результатів випливає, що потрапляння окремого електрона в задану область екрана відбувається не обов'язково, а лише з певною імовірністю. Тому кількість електронів $\mathrm{d}N_x$, які потрапляють у задану гранично вузьку область ширини dx (рис. 3.2), можна виразити, як $\mathrm{d}N_x = N\mathrm{d}P_x$, де N – загальна кількість електронів, які пройшли крізь щілину, і $dP_x = dN_x/N$ – імовірність (шанс) потрапляння окремого електрона в область dx. Зрозуміло, що вона є прямо пропорційною ширині смужки, тож можна записати:

$$\mathrm{d}P_x = \rho(x)\mathrm{d}x, \tag{3.8}$$

де величина

$$\rho(x) = \frac{\mathrm{d}P_x}{\mathrm{d}x} = \frac{\mathrm{d}N_x}{N\mathrm{d}x} \tag{3.8a}$$

називається ***густиною імовірності*** або ***функцією розподілу*** і визначає імовірність потрапляння окремого електрона в смужку одиничної ширини, розташовану в околі заданої точки. Відповідно, імовірність перебування частинки у скінченній області шириною $[x_1, x_2]$ визначається, як

$$P = \int_{x_1}^{x_2} \rho(x)\mathrm{d}x.$$

Таким чином, кількість електронів, які потрапляють в елементарну область dx на екрані, виражається через густину імовірності $\rho(x)$, як

$$\mathrm{d}N_x = N\rho(x)\mathrm{d}x.$$

Порівнюючи цей вираз із виразом (3.7), доходимо висновку, що густина імовірності потрапляння електрона в дану точку екрана є прямо пропорційною квадрату модуля його хвильової функції в цій точці:

$$\rho(x) \sim |\Psi(x)|^2 \tag{3.9}$$

Сама ж хвильова функція, як відмічалося вище, безпосереднього фізичного змісту не має. Це дозволяє підбирати (нормувати) амплітуду хвильової функції так, аби виконувалася рівність $\rho(x) = |\Psi(x)|^2$ і говорити, що

квадрат модуля хвильової функції частинки в будь-якій точці простору в даний момент часу дорівнює густині ймовірності перебування частинки в цій точці й у цей момент:

$$|\Psi(x,t)|^2 \equiv \Psi(x,t)\Psi^*(x,t) = \rho(x,t). \tag{3.9a}$$

Таке статистичне трактування хвильової функції, запропоноване М.Борном, було підтверджено всім подальшим розвитком квантової механіки і є загальновизнаним.

Отже, дебройлівські хвилі є не реальними фізичними хвилями, а "хвилями імовірності", або "хвилями інформації". Це накладає на хвильову функцію частинки досить жорсткі математичні обмеження:

1. Імовірність події, що може відбутися в даному місці і в даний час, не може мати одразу двох або більше значень. Тому

   хвильова функція має бути однозначною.

2. Положення та швидкість частинки не можуть змінюватися миттєво, тому

   хвильова функція разом із своїми похідними повинна бути неперервною (гладкою).

3. Імовірність за означенням не може бути більшою за одиницю, отже,

   хвильова функція має бути *обмеженою*.

Ця вимога виражається ***умовою нормування хвильової функції***:

$$\int |\Psi(x,t)|^2 \, \mathrm{d}x = 1, \qquad (3.10)$$

де інтегрування ведеться по всій області існування підінтегральної функції, тобто по області, де може знаходитися частинка. Умова (3.10) має простий зміст – вона формально виражає сам факт присутності частинки в заданій області простору протягом заданого інтервалу часу.

4. ПРИНЦИП НЕВИЗНАЧЕНОСТІ

Таким чином, через наявність у частинки хвильових властивостей неможливо одночасно точно вказати її координату та імпульс. Те саме стосується і пар деяких інших фізичних величин, які визначають стан мікрочастинки. Існування подібних обмежень відображає один із основних постулатів квантової фізики – ***принцип невизначеності***:

> ***існують такі пари фізичних величин – характеристик стану мікрочастинки, котрі точно указати разом є принципово неможливо; їх можна задати тільки з невизначеностями, добуток яких не може бути меншим за величину порядку сталої Планка.***

Математично цей принцип виражається ***співвідношеннями невизначеностей*** або ***нерівностями Гайзенберга***. Для пари координата-імпульс вони виглядають так:

$$\begin{aligned} \Delta x \cdot \Delta p_x &\geq \hbar, \\ \Delta y \cdot \Delta p_y &\geq \hbar, \\ \Delta z \cdot \Delta p_z &\geq \hbar. \end{aligned} \qquad (3.11)$$

З інших величин наведемо співвідношення невизначеностей для пари *енергія-час*:

$$\Delta E \cdot \Delta t \geq \hbar, \qquad (3.12)$$

де Δt – тривалість процесу або час існування («час життя») даного стану частинки, ΔE – невизначеність енергії частинки в цьому процесі чи стані.

Із приводу нерівностей Гайзенберга треба зауважити таке. Не лише координати, а й сама область просторово-часової локалізації квантової частинки не визначена точно. До того ж і критерії оцінки невизначеностей окремих величин у різних задачах можуть відрізнятися. Через це в теоретичні фізиці в співвідношеннях (3.11) і (3.12) замість $\hbar$ пишуть $\hbar/2$, а в підручниках із загальної фізики замість $\hbar$ можна зустріти також величину h.

Те, як "працює" принцип невизначеності можна прослідкувати на прикладі розглянутого вище дифракційного досліду з електронами. Положення кожного електрона в момент прольоту крізь щілину відоме з невизначеністю, рівною ширині щілини: $\Delta x = b$. При цьому імпульс електрона при підльоті до щілини, в принципі, точно відомий. Він визначається формулою $p = \sqrt{2meU}$ через кінетичну енергію $K = eU$, яку електрон отримує в прискорюючому полі з напругою U перед тим, як потрапити на щілину. Тому може здатися, що, зменшуючи ширину щілини, вдасться з будь-якою потрібною точністю визначити як імпульс, так і координату електрона. Але така можливість є примарною. Унаслідок дифракції в момент прольоту крізь щілину імпульс електрона неконтрольовано змінюється так, що він може потрапити в будь-яку точку на екрані. Це означає, що, визначаючи положення електрона в момент проходження щілини, ми вносимо в його імпульс невизначеність Δp_x рівну (див. рис. (3.3))

$$\Delta p_x = p \operatorname{tg} \varphi, \qquad (3.13)$$

де φ – кут дифракції.

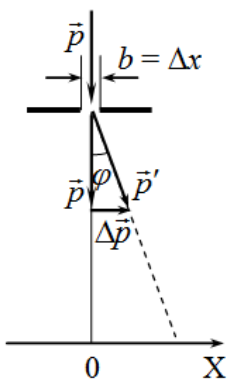


Рис. 3.3

Переважна більшість електронів після проходження щілини потрапляє в область центрального максимуму дифракції, межі якого визначаються умовою

$$b\sin\varphi = \lambda, \qquad (3.14)$$

де $b = \Delta x$ – невизначеність координати електрона, λ – його дебройлівська довжина хвилі. При цьому для мікрочастинок кути дифракції завжди малі. Тому, зробивши в (3.13) заміну $\mathrm{tg}\varphi = \sin\varphi$ і врахувавши умову (3.14) та співвідношення (3.3), отримуємо результат, який узгоджується із співвідношеннями невизначеності (3.11):

$$\Delta x \Delta p_x \approx p\lambda \quad \Rightarrow \quad \Delta x \Delta p_x \approx h.$$

Зміст принципу невизначеності. Аналізуючи розглянутий приклад, можна подумати, що невизначеності координати та імпульсу частинки є звичайними похибками експерименту, яких можна уникнути, вимірюючи положення частинки якимось іншим методом, наприклад, за допомогою мікроскопа. Але це не так. При спостереженні під мікроскопом об'єкт опромінюється світлом, тобто потоком фотонів, і реєструються відбиті від нього фотони, що потрапляють в об'єктив мікроскопа. З теорії відомо, що точність визначення положення частинки за допомогою мікроскопа обмежена довжиною хвилі використаного світла, так що $\Delta x \approx \lambda$. З іншого боку, відбиваючись від частинки, фотон передає їй імпульс, вносячи невизначеність $\Delta p_x \approx p = h/\lambda$. Отже, знову маємо $\Delta x \Delta p_x \approx h$.

Аналіз й інших принципово можливих способів одночасного визначення координат та імпульсу частинки теж приводить до співвідношення невизначеностей. Така універсальність співвідношень (3.11) свідчить, що принцип невизначеності пов'язаний не з технічною обмеженістю точності вимірювань, а із самою природою мікрочастинок. Електрон не є частинкою в класичному сенсі (корпускулою), тому до нього не можна вповні застосовувати поняття та уявлення класичної фізики. Зокрема, зміну положення мікрочастинки в просторі з плином часу принципово не можна трактувати як рух точки по визначеній траєкторії. Тож нерівності Гайзенберга задають обмеження, які накладає природа мікрочастинок на застосування понять класичної механіки при описі їхнього руху. А в ширшому розумінні

принцип невизначеності відображає обмежену можливість використання законів класичної фізики для адекватного відображення властивостей матерії на субмікроскопічному рівні.

Деякі наслідки з принципу невизначеності. Із співвідношень Гайзенберга (3.11) випливає наступний парадоксальний для класичної механіки висновок. Якщо частинка є вільною, тобто може знаходитись у будь-якій точці простору, то її положення є повністю невизначеним, тож $\Delta x = \infty$, а $\Delta p = 0$. При цьому частинка може як рухатися, так і перебувати в спокої. Але коли частинка заперта в обмеженій області, то $\Delta x \neq \infty$, і за будь-яких умов невизначеність імпульсу $\Delta p \neq 0$, отже й сам імпульс $p \neq 0$. Це означає, що квантова частинка, заперта в обмеженій області простору, принципово не може перебувати у стані спокою. При цьому, якщо $\Delta x \approx \lambda$, то $\Delta p \approx \hbar/\lambda \sim p$, тобто,

> ***коли невизначеність місцеположення частинки є величиною порядку дебройлівської довжини хвилі, то невизначеність її імпульсу за порядком величини дорівнює самому імпульсу.***

Зокрема, така ситуація є характерною для електрона в атомі, що перебуває в незбудженому або невисокому збудженому квантовому стані. Тому можна записати

$$rp = rmv \approx \hbar, \qquad (3.15)$$

де r – «радіус» атома (середня відстань електрона від ядра), m – маса електрона, v – його швидкість. У не збудженому атомі Гідрогену швидкість електрона $v \approx 2 \cdot 10^6$ м/с, тому із співвідношення (3.15) для поперечника атома Гідрогену виходить оцінка $d = 2r \approx 10^{-10}$ м, що співпадає з експериментальними даними.

Наведемо також приклад конструктивного використання співвідношень (3.11) з області ядерної фізики. Невдовзі після відкриття атомного ядра з'ясувалося, що маси ядер ізотопів усіх хімічних елементів з великою точністю є кратними масі протона m_p – ядра найлегшого ізотопу Гідрогену (H), так що $m_{\text{я}} = Am_p$. Заряд ядра є кратним величині заряду електрона e: $q_{\text{я}} = Ze$, причому число A завжди суттєво перевищує Z (для легких елементів $A \approx 2Z$). Тому спочатку думали, що ядро складається з A масивних протонів і $(A - Z)$ легеньких електронів, які компенсують відповідну частку позитивного заряду ядра, практично не впливаючи на його масу. Але принцип невизначеності змусив відкинути таку можливість. Розміри атомного ядра $r_{\text{я}} \sim 10^{-14}$ м, отже, згідно з (3.11), електрони в ядрі мають рухатись із субсвітловою швидкістю $v \approx 3 \cdot 10^8$ м/с. А при таких швидкостях сили притягання з боку протонів не здатні утримати електрон у ядрі (див. п. 3.6, Задача 3.9). Тому в ядрі не може бути електронів, а відміна чисел Z і A зумовлена присутністю в ньому інших масивних частинок – нейтронів, які не мають електричного заряду.

Одним із проявів співвідношення невизначеностей «енергія-час» (3.12) є так зване природнє розширення спектральних ліній атомів. Як відзначалося в п. 3.1, атоми випромінюють світло у вигляді квантів (фотонів) при переходах електронів із одних дискретних енергетичних рівнів на інші. Тому спектри випромінювання атомів складаються з окремих спектральних ліній, частоти яких мали би точно визначатися правилом частот (3.1). Але збуджені стани атома є нестабільними – електрон перебуває на збудженому енергетичному рівні в середньому протягом так званого «часу життя рівня» $\tau \approx 10^{-8}$ с. Тому його енергія згідно з (3.12) має невизначеність $\Delta E \approx \hbar/\tau$, яка спричинює відповідну невизначеність частоти фотонів, що випромінюються:

$$\Delta\omega = \frac{\Delta E}{\hbar} \approx \frac{1}{\tau} \approx 10^8 \text{ с}^{-1}.$$

Це означає, що при переході електрона між заданими двома рівнями за будь-яких умов випромінюються фотони не з однією, а з усіма частотами в указаному інтервалі $\Delta\omega \sim 10^8$ с$^{-1}$. Тому спектральні лінії атомів являють собою не дослівно «лінії», а вузькі смуги. Слід зауважити, що в дійсності ширина спектральних ліній атомів ще більша, оскільки є й інші причини розширення, зокрема, тепловий рух атомів. Частоти видимого світла складають $\omega \sim 10^{15}$ с$^{-1}$, тому реальний ступінь монохроматичності випромінювання атомів $\omega/\Delta\omega$ є дуже високим. Але строго монохроматичного випромінювання в природі не існує. Таким чином, відома з оптики принципова немонохроматичність реального світла має квантовомеханічну природу.

Межі застосовності класичного способу опису. Дифракційні ефекти спостерігаються не лише для електронів або нейтронів, а й для складних систем – атомів і молекул. Тому співвідношення де Бройля (3.3) і Гайзенберга (3.11), (3.12) слід трактувати як фундаментальні і принципово чинні для будь-якого матеріального об'єкта. Але добре відомо, що в багатьох випадках, наприклад, при аналізі теплового руху молекул газу в посудині або розрахунках руху електронів в електронній лампі, не виникає жодної потреби враховувати їх хвильові властивості. А ось при розгляді руху електронів у атомі без цього неможливо обійтися. Для з'ясування, чому це так, розглянемо декілька прикладів.

5. ПРИКЛАДИ ЗАСТОСУВАННЯ ПРИЦИПУ НЕВИЗНАЧЕНОСТІ

Приклад 1. Нехай маємо броунівську частинку – порошинку масою $m = 10^{-10}$ кг, яка рухається зі швидкістю v=1 см/с. При густині $\rho = 2\ \text{г}/\text{см}^3$ її розмір складає $d \approx 0{,}05$ мм. Одразу занотуємо, що хоча таку порошинку можна побачити тільки в мікроскоп, вона складається з приблизно 10^{15} молекул, тобто є *макроскопічною* частинкою. Згідно з (3.3), довжина хвилі де Бройля неї становить $\lambda_Б \approx 10^{-22}$ м, що в сто мільйонів разів менше, ніж розміри атомного ядра. Тож про хвильові властивості порошинки немає мови. Спостерігаючи за частинкою через мікроскоп, ми можемо фіксувати її положення з невизначеністю $\Delta x \approx 0{,}5$ мкм (довжина світлової хвилі), що становить всього 1% від розмірів частинки. При цьому, згідно з (3.11), невизначеність швидкості $\Delta v \approx \hbar/m\Delta x = 2 \cdot 10^{-19}$ м/с. Отже, можна практично точно визначати положення частинки та її імпульс у кожен момент часу й описувати її рух на основі законів класичної механіки.

Приклад 2. Розглянемо тепер електрон ($m = 9{,}1 \cdot 10^{-31}$ кг) в атомі Гідрогену, де його швидкість складає $v \approx 2 \cdot 10^6$ м/с. При такій швидкості дебройлівська довжина хвилі електрона $\lambda \approx 3 \cdot 10^{-10}$ м/с є сумірною з поперечником атома $d \approx 1 \cdot 10^{-10}$ м. Очевидно, що за таких умов ігнорувати хвильові властивості електрона й говорити про його рух навколо ядра по певній траєкторії (орбіті) не випадає. Справді, аби говорити про орбіту електрона навіть наближено, його положення в кожен момент часу має бути задане хоча б із невизначеністю $\Delta x \sim 0{,}1d = 10^{-11}$ м. При цьому, згідно з (3.11), невизначеність швидкості електрона $\Delta v \sim 10^7$ м/с. Отже, невизначеність швидкості $\Delta v > v$, і говорити про рух електрона навіть по наближено заданій траєкторії із наближено заданою швидкістю позбавлено сенсу. Тому вживаний зворот «орбітальний рух електронів у атомах» є тільки звичною умовністю.

Приклад 3. Розглянемо ще й випадок, коли елементарна частинка рухається у великій області простору. Прикладом може бути електронно-променева трубка, в якій зображення створюється на люмінесцентному екрані при його бомбардуванні електронами. Екран трубки має мозаїчну структуру і складається з безлічі зерен – мікрокристаликів люмінесцентної речовини розміром $d \sim 10^{-6}$ м. При цьому незалежно від місця влучання електрона спалахує весь мікрокристалик. Тому немає жодної потреби фіксувати положення електрона в трубці точніше, ніж із невизначеністю $\Delta x \sim 10^{-6}$ м, при якій згідно з (3.11) невизначеність швидкості складає $\Delta v \sim 10^{2}$ м/с. Якщо прийняти відстань від електронної гармати, що формує електронний пучок, до екрана $l \sim 10$ ñì і швидкість електронів $v \sim 10^{6}$ м/с, то відносні невизначеності складають $(\Delta x/l) \sim 10^{-5}$ і $(\Delta v/v) \sim 10^{-4}$. Ці величини становлять теоретичні похибки розрахунку руху електронів по визначеній траєкторії за законами класичної механіки і очевидно є не істотними.

Так само можна чинити й при розгляді руху вільних електронів у металі або молекул газу в посудині. Але до мікроскопічних частинок навіть у таких випадках класичний підхід не можна застосовувати беззастережно через наявність у них внутрішніх квантових властивостей. Зокрема з цієї причини є неповною класична теорія електропровідності металів, хоча рух вільних електронів відбувається в об'ємі всього провідника, тобто в макроскопічній області простору.

Аналізуючи наведені приклади, можна зробити певні загальні висновки:

– матеріальні точки й тіла класичної механіки не виявляють хвильових властивостей в силу своєї макроскопічності, тобто, через те, що вони складаються із гігантської кількості мікроскопічних частинок;

– якщо істинно мікроскопічні (елементарні) частинки рухаються в макроскопічній області простору, то з певними застереженнями їх теж можна розглядати як матеріальні точки й застосовувати до них закони класичної механіки;

– але мікрочастинки, що локалізовані в мікроскопічних областях простору, демонструють яскраво виражену некласичну поведінку, і їхній рух можна досліджувати тільки на основі законів квантової механіки.

6. КОНТРОЛЬНІ ЗАПИТАННЯ

1. Звідки випливає, що електрони в атомі знаходяться поза ядром і рухаються навколо нього?
2. Які факти стосовно будови атома не здатна пояснити класична фізика? Чому?
3. У чому полягає причина неспроможності класичної фізики в поясненні ядерної моделі атома?
4. Які факти стосовно будови атома безпосередньо відображають постулати Бора?
5. Сформулюйте гіпотезу де Бройля та запишіть співвідношення, що її виражають.
6. За допомогою виразу (3.6) покажіть, що швидкість руху частинки дорівнює груповій швидкості її дебройлівської хвилі.
7. Що таке «хвильова функція» мікрочастинки та який зміст вона має?
8. Які математичні критерії має задовольняти хвильова функція частинки, виходячи з її змісту?
9. Обґрунтує умову нормування хвильової функції мікрочастинки (3.10).
10. Сформулюйте принцип невизначеності та запишіть співвідношення, що його виражають.
11. Чи можна трактувати величини, що входять у співвідношення невизначеностей (3.11), як похибки вимірювання відповідних величин? Чому?
12. До відкриття нейтрона розглядалася можливість того, що ядро атома складається з протонів та електронів. Спростуйте таку модель на основі принципу невизначеності.
13. За яких умов можна розглядати рух мікрочастинок класично, тобто як рух матеріальних точок по визначених траєкторіях?

6.4. РІВНЯННЯ ШРЬОДІНҐЕРА

У цій лекції мова йде про таке:

1. Рівняння Шрьодінґера
2. Частинка в одновимірному потенціальному ящику
3. Частинка у тривимірному ящику. Виродження
4. Лінійний гармонічний осцилятор
5. Тунельний ефект
6. Контрольні запитання

1. РІВНЯННЯ ШРЬОДІНҐЕРА

При розгляді руху частинки головним завданням є визначення її місцеположення в будь-який момент часу. В механіці Ньютона воно вирішується шляхом розв'язування основного рівняння класичної механіки – диференціального рівняння, яке виражає другий закон Ньютона:

$$\frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2}=\frac{\vec{F}(\vec{r},t)}{m}.$$

Інтегруванням цього рівняння можна визначити закон руху $\vec{r}=\vec{r}(t)$, а по тому й рівняння імпульсу частинки $p(t)=m(\mathrm{d}\vec{r}/\mathrm{d}t)$. Тому для класичної частинки можна точно встановити положення та імпульс у будь-який момент часу, а відтак й інші величини, що визначають її механічний стан. Але для мікрочастинки можна говорити не про точні значення $\vec{r}$ і $\vec{p}$ у кожен момент часу, а тільки про ймовірності їхніх можливих значень. Цю та іншу інформацію про стан частинки містить її хвильова функція, визначення якої і є основним завданням квантової механіки.

Існують різні підходи й математичні методи вирішення цього завдання і, відповідно, – різні версії квантової механіки. Найпростішою й наочною є так звана хвильова механіка, що ґрунтується на *рівнянні Шрьодінґера* – диференціальному рівнянні, яке задовольняє хвильова функція частинки в загальному випадку. В декартових координатах воно має вигляд:

$$\frac{\hbar^2}{2m}\left(\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial y^2}+\frac{\partial^2\Psi}{\partial z^2}\right)-U\Psi=-i\hbar\frac{\partial\Psi}{\partial t}. \qquad (4.1)$$

Тут $\Psi=\Psi(x,y,z,t)$ – хвильова функція; $U=U(x,y,z,t)$ – функція потенціальної енергії частинки, котра визначає силове поле, в якому вона знаходиться; $i=\sqrt{-1}$ – уявна одиниця; $\hbar$ – стала Планка, m – маса частинки.

Рівняння (4.1) можна записати більш зручно та інваріантно щодо вибору системи координат, позначивши вираз у дужках через символ оператора Лапласа:

$$\frac{\hbar^2}{2m}\nabla^2\Psi - U\Psi = -i\hbar\frac{\partial\Psi}{\partial t}. \qquad (4.2)$$

Стосовно цього рівняння слід одразу зауважити таке. По-перше, його не можна вивести математично з якихось інших формальних положень. Тому

рівняння Шрьодінґера (4.2) є основним рівнянням квантової (хвильової) механіки,

тобто – постулатом, вірогідність якого доводиться тим, що його наслідки підтверджуються дослідом. По-друге, механіка Шрьодінґера є нерелятивістською теорією, тобто, рівняння (4.2) не є чинним при швидкостях сумірних із граничною швидкістю c.

У квантовій механіці важливе місце посідає дослідження стану частинок, які знаходяться в стаціонарних (не залежних від часу) силових полях $U = U(x, y, z)$. За такої умови розподіл імовірностей для величин, які визначають стан частинки, теж є стаціонарним і, як доводиться в математиці, хвильову функцію частинки можна записати у вигляді:

$$\Psi(x, y, z, t) = \psi(x, y, z,)e^{-i\omega t}, \qquad (4.3)$$

äå $\omega = E/\hbar$, $i = \sqrt{-1}$ і $\psi(x, y, z)$ – "координатна" хвильова функція, що є розв'язком більш простого, ніж (4.2), *стаціонарного рівняння Шрьодінґера*:

$$\nabla^2\psi + \frac{2m}{\hbar^2}(E - U)\psi = 0, \qquad (4.4)$$

де E – повна енергія частинки.

Слід зазначити, що інтегрування рівнянь (4.2) і (4.4) звичайно пов'язане з великими математичними труднощами. Тому далі детальний розгляд проводиться тільки для декількох найпростіших задач важливих для розуміння квантовомеханічного способу опису та загальних рис поведінки мікрочастинок.

2. ЧАСТИНКА В ОДНОВИМІРНОМУ ПОТЕНЦІАЛЬНОМУ ЯЩИКУ

Постановка задачі. Уявімо маленьку пружну кульку, котра рухається без тертя всередині ящика у вигляді прямого паралелепіпеда з абсолютно жорсткими непроникними стінками, що лежить на горизонтальній площині. За таких умов усередині ящика кулька рухається, як вільна частинка, і зіткнення зі стінками відбуваються пружно, без зміни величини імпульсу та кінетичної енергії. Задля спрощення викладок будемо вважати, що рух відбувається вздовж одного з ребер ящика довжини l.

При пружному зіткненні із непроникною стінкою ящика кінетична енергія кульки, якою б вона не була, повністю переходить у потенціальну енергію пружної деформації, а потім на стадії відскоку відновлюється до вихідної величини. Але при необмеженій жорсткості стінок деформації є гранично малими, тому перетворення енергії відбуваються в одній точці – точці дотику кульки до стінки ящика. В такому разі можна вважати, що потенціальна енергія кульки $U(x)$ скрізь всередині ящика дорівнює нулю, а на стінках стрибком зростає і може бути якою завгодно, як показано на рис. 4.1. Отже, можна говорити, що частинка в ящику рухається в такому силовому полі, де її потенціальна енергія $U = 0$ в області шириною l і $U \to \infty$ за її межами. Фраза досить громіздка. Тому для характеристики типових силових полів використовують відповідні жаргонні терміни. Зокрема, описане поле називають ***потенціальним ящиком*** або потенціальною ямою з нескінченно високими вертикальними стінками.

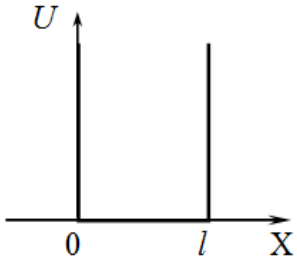


Рис. 4.1

Характеристики руху кульки в ящику – енергія та імпульс – ніяк не залежать від її маси та розмірів ящика. Вони задаються із зовні при приведенні культи в рух. Це стосується й мікрочастинки за умов, коли вона не виявляє квантових властивостей (див. Лекція 3, п. 3.4). Приміром, енергія молекул газу в посудині визначається не розмірами посудини чи масами молекул, а температурою газу.

Але поведінка мікрочастинки, що заперта в мікроскопічному "ящику", не є очевидною. Вона підпорядкована законам не класичної, а квантової механіки, і її треба розглядати на основі рівняння Шрьодінґера (4.4). Складемо його, спрямувавши координатну вісь OX уздовж напрямку руху частинки. В такому разі хвильова функція є функцією тільки однієї змінної $\psi(x)$ і $\nabla^2\psi = \left(\mathrm{d}^2\psi / \mathrm{d}x^2\right)$. Через непроникність стінок частинка не може опинитися за межами ящика ні за яких умов. Отже, хвильова функція існує тільки в області $x \in [0; l]$, де $U(x) = 0$. Тому рівняння (4.4) має вигляд:

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + k^2\psi = 0, \tag{4.5}$$

де уведено позначення

$$k^2 = \frac{2mE}{\hbar^2}. \tag{4.6}$$

При цьому з вимоги неперервності хвильової функції випливає, що вона дорівнює нулю не тільки за межами ящика, а й на його стінках:

$$\psi(0) = \psi(l) = 0. \tag{4.7}$$

Хвильові функції. Квантові стани. З математики відомо, що розв'язками рівняння (4.5) є гармонічні функції, тож можна записати:

$$\psi(x) = A\sin(kx + \alpha), \qquad (4.8)$$

де A та α – довільні константи (сталі інтегрування). Але з усієї множини розв'язків (4.8) хвильовими функціями частинки є лише ті, що задовольняють вимоги однозначності, неперервності та обмеженості (див. Лекція 3, п. 3.3).

Усі функції (4.8) є однозначними та обмеженими, але неперервними з них є лише ті, що задовольняють умови (4.7). При цьому з умови $\psi(0) = 0$ випливає $A\sin\alpha = 0 \Rightarrow \alpha = 0$, отже

$$\psi(x) = A\sin kx. \qquad (4.9)$$

А з другої умови $\psi(l) = 0$, маємо:

$$\sin kl = 0 \quad \Rightarrow \quad kl = \pm\pi n.$$

Отже, параметр k може мати лише дискретні значення

$$k_n = \pm\frac{\pi}{l}n, \qquad n = 1, 2, 3, \ldots, \qquad (4.10)$$

і зі всієї множини розв'язків (4.8) хвильовими функціями частинки є лише окремі (дискретні) функції

$$\psi_n(x) = \pm A\sin\left(\frac{\pi x}{l}n\right), \quad n = 1, 2, \ldots.$$

Відмітимо, що у множині чисел n відсутнє значення n = 0, хоча формально воно задовольняє вимозі $\psi(l) = 0$. Це й зрозуміло, адже при n = 0 хвильова функція $\psi_0(x) = 0$ при будь-яких значеннях x, у тому числі і всередині ящика. Отже, значення n = 0 є тривіальним і відповідає порожньому ящику.

Амплітуди A хвильових функцій визначаються з умови нормування (4.7):

$$A^2\int_0^l \sin^2\left(\frac{\pi x}{l}n\right)\mathrm{d}x = 1 \quad \Rightarrow \quad \frac{A^2 l}{2} = 1 \quad \Rightarrow \quad A = \sqrt{\frac{2}{l}}.$$

Таким чином, хвильові функції частинки в ящику утворюють дискретний набір:

$$\psi_n(x) = \pm\sqrt{\frac{2}{l}}\sin\left(\frac{\pi x}{l}n\right), \quad n = 1, 2, 3, \ldots \qquad (4.11)$$

Хвильова функція повністю визначає стан мікрочастинки. Тому дискретність функцій (4.11) означає, що для частинки в ящику можливі не будь-які, а лише окремі дозволені стани. Це явище називається ***квантуванням***, а самі дозволені стани – ***квантовими станами***.

Енергетичний спектр. Кожна хвильова функція ψ_n і кожен квантовий стан частинки в ящику визначається відповідним значенням параметра k_n, який пов'язаний із енергією частинки формулою (4.6). Тому, врахувавши вираз (4.10), отримаємо:

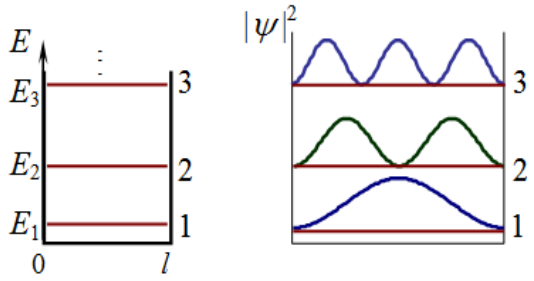
$$E_n = \frac{\pi^2\hbar^2}{2ml^2}n^2, \quad n = 1, 2, 3, \ldots \qquad (4.12)$$

Отже, енергія частинки в ящику теж є квантованою: в кожному дозволеному стані вона може мати тільки відповідне значення з набору (4.12), який називається ***енергетичним спектром*** частинки. Кожне дозволене значення енергії E_n утворює відповідний ***енергетичний рівень***, а число n називається ***квантовим числом*** цього рівня. Наочно енергетичні рівні зображують горизонтальними відрізками на енергетичній діаграмі (рис. 4.2а).

Іншою особливістю енергетичного спектра (4.12) є те, що найнижчий енергетичний рівень не дорівнює нулю:

$$E_1 = \frac{\pi^2\hbar^2}{2ml^2}. \qquad (4.12a)$$

Оскільки енергія частинки в ящику є чисто кінетичною, це означає, що частинка ні за яких умов не може перебувати у стані спокою. Такий висновок є дуже дивним для класичної механіки, але він повністю узгоджується із принципом невизначеності. Справді, якщо частинка нерухома, то її імпульс заданий точно: $p = 0$, тож невизначеність імпульсу $\Delta p = 0$. Але це суперечить співвідношенням Гайзенберга (3.10): оскільки невизначеність координати частинки в ящику $\Delta x \le l$, то повинно бути $\Delta p \ge \hbar/l$. Якщо прийняти мінімальний імпульс частинки $p_{\min} \sim \hbar/l$ (див. Лекція 3, п. 1.3), то для мінімальної енергії частинки отримаємо $E_{\min} \sim \hbar^2/2ml^2$, що за порядком величини узгоджується з виразом (4.12a).



а) ***б)***

Рис. 4.2

Розглянуті особливості поведінки частинок не спостерігаються в класичній механіці, навіть якщо йдеться про рух мікроскопічних частинок, наприклад, молекул газу в посудині. Для з'ясування причин цього оцінимо найменшу можливу кінетичну енергію E_1 та швидкість руху v_1 електрона ($m \approx 10^{-30}$ кг) в ящику ширини $l = 10^{-10}$ м і молекули кисню O_2 ($m \approx 5{,}3 \cdot 10^{-26}$ кг) в посудині розміром l = 10 см.

Для електрона маємо:

$$E_1 = \frac{mv_1^2}{2} = \frac{3{,}14^2 \cdot 1{,}05^2 \cdot 10^{-68}}{2 \cdot 9{,}1 \cdot 10^{-31} \cdot 10^{-20}} = 6 \cdot 10^{-18}\,\text{Дж} \approx 37\,\text{еВ}, \text{ і } v_1 = 3{,}6 \cdot 10^6\,\text{м/с}.$$

Це дуже багато. Наприклад, величина E_1 більш ніж у 100 разів перевищує енергію теплового руху частинок при кімнатній температурі. Натомість аналогічні розрахунки для молекули кисню в посудині дають недоступні для спостереження величини $E_1 \approx 10^{-40}$ Дж і $v_1 \approx 6 \cdot 10^{-8}$ м/с.

Це дуже багато. Наприклад, величина E_1 більш ніж у 100 разів перевищує енергію теплового руху частинок при кімнатній температурі. Натомість аналогічні розрахунки для молекули кисню в посудині дають недоступні для спостереження величини $E_1 \approx 10^{-40}$ Дж і $v_1 \approx 6 \cdot 10^{-8}$ м/с.

Сказане стосується й дискретності енергетичного спектра. Для переведення електрона в ящику ширини 10^{-10} м з першого на другий рівень необхідно надати йому енергію $\sim$ 100 еВ (в атомі Гідрогену ≈ 10 еВ). Ця величина відповідає енергії фотонів жорсткого ультрафіолетового випромінювання і в $\sim 10^4$ разів перевищує середню енергію, якою обмінюються частинки внаслідок теплового руху. А для молекули O_2 в посудині розміром $l = 10^{-10}$ м треба всього $\sim 10^{-21}$ еВ, тобто, ні про яку дискретність рівнів говорити не доводиться. Зауважимо також, що фактична відсутність квантових властивостей у молекули газу зумовлена макроскопічними розмірами ящика (посудини), в якому знаходиться молекула.

Рух частинки в ящику. Частинка в ящику рухається між стінками із сталою швидкістю. Тому класична частинка (кулька) в усіх точках, окрім стінок, буває однаково часто, тобто з однаковою імовірністю. Біля стінок ця імовірність більша, оскільки під час зіткнення частинка витрачає певний час на гальмування та наступний розгін при відскоку. А от квантова частинка і в цьому відношенню поводиться інакше. Згідно з (4.11), густина ймовірності перебування частинки в заданій точці ящика визначається функцією

$$|\psi(x)|^2 = \frac{2}{l}\sin^2(\pi n x/l).$$

Із цього виразу випливає, що для квантової частинки в ящику існують точки, в яких $|\psi|^2 = 0$. Положення таких точок визначається умовою

$$\frac{\pi n x_i}{l} = i\pi \quad \Rightarrow \quad x_i = i\frac{l}{n}, \quad i = 0, 1, 2, \ldots n.$$

Наявність таких точок ілюструє рис. 4.2б, на якому зображено графік $|\psi(x)|^2$ для трьох перших квантових станів частинки в ящику.

Таким чином для всіх станів, окрім першого (n = 1), *всередині* ящика є точки, в яких частинку не можна виявити ***ні за яких умов***. Така поведінка частинки є несумісною з поняттям руху по заданій траєкторії. Але її можна збагнути на основі хвильових уявлень. А саме. Енергія частинки в ящику є чисто кінетичною

$$E = \frac{p^2}{2m},$$

тож зробивши таку заміну у виразі (4.6), отримаємо:

$$k^2 = \frac{p^2}{\hbar^2} \quad \Rightarrow \quad k = \pm\frac{p}{\hbar} = \pm\frac{2\pi}{\lambda},$$

де, відповідно до (3.3), λ – дебройлівська довжина хвилі частинки. Це означає, що параметр k в рівнянні (4.9) є хвильовим числом дебройлівських хвиль, які разом із частинкою поширюються в ящику і відбиваються від його стінок. Але теорія хвиль говорить (Див. Лекція 4.7, п.7.4), що між двома відбиваючими стінками можуть існувати лише хвилі із дискретними довжинами λ_n, при накладанні яких утворюються стаціонарні стоячі хвилі. При цьому на відстані між стінками має укладатися ціла кількість півхвиль:

$$l = n\frac{\lambda_n}{2} \quad \Rightarrow \lambda_n = \frac{2l}{n}, \; n = 1, 2, 3, \ldots .$$

І якщо підставити сюди співвідношення де-Бройля (3.3), то для імпульсу частинки вийде

$$p_n = \frac{\pi\hbar}{l}n, \qquad (4.13)$$

а для енергії –

$$E_n = \frac{\pi^2\hbar^2}{2ml^2}n^2,$$

що збігається з виразом (4.12). Отже, ефект квантування станів можна трактувати як наслідок утворення стоячих дебройлівських хвиль при відбиванні частинки від стінок ящика. При цьому точки, в яких $|\psi(x)|^2 = 0$, є вузлами вказаних стоячих хвиль.

На завершення зазначимо, що хвильове поле на кшталт асоційованих із частинкою стоячих хвиль із дискретним набором довжин хвилі утворюється не лише при вільному русі частинки між перпендикулярними до напрямку руху «стінками», а й у більш складних силових полях, які утримують частинку в обмеженій області простору. Відповідно, стани частинки теж квантуються, але хвильові функції та правила квантування є складнішими. Отже,

квантування станів є універсальним ефектом і спостерігається завжди, коли мікрочастинка може рухатися тільки в обмеженій області простору.

Одним із проявів цього ефекту є дискретність станів електронів в атомах, які утримуються в околі ядра кулонівською силою.

Таким чином, у квантовій механіці існування дискретних дозволених станів електронів у атомах не постулюється, як в теорії Бора, а природньо випливає з основного рівняння квантової механіки.

3. ЧАСТИНКА У ТРИВИМІРНОМУ ПОТЕНЦІАЛЬНОМУ ЯЩИКУ

Квантові стани. Розглянемо тепер випадок, коли частинка перебуває в тривимірному ящику з ребрами l_1, l_2, l_3 рис. 4.3 і може в ньому вільно рухатись у будь-якому напрямку. В цьому випадку її хвильова функція є функцією трьох змінних (x, y, z), тож рівняння Шрьодінґера має вигляд:

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}+k^2\psi=0, \quad k^2=\frac{2mE}{\hbar^2}. \tag{4.14}$$

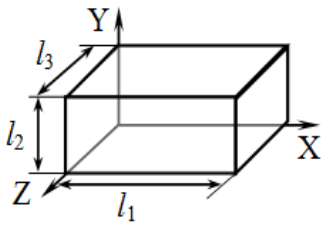


Рис. 4.3

Це рівняння легко розв'язується методом відокремлення змінних. А саме. Всередині ящика $U(x, y, z)=0$ і ніякі сили на частинку не діють. Тому координати частинки є незалежними: на можливі значення однієї з них ніяк не впливають дві інші. На мові теорії ймовірностей те, що частинка має задане значення однієї координати при будь-яких інших, є незалежною простою подією, а ймовірність перебування частинки в заданому місці, тобто імовірність того, що вона одночасно має задані значення всіх трьох координат, є складною подією. При цьому в теорії доводиться, що ймовірність складної події дорівнює добутку імовірностей простих незалежних подій, з яких вона "складається". Це, враховуючи ймовірнісний зміст хвильової функції, дозволяє шукати розв'язки рівняння (4.14) у вигляді добутку "однокоординатних" функцій, кожна з яких визначає положення частинки тільки відносно однієї координатної осі:

$$\psi(x, y, z)=\psi_1(x)\cdot\psi_2(y)\cdot\psi_3(z).$$

Підставляючи цей вираз у рівняння (4.14), маємо:

$$\psi_2\psi_3\frac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2}+\psi_1\psi_3\frac{\mathrm{d}^2\psi_2}{\mathrm{d}y^2}+\psi_1\psi_2\frac{\mathrm{d}^2\psi_3}{\mathrm{d}z^2}=-k^2\psi_1\psi_2\psi_3 \quad \Rightarrow$$
$$\frac{1}{\psi_1}\frac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2}+\frac{1}{\psi_2}\frac{\mathrm{d}^2\psi_2}{\mathrm{d}y^2}+\frac{1}{\psi_3}\frac{\mathrm{d}^2\psi_3}{\mathrm{d}z^2}=-k^2. \qquad (4.15)$$

Отримана рівність містить не числа, а функції $\psi_1(x),\psi_2(y),\psi_3(z)$ і має бути чинною при всіх значеннях змінних x,y,z. А це можливо лише за умови, що аналогічні рівності виконуються окремо для кожної із функцій:

$$\frac{1}{\psi_1}\frac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2}=-k_x^2; \qquad \frac{1}{\psi_2}\frac{\mathrm{d}^2\psi_2}{\mathrm{d}y^2}=-k_y^2; \qquad \frac{1}{\psi_3}\frac{\mathrm{d}^2\psi_3}{\mathrm{d}z^2}=-k_z^2,$$

або

$$\begin{cases}\dfrac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2}+k_x^2\psi_1=0;\\ \dfrac{\mathrm{d}^2\psi_2}{\mathrm{d}y^2}+k_y^2\psi_2=0;\\ \dfrac{\mathrm{d}^2\psi_3}{\mathrm{d}z^2}+k_z^2\psi_3=0.\end{cases} \qquad (4.16)$$

При цьому відповідно до (4.15) параметри k_x,k_y,k_z повинні задовольняти умову:

$$k_x^2+k_y^2+k_z^2=k^2. \qquad (4.17)$$

Як бачимо, вихідне рівняння (4.14) трансформувалося в систему з трьох “однокоординатних” рівнянь (4.16), кожне з котрих є ідентичним рівнянню (4.5) з усіма наслідками. Зокрема, однокоординатні хвильові функції мають вигляд (4.11), відповідно до чого повні хвильові функції
частинки в тривимірному ящику виражаються, як

$$\psi(x,y,z)=\sqrt{\frac{8}{l_1 l_2 l_3}}\sin\left(\frac{\pi x}{l_1}n_1\right)\cdot\sin\left(\frac{\pi y}{l_2}n_2\right)\cdot\sin\left(\frac{\pi z}{l_3}n_3\right). \qquad (4.18)$$

Так само енергія частинки, згідно з (4.12) і (4.17), виражається, як

$$E=\frac{\hbar^2}{2m}\left(k_x^2+k_y^2+k_z^2\right),$$

або, враховуючи вираз (4.10, як

$$E=\frac{\pi^2\hbar^2}{2m}\left(\frac{n_1^2}{l_1^2}+\frac{n_2^2}{l_2^2}+\frac{n_3^2}{l_3^2}\right),\quad n_1,\ n_2,\ n_3\ =\ 1,\ 2,\ 3,\ \ldots\ , \qquad (4.19)$$

де квантові числа $n_1,\ n_2,\ n_3$ не залежать одне від одного.

Вираз (4.19) є досить прозорим. Якщо в ньому розкрити дужки і врахувати співвідношення (4.13), то вийде:

$$E=\frac{\left(p_x^2+p_y^2+p_z^2\right)}{2m}=\frac{p^2}{2m},$$

де $\left(p_x,p_y,p_z\right)$ – проекції, а p – модуль імпульсу частинки.

Це висвітлює фізичний зміст проведеної при розв’язуванні рівняння (4.14) процедури відокремлення змінних. Вона по суті означає звичний для механіки поділ руху частинки в довільному напрямі на три складові рухи уздовж координатних осей. При цьому кожна однокоординатна хвильова функція ψ_1, ψ_2, або ψ_3 визначає один із таких рухів.

Виродження. Можливість руху в довільному напрямку урізноманітнює квантові стани частинки в тривимірному ящику й ускладнює її енергетичний спектр. Для спрощення розглянемо окремий випадок частинки в кубічному ящику, коли $l_1 = l_2 = l_3 = l$, і енергетичні рівні (4.19) визначаються формулою

$$E = E_0(n_1^2 + n_2^2 + n_3^2), \quad E_0 = \frac{\pi^2\hbar^2}{2ml^2}.$$

Відповідно до цього виразу в таблиці 1 подані енергії та квантові числа для декількох перших енергетичних рівнів частинки в кубічному ящику.

Таблиця 1.

Рівень	Енергія	Квантові числа			Кратність
		n_1	n_2	n_3	K
1	$3E_0$	1	1	1	1
2	$6E_0$	2	1	1	3
		1	2	1	
		1	1	2	
3	$9E_0$	2	2	1	3
		1	2	2	
		2	1	2	
4	$11E_0$	3	1	1	3
		1	3	1	
		1	1	3	
5	$12E_0$	2	2	2	1

Впадає в око, що для деяких рівнів є декілька наборів чисел n_1, n_2, n_3, отже, і хвильових функцій (4.18) та квантових станів частинки. Така ситуація є типовою для квантової механіки, причому

різні квантові стани, що відповідають одному енергетичному рівню, як і самі такі рівні, називаються виродженими.

Відповідно,

кількість станів із заданим значенням енергії називається кратністю виродження K енергетичного рівня. Для невиродженого рівня $K = 1$.

Кратності виродження розглянутих рівнів частинки в тривимірному ящику вказані в останній колонці табл. 1.

Виродження станів квантової частинки в ящику є дуже специфічним ефектом. При кожному значенні енергії та імпульсу існує тільки K різних комбінацій чисел n_1, n_2, n_3 і проекцій (4.13) імпульсу на осі. А це означає, що, хоча всередині ящика на частинку не діють ніякі сили, для неї є лише декілька дозволених напрямків руху. Класична частинка (кулька) за таких умов може рухатися в будь-якому з нескінченної кількості напрямків. Тому в класичній механіці поняття виродження станів позбавлене змісту.

Зауважимо також, що ефект виродження станів є типовим для квантових систем і відіграє істотну роль у різних квантових явищах.

4. ЛІНІЙНИЙ ГАРМОНІЧНИЙ ОСЦИЛЯТОР

Класичний осцилятор. У класичній механіці лінійним гармонічним осцилятором називають матеріальну точку, що здійснює вільні гармонічні коливання вздовж однієї осі (лінії), тобто, рухається за законом:

$$x = x_m \cos(\omega t + \varphi).$$

Цей рух здійснюється під дією сили

$$F_x = m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -m\omega^2 x_m \cos(\omega t + \varphi),$$

або

$$F_x = -kx, \quad k = m\omega^2. \tag{4.20}$$

Така сила є консервативною й утворює потенціальне поле $U(x) = -\int F_x \,\mathrm{d}x$, яке визначається формулою

$$U(x) = \frac{kx^2}{2} \tag{4.21}$$

і на графіку зображується параболою (рис. 4.3). Тому поле виду (4.21) називають параболічним або осциляторним полем.

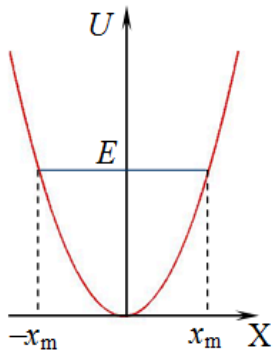


Рис. 4.3

Повна енергія осцилятора зберігається ($E = \text{const}$) і на рис. 4.3 зображується горизонтальним відрізком. Координати точок його перетину з параболою $U(x)$ визначають амплітуду коливань осцилятора: $A = x_\text{m}$. При цьому повна енергія E і амплітуда A залежать тільки від початкових умов (початкового відхилення осцилятора від положення рівноваги або наданої йому початкової швидкості) і можуть бути якими завгодно.

Відповідно до виразу (4.20), частота осцилятора ω визначається силовою константою поля k та масою осцилятора m:

$$\omega = \sqrt{\frac{k}{m}}. \tag{4.22}$$

Враховуючи цю формулу, вираз потенціальної енергії осцилятора (4.21) можна записати у вигляді:

$$U(x) = \frac{m\omega^2 x^2}{2}. \tag{4.23}$$

Квантовий осцилятор. Енергетичний спектр. Для мікрочастинок сила не є визначеною, тому в квантовій механіці лінійний гармонічний осцилятор означають як частинку, що знаходиться в параболічному силовому полі (4.23). Відповідно, рівняння Шрьодінґера (4.4) для лінійного осцилятора має вигляд:

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}+\frac{2m}{\hbar^2}\left(E-\frac{m\omega^2x^2}{2}\right)\psi=0\,. \qquad (4.24)$$

Розв'язування цього рівняння та його розв'язки – хвильові функції квантового осцилятора – є досить складними. Але й без строгого математичного дослідження можна передбачити, що стаціонарні стани квантового осцилятора є квантованими. Справді, при віддаленні від положення рівноваги потенціальна енергія осцилятора зростає необмежено, тож він ні за яких умов не може необмежено віддалитися від положення рівноваги $x=0$. Отже, осцилятор можна розглядати як частинку в потенціальному ящику з параболічними стінками, тому його стани мають бути квантованими, як про це говорилося в п. 4.2.

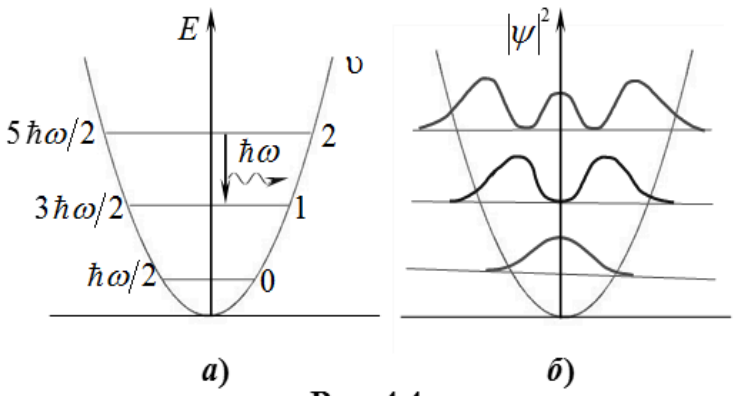
Математика підтверджує такий прогноз: однозначні й неперервні розв'язки рівняння (4.24) – хвильові функції осцилятора $\psi_\upsilon(x)$ – існують лише при окремих значеннях параметра E, що утворюють дискретний енергетичний спектр осцилятора відповідно до умови:

$$E_\upsilon=\left(\upsilon+\frac{1}{2}\right)\hbar\omega,\quad \upsilon=0,1,2,\ldots, \qquad (4.25)$$

де υ – квантове число осцилятора, або ***коливальне квантове число***, а величина $\omega=\sqrt{k/m}$ по аналогії з класичним осцилятором називається циклічною частотою квантового осцилятора.

На рис. 4.4а показана картина енергетичних рівнів, яка ілюструє типові риси та деякі особливості енергетичного спектра квантового гармонічного осцилятора. Зокрема, вираз (4.25) показує, що осцилятор, як і частинка в прямокутному ящику, не може перебувати у стані спокою. Мінімальна енергія його коливань (енергія основного стану) дорівнює

$$E_0=\frac{1}{2}\hbar\omega. \qquad (4.25a)$$



а) *б)*

Рис. 4.4

Такими квантовими осциляторами є атоми в твердих тілах і рідинах, оскільки їхній тепловий рух являє собою малі коливання навколо положень рівноваги. Отже навіть при температурі T = 0 К рух атомів не припиняється – вони здійснюють так звані «нульові коливання» з енергією E_0. Але атоми не можуть передавати цю енергію один одному, тому квантові нульові коливання не мають нічого спільного з тепловим рухом атомів і не суперечать поняттю абсолютного нуля температури, так само як не суперечить йому рух електронів всередині атомів.

Енергетичні рівні збуджених станів осцилятора, які можна записати у вигляді

$$E_\upsilon = E_0 + n\hbar\omega, \quad n = 1,2,\ldots,$$

є еквідистантними – вони розташовані в енергетичній шкалі на однаковій відстані (рис. 4.4а)

$$\Delta E = E_{\upsilon+1} - E_\upsilon = \hbar\omega. \tag{4.26}$$

Отже, при переході гармонічного осцилятора з будь-якого енергетичного рівня на сусідній нижчий рівень згідно із законом збереження енергії, випромінюється порція (квант) енергії $\hbar\omega$. Це пояснює квантову гіпотезу Планка та правило частот Бора.

Теорія і дослід показують, що можливі тільки такі зміни стану осцилятора, при яких коливальне квантове число змінюється на одиницю:

$$\Delta\upsilon = \pm 1.$$

Згідно з цим "правилом відбору", можливі переходи лише між сусідніми енергетичними рівнями осцилятора, і згідно з (4.26), атоми твердого тіла мали б обмінюватися квантами енергії однієї єдиної частоти. Але насправді так відбувається тільки при дуже низьких температурах. У загальному ж випадку в спектрах коливань атомів присутні цілі смуги близьких частот. Це пов'язано з ангармонізмом – відхиленням взаємодії між атомами від гармонічного закону (4.21) при підвищених температурах. Слід зазначити, що величина квантів коливальної енергії набагато менша, ніж енергія, що випромінюється чи поглинаються при переходах електронів між рівнями в атомах. Тому частоти коливальних переходів атомів лежать у далекій інфрачервоній області спектра.

Атоми здійснюють коливання не тільки в конденсованій речовині, а і в окремих молекулах. Тож при зміні стану молекули відбувається не тільки перехід електронів в атомах з одних електронних рівнів на інші, а й перехід самих атомів із одних коливальних рівнів на інші. Через це при переходах електронів у молекулах між заданими двома рівнями випромінюються (чи поглинаються) фотони не з однією частотою, а з цілою низкою близьких частот, які відповідають переходам атомів між різними коливальними рівнями. Тому молекулярні спектри є не лінійчастими, як атомарні, а смугастими.

Принагідно зазначимо, що експериментальні дослідження молекулярних спектрів не лише підтверджують квантову теорію осцилятора, а й дають цінну інформацію про взаємодію та розташування атомів у молекулах.

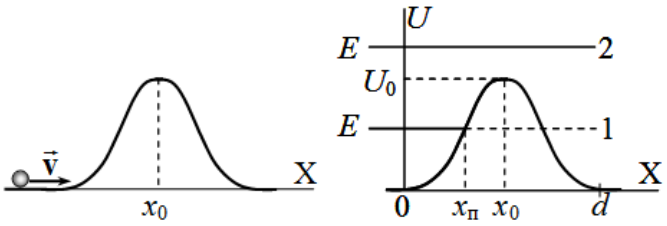
Хвильові функції. Із рівняння (4.24) можна отримати аналітичні вирази хвильових функцій гармонічного осцилятора $\psi_\upsilon(x)$, і визначити розподіл ймовірностей $|\psi(x)|^2$ знаходження осцилятора в різних точках простору для різних квантових станів.

Для перших трьох станів ($\upsilon = 0, 1, 2$) графіки $|\psi(x)|^2$ показано на рис. 4.4б. Загалом вони нагадують аналогічні графіки рис. 4.2б для частинки в ящику з вертикальними стінками. Зокрема, в області руху існують точки з максимальною імовірністю знаходження осцилятора та «заборонені точки», в яких осцилятор виявити неможливо. Проте є й істотна відміна: графіки $|\psi(x)|^2$ виходять за межі графіка $U(x)$, що означає можливість перебування осцилятора в точках з координатами $|x| > x_{\text{m}}$. У таких точках на графіку $E < U$, що дає для кінетичної енергії неможливі значення $(p^2/2m) = E - U < 0$. Але цей парадокс є удаваним і пов'язаний із принципом невизначеності. Дійсно, кінетична енергія K визначається імпульсом, а потенціальна U – координатами, тобто величинами, що одночасно на можуть бути точно визначеними (див. лекція 3, п. 3.4). Тому поділ повної енергії $E = K + U$ на кінетичну та потенціальну є певною мірою умовним. При цьому, як доводять розрахунки, у "забороненій зоні" квантовомеханічна невизначеність кінетичної енергії осцилятора ΔK ніде не перевищує «дефіцит» $U - E$. Тому можливість виходу осцилятора в класично заборонену зону $|x| > x_{\text{m}}$ зовсім не означає від'ємність його кінетичної енергії.

5. ТУНЕЛЬНИЙ ЕФЕКТ

Бар'єри. Якщо на шляху частинки трапляється силове поле, що перешкоджає її рухові, то таке поле називають ***потенціальним бар'єром***.

Наприклад, гірка на шляху кульки, що вільно рухається по горизонтальній поверхні (рис. 4.5*а*), створює потенціальний бар'єр, показаний на рис. 4.5*б* графіком потенціальної енергії кульки $U(x)$ в залежності від координати. Бар'єр характеризується ***висотою*** U_0 (максимальною величиною потенціальної енергії) та шириною d (розміром області, в якій існує силове поле; зазвичай ця область не визначена точно, тому в кожному випадку вводять певну ефективну ширину бар'єра).



б)

Рис. 4.5

При розгляді руху частинки в напрямку бар'єра виникає природнє запитання – чи подолає вона бар'єр і буде продовжувати рух у тому ж напрямку, чи ні? Для класичної частинки відповідь є однозначною й залежить від співвідношення між повною енергією частинки та висотою бар'єра. При русі частинки в потенціальному полі повна енергія, що складається з кінетичної та потенціальної, зберігається: $E = K + U = \text{const}$. Тому, якщо $E < U_0$ (горизонтальна лінія 1 на рис. 4.5*б*), то частинка «відіб'ється» від бар'єра, тобто дістанеться тільки точки $x = x_1$, де $U = E$ і $K = 0$, і почне рухатись у зворотному напрямі. Якщо ж $E > U_0$ (лінія 2), то навіть при $x = x_0$ (на вершині гірки) кінетична енергія кульки $K = E - U_0 > 0$. Тож при $E > U_0$ кулька обов'язково подолає бар'єр і продовжить рух далі. В окремому випадку $E = U_0$ в точці $x = x_0$ кінетична енергія $K = 0$, тобто кулька дістанеться вершини і зупиниться. Отже, формально вона подолає бар'єр, бо не буде відкинута назад. Але рівновага кульки в точці максимуму потенціальної енергії (на вершині гірки) є нестійкою, і в дійсності вона неконтрольовано скотиться в той чи інший бік. Тому цей випадок не являє інтересу.

Розглянемо тепер рух квантової частинки за наявності потенціального бар'єра. Для цього треба скласти та розв'язати рівняння Шрьодінґера (4.4). Аби максимально спростити математику, розглянемо бар'єр у вигляді прямокутної сходинки (рис. 4.6). У такому ідеалізованому полі потенціальну енергію в усьому просторі неможливо виразити аналітично, тобто якоюсь однією функцією $U(x)$. Тому рух частинки до бар'єра (область 1, $x < 0$) і в області бар'єра (область 2, $x \ge 0$) описують окремими рівняннями (4.5):

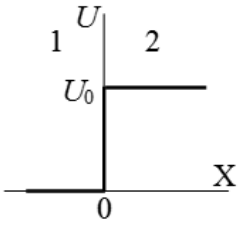


Рис. 4.6

$$\begin{aligned} &\textit{область}\ 1: \quad \psi_1'' + k_1^2 \psi_1 = 0, \quad k_1 = \sqrt{2mE}\,/\hbar; \\ &\textit{область}\ 2: \quad \psi_2'' + k_2^2 \psi_2 = 0, \quad k_2 = \sqrt{2m(E-U_0)}\,/\hbar. \end{aligned} \qquad (4.26)$$

При цьому, позаяк $\psi_1(x)$ і $\psi_2(x)$ є "половинками" хвильової функції тієї самої частинки, вони на межі областей ($x=0$) мають «зшиватися» – гладко переходити одна в одну, тобто задовольняти умови

$$\psi_1(0)=\psi_2(0); \quad \psi_1'(0)=\psi_2'(0). \tag{4.27}$$

Із математики відомо, що загальні розв'язки рівнянь (4.26) мають вигляд

$$\begin{aligned} \psi_1 &= A_1 \mathrm{e}^{ik_1x} + B_1 \mathrm{e}^{-ik_1x}, \\ \psi_2 &= A_2 \mathrm{e}^{ik_2x} + B_2 \mathrm{e}^{-ik_2x}. \end{aligned} \tag{4.28}$$

Фізичний зміст цих розв'язків стає зрозумілим, якщо записати повні вирази $\Psi(x,t)$ функцій (4.28) за формулою (4.3)

$$\Psi(x,t) = A\mathrm{e}^{-i(\omega t - kx)} + B\mathrm{e}^{-i(\omega t + kx)}.$$

Така функція описує сукупність двох хвиль, які поширюються у протилежних напрямках осі X. Отже у виразах (4.28) перший доданок функції ψ_1 відповідає рухові частинки у напрямку бар'єра, а другий – її рухові у зворотньому напрямку після відбивання від стінки бар'єра. Так само перший доданок у ψ_2 відповідає рухові частинки в області бар'єра після подолання стінки. А от другого доданку в ψ_2, себто зворотньої хвилі, не може бути, бо далі ніяких перешкод на шляху частинки немає. Тому $B_2=0$, і $\psi_2 = A_2 \mathrm{e}^{ik_2x}$.

Суттєво також, що реально прослідкувати за окремою частинкою в потоці, що падає на бар'єр, неможливо. Можна говорити лише про коефіцієнти відбивання та проходження – відсотки частинок, які відбиваються від бар'єра та долають його. Ці коефіцієнти визначаються відносними амплітудами хвильових функцій $b_1 = B_1/A_1$ і $a_2 = A_2/A_1$. Відтак розв'язки (4.28) можна подати у вигляді

$$\begin{aligned} \psi_1 &= \mathrm{e}^{ik_1x} + b_1 \mathrm{e}^{-ik_1x}, \\ \psi_2 &= a_2 \mathrm{e}^{ik_2x}. \end{aligned} \tag{4.28a}$$

Наклавши на ці функції умови «зшивання» (4.27), визначимо відносні амплітуди:

$$\begin{aligned} 1+b_1 &= a_2, \\ k_1(1-b_1) &= k_2 a_2; \end{aligned} \quad \Rightarrow \quad a_2 = \frac{2k_1}{k_1+k_2}, \quad b_1 = \frac{k_1-k_2}{k_1+k_2}. \tag{4.29}$$

Ці вирази дозволяють аналізувати поведінку потоку частинок, що налітають на бар'єр, у залежності від значень k_1 і k_2 (формули (4.26)), тобто від енергії частинок E та висоти бар'єра U_0. Зосередимося лише на випадку $E < U_0$. За такої умови величина k_2 є уявною, і її доцільно записати як

$$k_2 = ik, \ \text{äå} \ \ k = \frac{1}{\hbar}\sqrt{2m(U_0 - E)}, \ i = \sqrt{-1}. \tag{4.30}$$

Тоді амплітуда b_1 у виразах (4.29) дорівнює:

$$b_1 = \frac{k_1 - ik}{k_1 + ik}.$$

При цьому ймовірність відбивання

$$\left| b_1 \mathrm{e}^{-ik_1 x} \right|^2 = b_1 \mathrm{e}^{-ik_1 x} \cdot b_1^* \mathrm{e}^{ik_1 x} = b_1 b_1^* = \frac{k_1 - ik}{k_1 + ik} \cdot \frac{k_1 + ik}{k_1 - ik} = 1,$$

тож всі частинки, що налітають на бар'єр, відбиваються від нього. Оскільки $E < U_0$, такий результат є цілком очікуваним. Але при цьому, згідно з виразами (4.28*a*) і (4.30), $\psi_2 = a_2 \mathrm{e}^{-kx}$, де k – дійсне число. Тому

$$\left|\psi_2\right|^2 = \left|a_2\right|^2 \cdot \mathrm{e}^{-2kx} \neq 0. \tag{4.31}$$

Це означає, що існує ймовірність проникнення частинки в заборонену в класичній механіці область "під бар'єром", де кінетична енергія $K = E - U_0 < 0$. Пояснення цього суто квантового феномена таке саме, як і у випадку квантового осцилятора (див. п. 4.4).

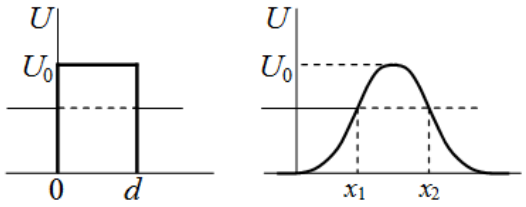
Тунельний ефект. Як випливає з (4.31), імовірність заглиблення частинки під бар'єр стрімко спадає із збільшенням відстані x. Тому частинки, після подолання стінки заглиблюються в область бар'єра тільки на дуже малу відстань і повертаються назад. При цьому глибина заглиблення тим більша, чим менша величина k (4.31), тобто чим більша початкова кінетична енергія $K_0 = E$ (і швидкість) частинки. Описана поведінка квантової частинки зовні нагадує відбивання від поверхні води легкої кульки, що падає з деякої висоти: вона спочатку занурюється на певну глибину, а потім вискакує назад. При цьому глибина занурення зростає при збільшенні швидкості входження кульки у воду.

Якщо на шляху частинки трапляється бар'єр скінченної, причому дуже малої ширини d, існує помітна імовірність того, що вона заглибиться в область бар'єра на відстань $x = d$. В такому разі частинка подолає бар'єр і зможе рухатися далі, хоча при $E < U_0$ її енергія для цього формально є недостатньою.

Проходження частинок крізь бар'єр при енергії, меншій за його висоту, називається тунельним проходженням або тунельним ефектом.

Імовірність тунельного проходження, інакше – *прозорість* бар'єра D, є пропорційною ймовірності проникнення частинки під бар'єр на відстань $x = d$ і для прямокутного бар'єра (рис. 4.7*а*) наближено виражається формулою

$$D \approx e^{-\frac{2}{\hbar}\sqrt{2m(U_0-E)}\cdot d} \qquad (4.32)$$



а) ***б)***

Рис. 4.7

У випадку бар'єра довільної заданої форми $U = U(x)$ (рис. 4.7*б*) прозорість виражається (див. п. 4.7, задача 4.5), як

$$D \approx e^{-\frac{2}{\hbar}\int_{x_1}^{x_2}\sqrt{2m(U(x)-E)}\cdot dx} \qquad (4.32a)$$

З наведених формул видно, що ймовірність тунелювання є дуже чутливою до маси частинки та характеристик бар'єра, зокрема, ширини. Наприклад, для електронів при *ефективній висоті* прямокутного бар'єра

$U_0 - E = 1{,}0$ еВ розрахунки за формулою (4.32) для різної ширини d дають:

d, нм	0,2	0,5	1,0	2,0
D	0,135	0,007	$4{,}5\cdot10^{-5}$	$2{,}1\cdot10^{-10}$

Отже, прозорість бар'єра є помітною лише при гранично малій його ширині, і може здатися, що на практиці спостерігати тунельний ефект неможливо. Але це не так. Достатньо вузькі й помірно високі бар'єри створюються для валентних електронів ядрами атомів у твердих тілах. При цьому через велику концентрацію електронів ($\sim 10^{23}$ 1/м^3), навіть при дуже малій прозорості бар'єрів відбувається велика кількість тунельних переходів. Це зумовлює низку характерних тунельних явищ, які не лише спостерігаються в експерименті, а й використовуються на практиці.

6. КОНТРОЛЬНІ ЗАПИТАННЯ

1. Отримайте рівняння Шрьодінгера для стаціонарних станів (4.4) із загального рівняння (4.2) за допомогою підстановки (4.3).
2. Чим поведінка мікроскопічної частинки в потенціальному ящику принципово відрізняється від руху маленької кульки між стінками всередині маленької коробочки?
3. Отримайте енергетичний спектр (4.12) частинки в одновимірному ящику за допомогою співвідношень де-Бройля (3.3).
4. Доведіть за допомогою принципу невизначеності, що мікрочастинка в потенціальному ящику не може перебувати в стані спокою.
5. Оцініть найменшу можливу енергію частинки в одновимірному потенціальному ящику за допомогою принципу невизначеності.
6. Запишіть формулу енергетичних рівнів частинки маси m, яка знаходиться в двовимірному квадратному потенціальному ящику зі стороною a. Чому дорівнює енергія основного стану такої частинки?
7. Які енергетичні рівні називаються виродженими? Що таке кратність виродження?
8. Чому дорівнює максимальна можлива кратність виродження енергетичного рівня для частинки в квадратному потенціальному ящику?
9. Згідно з (4.25) і (4.25a), атоми твердого тіла, котрі можна трактувати як квантові осцилятори, навіть при температурі T = 0 К перебувають у русі. Як це узгоджується з уявленням про абсолютний нуль температури як температуру, при якій припиняється тепловий рух частинок речовини?
10. Скільки частот можна спостерігати в спектрі випромінювання лінійного гармонічного квантового осцилятора?
11. Чим принципово відрізняється поведінка квантової частинки та матеріальної точки (класичної частинки), коли на шляху їхнього руху трапляється потенціальний бар'єр?
12. Чим відрізняється поведінка класичної та квантової частинок при зіткненні з потенціальним бар'єром типу сходинка, висота якого більша за кінетичну енергію частинки?
13. У чому полягає тунельний ефект? Що називається прозорістю бар'єра?
14. Як і в скільки разів зміниться потік частинок, які тунелюють крізь прямокутний бар'єр, якщо в 10 разів збільшиться: а) ширина бар'єра; б) його ефективна висота.
15. Порівняйте імовірності тунельного проходження крізь один і той самий бар'єр для електрона (m_e = 9,1·10^{-31} кг) і для та протона (m_p = 1,66·10^{-27} кг).

6.5. АТОМ ГІДРОҐЕНУ

Атоми – найменші частинки речовини, що відображають усі фізико-хімічні властивості хімічних елементів – мають мікроскопічні розміри і маси . Тому послідовна теорія атома можлива лише на основі квантової механіки.

Найпростішим із атомів є атом Гідроґену. Він складається з одного електрона, котрий під дією кулонівської сили рухається навколо ядра, яке через велику масу можна вважати нерухомим. За допомогою рівняння Шрьодінґера задача про атом Гідроґену розв’язується точно, але математично є досить складною. Тому далі розглядаються тільки кінцеві теоретичні результати та обговорюється їхній зміст.

1. Квантові числа

2. Енергетичний та оптичний спектри атома Гідроґену

3. Квантування моменту імпульсу

4. Виродження енергетичних рівнів

5. Розподіл електронної густини в атомі Гідроґену

6. Орбітальний магнітний момент електрона

7. Контрольні запитання

1. КВАНТОВІ ЧИСЛА

Електрон в атомі Гідроґену знаходиться в полі ядра

$$U(\vec{r}) = k\frac{q_e q_{\ddot{y}}}{r} = -k\frac{e^2}{r}, \quad k = \frac{1}{4\pi\varepsilon_0},$$

де r – відстань електрона від центра ядра.
Це поле є центральним, тому задачу розглядають у сферичних координатах, і рівняння Шрьодінґера (4.4) має вигляд:

$$\nabla^2\psi + \frac{2m}{\hbar^2}\left(E + \frac{ke^2}{r}\right)\psi = 0, \quad k = \frac{1}{4\pi\varepsilon_0}. \qquad (5.1)$$

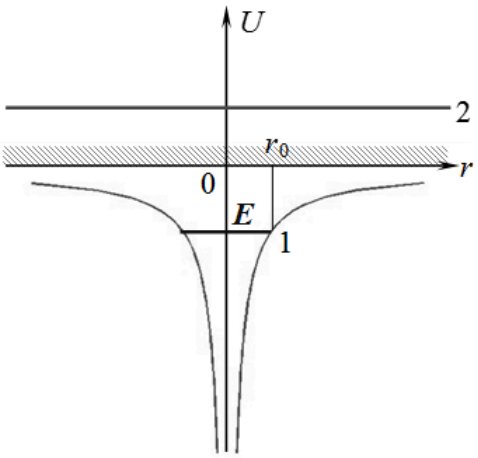


Рис. 5.1

На рис. 5.1 показаний графік залежності $U(r)$ у будь-якій площині, що проходить через центр ядра. Розглядаючи його, можна зробити деякі правдоподібні припущення щодо поведінки електрона, не розв'язуючи рівняння (5.1). А саме. Якщо повна енергія електрона $E < 0$, (лінія 1 на рис. 5.1) то при віддаленні від ядра на відстань r_0 кінетична енергія електрона $K = E - U$ поступово зменшується аж до нуля так, що він не може віддалитися від ядра значно далі, ніж на відстань $r = r_0$. Отже, при $E < 0$ електрон є зв'язаним із ядром, іншими словами, – знаходиться в потенціальній ямі. Тому його стани мають бути квантованими. А ось при $E \geq 0$ (лінія 2 на рис. 5.1) електрон має достатню кінетичну енергію, аби віддалитися від ядра на необмежену відстань. Такі стани є не квантованими і відповідають іонізації атома.

Розв'язування рівняння Шрьодінґера для атома Гідроґену підтверджує сказане: однозначні та неперервні і гладкі розв'язки рівняння (5.1) існують при будь яких значеннях енергії $E \geq 0$, але лише при дискретних значеннях $E = E_n$, якщо $E < 0$. Отже, не іонізований атом може перебувати тільки в деяких дозволених квантових станах. Дозволеним станам відповідає дискретний набір хвильових функцій $\psi(\vec{r}) = \psi_{n,l,m_l}(\vec{r})$, кожна з яких визначається трьома квантовими числами – ***головним n, орбітальним l і магнітним m_l***. При цьому, позаяк рух електрона в полі ядра не є вільним, існує певна ієрархія квантових чисел. А саме. Головне квантове число n може набувати будь-яких значень. Натомість можливі значення l обмежені величиною n, а значення m_l – величиною l, відповідно до наступних умов:

$$\begin{aligned} &n = 1, 2, 3, \ldots; \\ &l = 0, 1, 2, \ldots, (n-1) \quad - \quad \textit{всього } n \textit{ значень}; \\ &m_l = 0, \pm 1, \pm 2, \ldots, \pm l \quad - \quad \textit{всього } 2l+1 \textit{ значень}. \end{aligned} \qquad (5.2)$$

2. ЕНЕРГЕТИЧНИЙ ТА ОПТИЧНИЙ СПЕКТРИ АТОМА ГІДРОҐЕНУ

Головне квантове число ***n*** визначає можливі значення енергії (енергетичний спектр) атома Гідроґену, згідно із формулою:

$$E_n = -\frac{k^2 m e^4}{2\hbar^2}\frac{1}{n^2}, \qquad k = \frac{1}{4\pi\varepsilon_0}. \tag{5.3}$$

При розрахунках зручно користуватися виразом (5.3) в числовому вигляді:

$$E_n = -\frac{13{,}6}{n^2}\ (\text{еВ}). \tag{5.3a}$$

Принагідно відмітимо, що іони, котрі мають тільки один електрон (так звані водневоподібні іони) відрізняються від атома Гідроґену лише величиною заряду ядра: $q_{\text{ў}} = Ze$, де Z – порядковий номер елемента в таблиці Менделеєва. Тому вони мають аналогічний енергетичний спектр:

$$E_n = -\frac{k^2 m Z^2 e^4}{2\hbar^2}\frac{1}{n^2}. \tag{5.3$б$}$$

Схема енергетичних рівнів атома Гідроґену показана на рис. 5.2. Енергія основного рівня $(n=1)$ складає $E_1 = -13{,}6$ еВ. Щоб іонізувати атом (відривати електрон від ядра), електрон необхідно перевести в стан з енергією $E \geq 0$. Мінімальна необхідна для цього енергія, що називається ***енергією іонізації***, дорівнює $E_{\text{і}} = -E_1 = 13{,}6$ еВ. Це теоретичне значення збігається з величиною, знайденою дослідним шляхом.

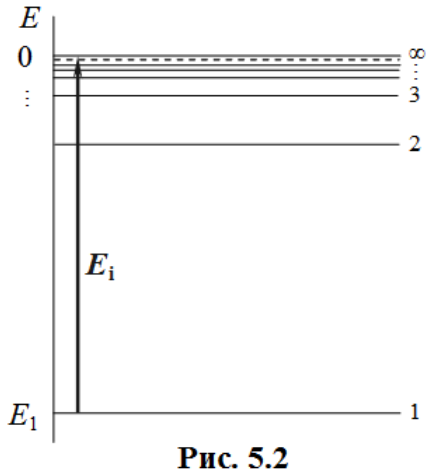


Рис. 5.2

З енергетичного спектра можна отримати й оптичний спектр атомарного Гідроґену – визначити всі частоти або довжини хвиль, які можуть випромінювати збуджені атоми. Енергія фотона, що випромінюється, дорівнює зміні енергії атома при переході електрона з якогось рівня n_2 на будь-який інший рівень $n_1 < n_2$: $\hbar\omega = E_{n_2} - E_{n_1}$. Тож із виразу (5.3) випливає наступна загальна спектральна формула:

$$\omega = \frac{k^2 m e^4}{2\hbar^3}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right). \tag{5.4}$$

Отже, спектр атома складається з окремих частот (спектральних ліній), тобто є лінійчастим. Вираз (5.4) записують компактніше, увівши так звану ***сталу Ридберґа***:

$$R = \frac{k^2 m e^4}{2\hbar^3} = 2{,}07 \cdot 10^{16} \text{ рад/с}.$$

Відтак

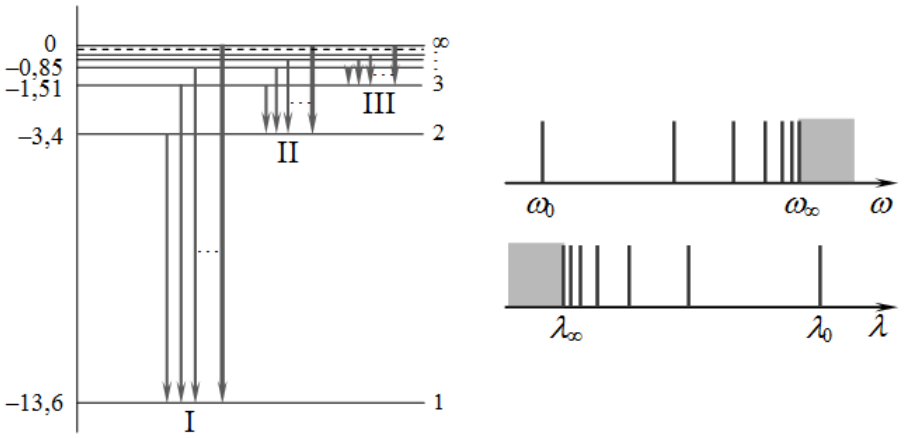
$$\omega = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right). \qquad (5.5)$$

Зауважимо, що ця формула, котра називається ***узагальненою формулою Бальмера***, була відома з дослідів ще до створення квантової механіки. При цьому теоретичне та експериментальне значення сталої Ридберга R збігаються. Зважаючи на високу точність спектроскопічних вимірювань, це свого часу стало важливим свідченням достовірності рівняння Шрьодінґера.

Для водневоподібних іонів спектральна формула має вигляд

$$\omega = RZ^2\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right), \qquad (5.5a)$$

де Z – порядковий номер елемента в періодичній системі Менделеєва.



а) ***б)***

Рис. 5.3

Згідно з (5.5) і (5.5а), спектральні лінії (частоти) випромінювання Гідроґену та воднево-подібних іонів природньо групуються у ***спектральні серії*** – виділені сукупності, зумовлені переходами електрона на заданий рівень n_i з усіх більш високих рівнів $n > n_i$, як показано на рис. 5.3а для перших трьох серій. Лінії кожної серії в шкалі частот (або довжин хвилі) розташовані упорядковано й однотипно, як це схематично показано на рис. 5.3б. Серія починається *головною* лінією, що зумовлена переходом електрона на заданий рівень n_i з наступного рівня $n_i + 1$. Ця лінія має найменшу частоту ω_0 і найбільшу довжину хвилі λ_0 в серії. Далі частоти поступово збільшуються, а довжини хвилі зменшуються так, що спектральні лінії подібно до енергетичних рівнів, поступово зближуються і при частоті ω_∞ (або довжині хвилі λ_∞) переходять у суцільну смугу. Величини ω_∞ і λ_∞ визначають *короткохвильову межу* серії й у виразі (5.5) відповідають умові $n_2 \to \infty$. (*Зауважимо, що через обмежену роздільну здатність ока лінійчастий спектр насправді не має чіткої межі і переходить у неперервний при скінченній величині* n_2).

Для ілюстрації нижче наведені конкретні спектральні формули та величини λ_0 і λ_∞ для перших трьох серій атома водню: 1) серії Лаймана ($n_1 = 1$), 2) серії Бальмера ($n_1 = 2$) і 3) Серія Пашена ($n_1 = 3$):

1. $$\omega = R\left(1 - \frac{1}{n^2}\right),\ n = 2,3,\ldots,\infty \quad \Rightarrow \quad \omega_0 = \frac{3R}{4},\ \omega_\infty = R \quad \Rightarrow$$
$$\Rightarrow \lambda_0 \approx 121\ \text{нм},\ \lambda_\infty \approx 91\ \text{нм}.$$

2. $$\omega = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right),\ n = 3,4,\ldots,\infty \quad \Rightarrow \quad \omega_0 = \frac{5R}{36},\ \omega_\infty = \frac{R}{4} \quad \Rightarrow$$
$$\Rightarrow \lambda_0 \approx 656\ \text{нм},\ \lambda_\infty \approx 364\ \text{нм}.$$

3. $$\omega = R\left(\frac{1}{3^2} - \frac{1}{n^2}\right),\ n = 4,5,\ldots,\infty \quad \Rightarrow \quad \omega_0 = \frac{7R}{144},\ \omega_\infty = \frac{R}{9} \quad \Rightarrow$$
$$\Rightarrow \lambda_0 \approx 1873\ \text{нм},\ \lambda_\infty \approx 819\ \text{нм}.$$

Із наведених числових величин видно, що вся перша спектральна серія гідроґену знаходиться в ультрафіолетовій, а вся третя серія (і наступні) – в інфрачервоній області спектра, тому безпосередньо спостерігати їх оком не можна. Натомість частина ліній другої (бальмерівської) серії потрапляє у видиму область і є доступною для візуального спостереження.

На завершення зауважимо, що такі самі результати стосовно енергетичного та оптичного спектрів атома гідроґену дає і внутрішньо суперечлива напівкласична борівська теорія. Але на відміну від неї, квантовомеханічний розгляд є послідовним, не вимагає спеціальних припущень на кшталт постулатів Бора і дає багато іншої інформації про атом, яка є недоступною для борівської теорії, тим паче – класичної механіки.

3. КВАНТУВАННЯ МОМЕНТУ ІМПУЛЬСУ

Орбітальний момент імпульсу електрона. Рух класичної частинки по замкненій траєкторії (орбіті) визначається вектором моменту імпульсу, що дорівнює векторному добутку радіуса-вектора частинки та вектора імпульсу:

$$\vec{L} = [\vec{r}\,\vec{p}] \equiv [\vec{r}, m\vec{v}].$$

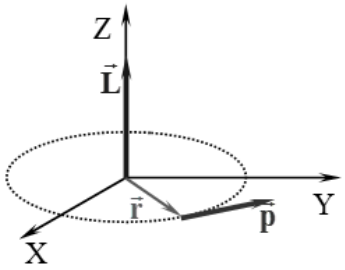


Рис. 5.4

Вектор $\vec{L}$ напрямлений перпендикулярно до площини орбіти (рис. 5.4), тож визначає її просторову орієнтацію.

Рух електрона в атомі навколо ядра (*орбітальний* рух) теж є замкненим і характеризується ***орбітальним моментом імпульсу***. При цьому з рівняння Шрьодінґера (5.1) випливає, що орбітальний момент імпульсу електрона в атомі, як і енергія, квантується Зауважимо, що поширені вирази «орбітальний рух» і «орбітальний момент» не слід сприймати буквально, позаяк електрон в атомі не має визначеної траєкторії руху, тобто орбіти.

Просторове квантування. Розв'язки рівняння Шрьодінґера показують, що не тільки модуль, а й проекція моменту імпульсу L_z є квантованою. При цьому можливі значення L_z задаються магнітним квантовим числом m_l відповідно до ***умови квантування проекції моменту імпульсу***:

$$L_z = \hbar m_l, \quad m_l = 0, \pm 1, \pm 2, \cdots ..., \pm l. \tag{5.7}$$

Квантування проекції означає, що вектор моменту імпульсу $\vec{L}$ може бути напрямленим тільки під такими кутами ϑ до осі OZ, для яких

$$\cos\vartheta = \frac{L_z}{L} = \frac{m_l}{\sqrt{l(l+1)}}. \tag{5.8}$$

Іншими словами – можливі не всі орієнтації вектора моменту імпульсу в просторі. Тому ефект квантування проекції моменту імпульсу називають ***просторовим квантуванням***.

Із умови (5.7) випливає, що можливі орієнтації моменту імпульсу є симетричними відносно площини, перпендикулярної до осі просторового квантування OZ і їхня кількість дорівнює $2l+1$. Наприклад, при $l = 1$ $m_l = 0, \pm 1$. При цьому проекція може мати тільки значення $L_z = 0$ і $\pm\hbar$, які відповідають кутам $\vartheta = 90°, 45° \text{ і } 135°$ (рис. 5.5).

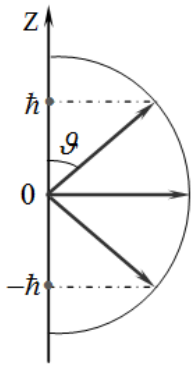


Рис. 5.5

Просторове квантування ставить два принципові запитання. Перше. Відомо, що для однозначного визначення вектора у тривимірному просторі треба задати три числа, наприклад, проекції на три осі координат, або модуль і проекції на дві осі. Натомість умови квантування (5.6) і (5.7) дають тільки дві величини – L і L_z і відтак визначають не один, а нескінченну множину векторів $\vec{L}$, напрямлених під заданим кутом ϑ до осі OZ. Тому може здатися, що рівняння Шрьодінґера не дає всіх відомостей про орбітальний момент імпульсу електрона в атомі. Але через принцип невизначеності повніша інформація є принципово недоступною. Дійсно, уявімо що напрям вектора $\vec{L}$ відомий точно, як у випадку класичної частинки (рис. 5.4). Тоді, спрямувавши координатну вісь OZ уздовж $\vec{L}$, можна було би твердити, що імпульс електрона $\vec{p}$ перпендикулярний до осі OZ і траєкторія лежить у площині XOY . В такому разі координата z і проекція імпульсу p_z в будь-який момент часу мали би точні значення $z = 0$, $p_z = 0$, що суперечить принципу невизначеності (3.11). Відтак маємо зробити наступний висновок:

момент імпульсу мікрочастинки принципово не може мати однозначно заданого напрямку.

Друге питання стосується напрямку осі OZ, відносно якої відбувається просторове квантування. Позаяк вільний простір є ізотропним, то всі напрямки в ньому рівноправні, і координатну вісь OZ (вісь просторового квантування) можна спрямувати, як забажається. Відтак напрошується парадоксальна думка, що напрямки вектора $\vec{L}$ електрона залежить від нашого бажання. Але ніякого парадоксу насправді немає, бо запитання про напрям моменту імпульсу електрона апріорі є беззмістовним. Аби взнати, яким є вектор $\vec{L}$, треба провести відповідні вимірювання, тобто здійснити контрольований вплив на атом і зареєструвати реакцію (відгук) атома на цей вплив, у чому й полягає сутність процесу вимірювання як такого. Вказаний вплив здійснює силове поле, що створюється вимірювальним приладом або установкою в процесі вимірювання. Але в такому разі в просторі з'являється фізично виділений напрям – напрям створеного в приладі силового поля. Саме відносно цього виділеного напряму і відбувається просторове квантування.

Наостанку скажемо, що умови (5.6), (5.7) є універсальними – вони виконуються не лише для орбітального моменту імпульсу електрона в атомі Гідроґену, а й для моменту імпульсу будь-якого іншого походження. Тож момент імпульсу за самою природою є дискретною (квантованою) фізичною величиною, так, як, скажімо, електричний заряд. При цьому стала Планка $\hbar$, подібно до елементарного заряду e, виступає як природня міра (квант) моменту імпульсу. Але через гранично малу величину $\hbar$ дискретність моменту імпульсу є суттєвою лише для елементарних частинок.

4. ВИРОДЖЕННЯ ЕНЕРГЕТИЧНИХ РІВНІВ

Кратність виродження рівня. Згідно з формулою (5.3), енергія електрона в атомі Гідроґену залежить тільки від головного квантового числа n. Натомість хвильові функції, тож і квантові стани, визначаються ще й орбітальним l і магнітним m_l числами, котрі за умовами (5.2) при заданому n можуть мати декілька різних значень. Тому квантові стани та енергетичні рівні електрона в атомі Гідроґену є виродженими (див. Лекція 4, п. 4.3). Виняток становить основний рівень, якому відповідає єдиний квантовий стан ($n=1$, $l=0$, $m_l=0$).

Підрахуємо кратність виродження K енергетичних рівнів у атомі Гідроґену. Згідно із (5.2), (5.6) і (5.7), кожному можливому значенню модуля моменту імпульсу електрона L і квантового числа l відповідає $N_l=2l+1$ квантових станів, які відрізняються значеннями L_z, тобто – орієнтацією моменту імпульсу відносно осі квантування. Отже, загальна кількість станів K, які відповідають заданому енергетичному рівню, складає

$$K=\sum_l N_l=\sum_0^{n-1}(2l+1).$$

Доданки в цій сумі утворюють арифметичну прогресією з першим членом 1 і останнім $2n-1$. Тому, згідно з відомою формулою суми арифметичної прогресії, виходить

$$K=n^2. \qquad (5.9)$$

Одразу зауважимо, що в дійсності кількість станів, які відповідають заданому значенню n, є вдвічі більшою:

$$N=2n^2. \qquad (5.9a)$$

Це пов'язано з існуванням у електрона власного моменту імпульсу – спіну (див. далі, Лекція 6, п. 6.3), чого не враховує рівняння Шрьодінґера.

Для позначення вироджених станів електрона в атомі використовують спеціальну символіку, в якій відображають головне квантове число n та орбітальне квантове число l, причому головне – цифрою, а орбітальне – латинською літерою за такою схемою:

l	0	1	2	3	4	5
Символ	s	p	d	f	g	h

Згідно з цією схемою, основний стан електрона в атомі водню ($n=1$, $l=0$) позначається як $1s$, стан з $n=2$, $l=1$ – як $2p$, стан з $n=4$, $l=2$ – як $4d$, тощо.

Правила відбору. Як відмічалося раніше, при переході електрона з даного енергетичного рівня на нижчий випромінюється фотон відповідної енергії. Але фотон, як і електрон, має власний момент імпульсу (спін), і відносить з атома не лише енергію, а й момент імпульсу. Тому при переході електрона в атомі з одного енергетичного рівня на інший можливі лише такі зміни його стану, що відповідають величині та можливим орієнтаціям моменту імпульсу випущеного фотона і узгоджуються із законом збереження моменту імпульсу. Через це електронні переходи в атомі задовольняють певним *правилам відбору* – обмеженням на можливі зміни квантових чисел. Для атома водню вони такі:

$$\Delta l=\pm 1;\quad \Delta m_l=0,\pm 1. \qquad (5.10)$$

Для головного квантового числа n ніяких обмежень не існує.

Завдяки правилам відбору можливі переходи тільки між «сусідніми» квантовими станами: із s-станів можливі переходи тільки в p-стани, із p-станів – тільки в s- або d-стани, тощо. Через виродження рівнів правила відбору ніяк не впливають на частоти випромінювання вільних атомів Гідроґену, але суттєво відбиваються на спектрах багатоелектронних атомів і на поведінці атомів у магнітному та електричному полях.

5. РОЗПОДІЛ ЕЛЕКТРОННОЇ ГУСТИНИ В АТОМІ ГІДРОҐЕНУ

У напівкласичній борівській теорії вважається, що електрони в атомах рухаються по певних стаціонарних орбітах. Але, як вже не раз говорилося, таке уявлення не відповідає дійсності: просторова локалізація електронів у атомі визначається не траєкторіями, а конфігурацією хвильових функції, або інакше – ***орбіталями***.

Із рівняння Шрьодінґера (5.1) випливає, що хвильові функції електрона в атомі водню визначаються квантовими числами n, l, m_l і є досить складними, тому не будемо їх виписувати та детально аналізувати. Відмітимо лише, що головне квантове число n показує, як далеко від ядра може в середньому знаходитися електрон, а числа l і m_l визначають геометричну конфігурацію області його локалізації. При цьому всі s-орбіталі ($l = 0$) є сферично симетричними, -орбіталі ($l = 1$) нагадують «гантель», а для більш високих станів мають ще складнішу форму. Залежність імовірності знаходження електрона на відстані r від ядра $P_r(r)$ для s- та p-станів перших двох рівнів атома водню показана на рис. 5.6а, де r_1 – радіус першої борівської орбіти електрона в атомі водню. З цього рисунка видно, що область можливого знаходження електрона, тож і розміри атома, не є чітко визначеними. За розміри атома r_a можна прийняти найбільш імовірну відстань електрона від ядра, тобто відстань, на якій величина P_r є найбільшою.

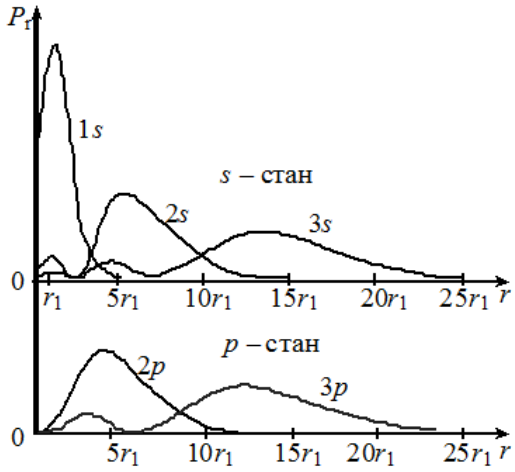
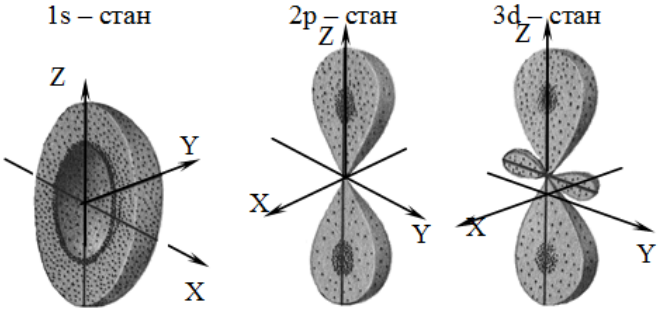


Рис. 5.6а

У такому контексті цікаво, що в стані $1s$ величина r_a дорівнює радіусу першої, а в стані $2p$ – другої борівської орбіти, причому така відповідність зберігається й для інших станів з максимальним значенням орбітального числа $l = n - 1$.

Таким чином, рух електрона в атомі («орбітальний» рух) мало нагадує рух матеріальної точки по траєкторії. При цьому, позаяк частота обертання електрона в атомі дуже велика ($\sim 10^{16}$ 1/с), ядро немов оповите неперервно розподіленим негативним зарядом, або «електронною хмариною». Деяке уявлення про її конфігурацію дає рис. 5.6б.



$l = 0,\ m_l = 0$ $l = 1,\ m_l = 0$ $l = 2,\ m_l = 0$

а) **б)** **в)**

Рис. 5.6б

6. ОРБІТАЛЬНИЙ МАГНІТНИЙ МОМЕНТ ЕЛЕКТРОНА

Орбітальний рух електрона в атомі створює циркуляцію заряду навколо ядра, себто – деякий ефективний електричний струм і відповідне ***орбітальне магнітне поле***. Із класичної електродинаміки відомо, що магнітні властивості малого витка зі струмом (магнітного диполя) визначаються магнітним моментом

$$\vec{p} = is\vec{n}, \qquad (5.11)$$

де i – величина струму, s – площа витка, $\vec{n}$ – орт (одиничний вектор) нормалі до площини витка, пов'язаний з напрямом струму правилом правого гвинта.

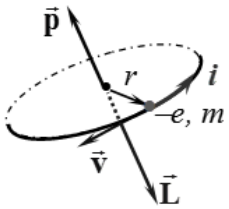


Рис. 5.7

Через невизначеність траєкторії класична формула (5.11) для знаходження ***орбітального магнітного моменту*** електрона в атомі, в принципі, є не коректною. Але квантовомеханічний розрахунок дає такий самий результат. Тому можемо тимчасово розглядати електрон як матеріальну точку із масою m та зарядом $-e$, яка рухається із великою швидкістю v по коловій траєкторії радіуса r й уподібнюється мікроскопічному витку площею $s = \pi r^2$ з деяким струмом i, як показано на рис. 5.7. Оскільки сила струму дорівнює кількості заряду, що переноситься через дану поверхню за 1 с, то, уявивши на шляху електрона деяку площадку, знайдемо:

$$i = e\nu = \frac{e}{T} = \frac{ev}{2\pi r},$$

де ν – частота, а $T = 2\pi r/v$ – період обертання електрона навколо ядра. Відтак із виразу (5.11) для модуля магнітного моменту отримаємо:

$$p = i \cdot \pi r^2 = \frac{evr}{2}.$$

При русі частинки по колу момент імпульсу $L = mvr$, тож $vr = L/m$, і вираз p можна записати, як $p = (e/2m)L$, або

$$p = gL, \qquad (5.12)$$

де константа

$$g = \frac{e}{2m}, \qquad (5.13)$$

називається ***орбітальним гіромагнітним співвідношенням***.

Оскільки заряд електрона від'ємний, напрям струму i є протилежним до напрямку руху електрона, тож вектори $\vec{p}$ і $\vec{L}$ антипаралельні. Тому

$$\vec{p} = -g\vec{L}, \qquad (5.14)$$

$$p_z = -gL_z. \qquad (5.14a)$$

Орбітальний момент імпульсу електрона квантується за правилами (5.6) і (5.7), отже згідно з (5.12) і (5.14a) аналогічно квантуються модуль і проекція орбітального магнітного моменту електрона:

$$p = \mu_B\sqrt{l(l+1)}, \qquad (5.15)$$

$$p_z = -\mu_B m_l, \qquad (5.15a)$$

де величина

$$\mu_B = g\hbar = \frac{e\hbar}{2m} = 9{,}27 \cdot 10^{-24}\ \text{А} \cdot \text{м}^2 \qquad (5.16)$$

називається ***магнетоном Бора*** і є природньою одиницею магнітного моменту електрона. Подібна природня одиниця використовується і для визначення магнітних моментів ядерних частинок – нуклонів. Вона називається ядерним магнетоном μ_{β} і визначається через масу протона:

$$\mu_{\beta} = \frac{e\hbar}{2m_p} = 5{,}05 \cdot 10^{-27}\ \text{À} \cdot \text{ì}^{\,2}. \qquad (5.16a)$$

Зауважимо, що через велику порівняно з електроном масу нуклонів $\mu_{Я} \ll \mu_{Б}$, тому магнітні властивості атома визначаються головним чином магнетизмом електронів.

7. КОНТРОЛЬНІ ЗАПИТАННЯ

1. Чому рух електрона навколо ядра в атомі не можна розглядати по аналогії з рухом Землі навколо Сонця?
2. Чи може атом поглинати світло будь-якої частоти? Чому?
3. Світіння газорозрядної трубки заповненої воднем спричинюється зіткненнями електронів із атомами водню. В якому випадку ці зіткнення є пружними, а в якому – непружними?
4. Назвіть квантові числа електрона в атомі Гідроґену та охарактеризуйте їх фізичний зміст.
5. Які значення може набувати кожне з квантових чисел? Яку максимальну кількість значень може мати магнітне квантове число m_l при заданому головному квантовому числі n?
6. Яку інформацію про енергетичний спектр електрона в атомі Гідроґену можна отримати, вимірявши його енергію іонізації?
7. Що таке спектральна серія? Головна лінія та короткохвильова межа серії? Запишіть вирази та розрахуйте довжини хвилі цих ліній для четвертої серії (серії Пфунда) Гідроґену.
8. Відомо, що в спектрі поглинання незбуджених атомів Гідроґену спостерігаються тільки лінії першої спектральної серії (серії Лаймана). Поясніть, чому?
9. Поясніть принцип класифікації квантових станів електрона в атомі Гідроґену за схемою цифра-літера. Чи може електрон бути в стані $2f$? $3d$?
10. Як записуються правила відбору для електронних переходів у атомі Гідроґену? Використовуючи прийняту класифікацію станів (цифра-літера), запишіть узагальнені позначення всіх електронних переходів у атомі Гідроґену, що утворюють серію Лаймана ($n_1 = 1$) та серію Пашена ($n_1 = 3$).
11. Які значення орбітального магнітного моменту може мати електрон у атомі водню? Запишіть (через магнетон Бора) цей момент для електрона, що перебуває в якомусь із s-станів, d-станів.

6.6. СПІН ЕЛЕКТРОНА

Нижче мова йде про таке:

1. Атом Гідроґену у зовнішньому магнітному полі
2. Неповнота рівняння Шрьодінґера
3. Спін
4. Принцип Паулі. Оболонкова структура атомів
5. Характеристичні рентгенівські спектри
6. Контрольні запитання

1. Атом гідроґену у зовнішньому магнітному полі

Розщеплення енергетичних рівнів. Із електродинаміки відомо, що при внесенні витка із струмом (магнітного диполя) в зовнішнє магнітне поле він отримує енергію

$$W = -\vec{p}\vec{B} = -p_z B, \tag{6.1}$$

де $\vec{B}$ – індукція магнітного поля, p_z – проекція магнітного моменту витка на напрям поля. При внесенні в магнітне поле атома Гідроґену його електрон так само взаємодіє з полем. При цьому, згідно з умовою квантування орбітального магнітного моменту (5.15*а*), енергія взаємодії з полем визначається як:

$$W = \mu_A B \cdot m_l, \tag{6.1a}$$

а енергетичні рівні атома — як

$$E_n = E_{n0} + \mu_A B \cdot m_l, \tag{6.2}$$

де E_{n0} – енергія рівня за відсутності поля.

На перший погляд може здатися, що відбувається просто зсув рівнів. Але через виродження ситуація є складнішою. Згідно з умовами (5.2), кожному рівню E_n відповідає n різних значень орбітального числа l, а кожному значенню l відповідає $2l+1$ різних значень m_l. Оскільки максимальна кількість значень орбітального числа $l_{\max} = n-1$, то повна кількість можливих значень m_l для заданого рівня складає $2l_{\max}+1 = 2n-1$. Тому кожен енергетичний рівень атома Гідроґену E_{n0} у зовнішньому магнітному полі *розщеплюється* на $2n-1$ симетрично розташованих компонент з енергіями

$$E_{n0},\ E_{n0} \pm \mu_A B,\ E_{n0} \pm 2\mu_A B,\ \ldots,\ E_{n0} \pm (n-1)\mu_A B. \tag{6.2a}$$

Виняток становить лишень основний рівень $1s$ $(l=0)$, на якому магнітний момент електрона $p=0$, і його взаємодія з магнітним полем відсутня. При цьому відстань між сусідніми компонентами будь-якого розщепленого енергетичного рівня атома однакова й визначається тільки величиною поля:

$$\Delta E_1 = \mu_A B. \tag{6.3}$$

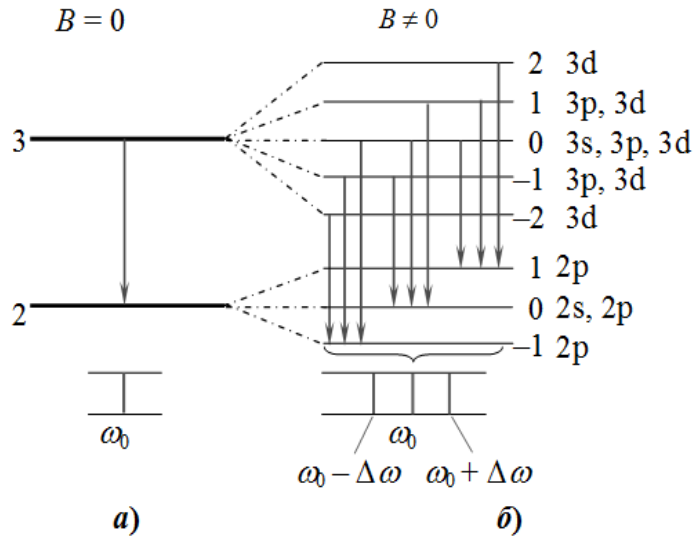


Рис. 6.1

На рис. 6.1 показана картина розщеплення рівня $n = 2$, коли кількість утворених компонент дорівнює 3, та рівня $n = 3$, для якого кількість компонент дорівнює 5. Принагідно відмітимо, що в системі утворених рівнів найвищий і найнижчий є невиродженими, а кратність виродження інших менша, ніж у вихідного рівня. Це видно з правої колонки на рисунку, в якій позначені всі стани, що відповідають кожній компоненті розщепленого рівня. Цей ефект, який називається *зняттям виродження*, спостерігається також у полях іншого походження і відіграє в квантових системах значну роль. Приміром, при об'єднанні атомів у кристал створюється так зване кристалічне поле, що призводить до зміни положення енергетичних рівнів електронів у вільних атомах і утворення енергетичних зон, поняття про які одним із основних у фізиці твердого тіла та фізиці напівпровідників.

Простий ефект Зеємана. Розщеплення енергетичних рівнів автоматично призводить до відповідного розщеплення спектральних ліній атомів у зовнішньому магнітному полі – явища, що називається ***ефектом Зеємана*** (Подібне явище спостерігається також в електричному полі і називається ефектом Штарка).

Розглянемо це явище для атома Гідроґену на прикладі електронного переходу між енергетичними рівнями $n = 3$ і $n = 2$, який схематично показаний стрілкою на рис. 6.1а.

У вільних атомах такий перехід дає одну спектральну лінію з частотою

$$\omega_0 = \frac{E_{03} - E_{02}}{\hbar}.$$

Але при вміщенні в магнітне поле дані енергетичні рівні розщеплюються (рис. 6.1б). Тому в спектрі замість однієї лінії з частотою ω_0 з'являється декілька інших ліній, зумовлених переходами між окремими компонентами розщеплених рівнів. На перший погляд нових ліній має бути стільки, скільки є комбінацій компонент розщеплених рівнів. Але насправді їх набагато менше через правила відбору (5.10), відповідно до яких можливі тільки три групи по три переходи, показані на рис. 6.1б. Оскільки всі компоненти розщеплених рівнів розміщені на однаковій відстані $\Delta E = \mu_A B$, випромінюються фотони всього з трьома різними енергіями: ΔE_0, $\Delta E_0 - \Delta E$, та $\Delta E_0 + \Delta E$. Отже, при вміщенні в магнітне поле в спектрі випромінювання атома замість кожної спектральної лінії з частотою ω_0 має з'явитися три лінії із частотами:

$$\omega_0 \text{ і } \omega = \omega_0 \pm \Delta\omega,$$

де

$$\Delta\omega = \frac{\Delta E_1}{\hbar} = gB. \qquad (6.4)$$

Таким чином, в магнітному полі кожна одиночна спектральна лінія («синглет») з частотою ω_0 має перетворитися на «триплет», який складається із центральної лінії з частотою ω_0 та двох симетричних сателітів, частоти котрих відрізняються від ω_0 на величину $\Delta\omega$, що залежить тільки від індукції магнітного поля (формула (6.4)). Таке розщеплення спектральних ліній за відповідних умов спостерігається експериментально й називається *простим* ефектом Зеємана. Правда, слід зазначити, що зеєманівське розщеплення є доволі малим (див. п. 6.7, Задача 6.1). Тому ефект Зеємана можна спостерігати тільки в сильних полях і за допомогою спектроскопа з високою роздільною здатністю.

Зазначимо також, що ефект Зеємана є експериментальним засвідченням просторового квантування моментів.

2. Неповнота рівняння Шрьодінґера

Квантова механіка, що ґрунтується на рівнянні Шрьодінґера (механіка Шрьодінґера) дає багату інформацію про будову та властивості не лише атома Гідроґену, а й складних атомів, молекул і конденсованої речовини. Разом із тим існують ефекти, які ця теорія пояснити не може, що вказує на її неповноту. Вкажемо на деякі з них.

Складний ефект Зеємана. Розглянутий вище простий ефект Зеємана є наслідком просторового квантування орбітального магнітного моменту електрона та правил відбору для електронних переходів. При цьому ані кількість дозволених переходів, ані різниця їхніх енергій (відстань між спектральними лініями (6.4)) не залежать від величини магнітного моменту електрона. Тому простий ефект Зеємана мав би спостерігатися для всіх атомів. Але насправді він є радше винятком, а не правилом. Для більшості атомів типовим є ***складний ефект Зеємана***, при якому різні спектральні лінії мають різні величину розщеплення та кількість розщеплених компонент. Отож, ефект Зеємана, підтверджуючи одне із принципових передбачень механіки Шрьодінґера – просторове квантування, – в той же час засвідчує, що магнетизм атома не можна пояснити лише орбітальним рухом електронів.

Тонка структура атомних спектрів. При спостереженнях за допомогою приладів із високою роздільною здатністю виявляється, що спектральні лінії атомів навіть за відсутності зовнішніх полів мають мультиплетну структуру, тобто складаються з декількох близько розташованих компонент. Спостерігаються дублети (подвійні лінії), триплети (потрійні лінії), квартети й таке інше. Зокрема трапляються і синглети – лінії, що не мають тонкої структури. У цьому контексті прикметно, що в спектрах Гідроґену та лужних металів усі лінії є дублетами, хоча мали би бути синглетними через орбітальне виродження енергетичних рівнів.

Досліди Штерна та Герлаха. Ці досліди, в яких вивчався вплив неоднорідного магнітного поля на атомарні пучки, дали прямі докази просторового квантування моментів. Схема досліду показана на рис. 6.2. Вузький пучок атомів досліджуваної речовини пропускався через сильне поперечне магнітне поле, індукція котрого $\vec{B}$ різко змінювалася вздовж осі OZ, і далі осаджувався на скляній пластинці-екрані.

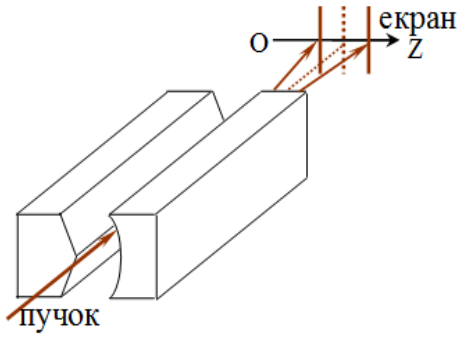


Рис. 6.2

Можливий результат досліду можна передбачити. За відсутності магнітного поля атоми рухаються без відхилення і створюють вузьку смужку в центрі екрана. Але, через наявність магнітного моменту $\vec{p}$, при ввімкненні указаного поперечного поля $\vec{B}$ на атоми вздовж осі OZ діє сила

$$F_z = p\frac{\partial B}{\partial z}\cos\vartheta = p_z\frac{\partial B}{\partial z}, \qquad (6.5)$$

де $(\partial B/\partial z)$ – градієнт поля, ϑ – кут між напрямками векторів $\vec{p}$ і $\vec{B}$ (див. Розділ 3, Лекція 3.9). Під дією цієї сили атоми в пучку відхиляються від напрямку вльоту в поле і потрапляють в різні точки екрана залежно від величини та напрямку магнітного моменту. При довільній орієнтації магнітних моментів атомів пучка значення p_z лежали б у неперервному інтервалі $[-p, p]$, і на них діяли би всі можливі сили (6.5) теж у неперервному інтервалі значень $\left[-p(\partial B/\partial z),\ p(\partial B/\partial z)\right]$. Тому на екрані мала би спостерігатися широка суцільна смуга. Проте в дійсності утворюється система вузьких симетричних смужок. Це прямо вказує на те, що можливі лише вибрані значення p_z, тобто має місце просторове квантування магнітного моменту атома. При цьому за правилами квантування моментів (5.15a) і (5.2) кількість можливих значень p_z, тож і розщеплених компонент у пучку, завжди мала би складати $2l+1$, тобто бути непарною. Але в дослідах спостерігалося також розщеплення і на парну кількість компонент. Зокрема це стосується срібла, гідроґену та інших атомів з одним валентним електроном, який знаходиться в s-стані $(l=0,\ m_l=0)$. Для таких елементів спостерігається дві компоненти, як показано на рис. 6.2, хоча розщеплення взагалі не мало би бути.

Дослід Ейнштейна – де-Гааза. Цей дослід мав на меті визначення гіромагнітного співвідношення для феромагнетиків у рамках класичної теорії магнетизму речовини. Зупинимося лише на загальній схемі та ідеї досліду. Залізний циліндр, підвішений на тонкій пружній нитці, був розміщений всередині соленоїда із змінним струмом, як показано на рис. 6.3. Під дією змінного магнітного поля соленоїда циліндр перемагнічувався, тобто напрямки магнітних моментів окремих атомів і циліндра в цілому періодично змінювалися. Відповідно змінювався й напрямок моменту сил, що діяв на циліндр з боку поля, через що він здійснював малі вимушені крутильні коливання. Для підсилення ефекту частоту струму підбирали рівною резонансній частоті циліндра, так що амплітуду коливань можна було виміряти достатньо точно. Далі, не вдаючись у деталі, вкажемо, що, вимірявши амплітуду коливань циліндра, за допомогою законів механіки можна знайти величину моменту імпульсу всього циліндра, та окремих атомів. Так само, за допомогою законів електродинаміки через величину струму та магнітну проникність матеріалу циліндра можна визначити відповідні магнітні моменти. Відтак можна встановити й відношення магнітного та механічного моментів атома, тобто гіромагнітне співвідношення.

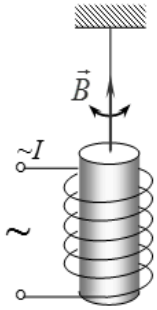


Рис. 6.3

Знайдена в описаному досліді величина гіромагнітного співвідношення виявилася з великою точністю рівною

$$g_s = \frac{e}{m}. \qquad (6.6)$$

Ця величина є вдвічі більшою за очікуване значення (5.13), тому її назвали *аномальним* гіромагнітним співвідношенням.

Прикметно, що природу аномального гіромагнітного співвідношення нездатна пояснити не тільки класична фізика, а й квантова механіка, що ґрунтується на рівнянні Шрьодінґера.

3. Спін електрона

Спін електрона. Наведені вище та інші розбіжності між передбаченнями механіки Шрьодінгера та дослідними фактами, як з'ясувалося, пов'язані з існуванням у електрона ***спіну***.

Гіпотеза про спін була висунута в 1925 р. американськими фізиками Гаудсмітом і Уленбеком і твердить, що

електрон, незалежно від руху, має власний момент імпульсу – спін – та відповідний власний (спіновий) магнітний момент.

(Термін спін, який перекладається як веретено або дзиґа, пов'язаний із первісним і, як з'ясувалося, **хибним** уявленням про обертання електрона навколо своєї осі).

Згідно із загальними умовами квантування моменту імпульсу (5.6) і (5.7), модуль і проекція спіну визначаються формулами

$$\begin{aligned} L_s &= \hbar\sqrt{s\left(s+1\right)}; \\ L_{sz} &= \hbar m_s \end{aligned} \qquad (6.7)$$

через спінове квантове число s і спінове магнітне квантове число m_s, які коротко називають одним словом ***спін***.

Значення цих квантових чисел можна встановити із наведених вище результатів дослідів Штерна та Герлаха. Відомо, що основні хімічні та фізичні властивості атомів визначаються практично тільки валентними електронами. Тому в атомах Гідрогену та срібла, що мають один валентний електрон у s-стані ($l=0$), орбітальний магнітний момент $p=0$, і ніякої взаємодії з магнітним полем не повинно бути. Отже, природньо припустити, що розщеплення пучків цих атомів у магнітному полі зумовлене власним магнітним моментом електрона – спіном. Зв'язок між квантовими числами m_s і s такий самий, як між m_l і l, тобто загальна кількість можливих проекцій спіну на напрям магнітного поля дорівнює $2s+1$. Оскільки для вказаних атомів у досліді спостерігається розщеплення пучка на дві компоненти, то

$$2s+1=2 \quad \Rightarrow \quad s=\frac{1}{2}. \qquad (6.8)$$

Цей результат привертає увагу тим, що квантове число s є не цілим і має єдине можливе значення. Відповідно, значення m_s, які лежать в інтервалі $\left[-s;s\right]$, складають $\mp(1/2)$. Це, згідно з (6.7), означає, що власний момент імпульсу електрона та його можливі проекції на вісь дорівнюють:

$$L_s=\hbar\frac{\sqrt{3}}{2}; \qquad L_{sz}=\pm\frac{\hbar}{2}. \qquad (6.9)$$

Існування спіну в електрона та його характеристики узгоджуються і з іншими дослідними фактами. Зокрема, аномальна величина гіромагнітного співвідношення g_s у феромагнетиках пояснюється тим, що їхні магнітні властивості зумовлені саме спіновими магнітними моментами атомів. Отже знайдена в дослідах Ейнштейна – де-Гааза величина $g_s = (e/m)$ є ***спіновим гіромагнітним співвідношенням*** для електрона. Відповідно, ***власний магнітний момент електрона*** визначається виразом:

$$\vec{p}_s = -g_s \vec{L}_s \qquad (6.10)$$

і, згідно з (6.6) і (6.9), має модуль і проекції

$$p_s = \mu_Б \sqrt{3}; \qquad p_{sz} = \pm \mu_Б, \qquad (6.11)$$

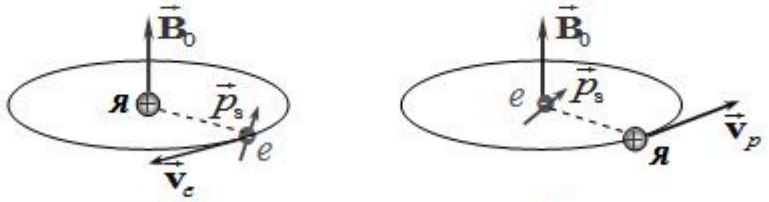
де $\mu_Б$ – магнетон Бора (5.16).

Отже, магнетон Бора є природньою мірою не тільки орбітального, а й власного магнітного моменту електрона.

Природа спіну. Спочатку спін пов'язували з обертанням електрона навколо власної осі, подібно до добового обертання Землі. Але така концепція виявилася хибною. По-перше, величина L_s, згідно з (6.9), має єдине значення, отже, не може змінюватися ні за яких умов. Це свідчить про те, що спін відображає не стан руху електрона, а його внутрішні властивості, так само, як заряд або маса. По-друге, аби при обертанні навколо своєї осі момент імпульсу кульки з розмірами і масою електрона мав величину (6.9), лінійна швидкість точок її поверхні мала би бути більшою за граничну швидкість $c = 3 \cdot 10^8$ м/с. Тож природньо, що нерелятивістська механіка Шрьодінгера не дає ніяких указівок на існування спіну. А ось з релятивістського хвильового рівняння, яке отримав Дірак, існування спіну випливає автоматично. А саме, з рівняння Дірака випливає, що хвильові функції електрона в атомі, крім відомих чисел n, l, m_l, визначаються ще одним квантовим числом із двома можливими значеннями $\pm 1/2$, тобто, – спіном. Таким чином, можна сказати, що спін є одночасно і квантовою, і релятивістською властивістю частинки.

Спін є властивістю не лише електрона, а й переважної більшості інших елементарних частинок. При цьому існують частинки як із напівцілим, так і з цілим значенням спіну. Зокрема, із добре знайомих частинок електрон, протон і нейтрон мають спін 1/2, а спін фотона дорівнює 1.

Спін-орбітальна взаємодія. Наявність спіну призводить до того, що в атомі Гідрогену кожному набору квантових чисел (n, l, m_l) відповідає не один, а два квантових стани з $m_s = 1/2$ і $m_s = -1/2$, які відрізняються орієнтацією спіну. Ці орієнтації часто умовно називають «спін угору» та «спін униз» і позначають стрілками $\uparrow$ і $\downarrow$. Оскільки ці два квантові стани належать одному значенню головного квантового числа n, іноді говорять, що наявність спіну приводить до подвоєння кратності виродження енергетичних рівнів. Але, строго говорячи, таке твердження є неточним, бо вказані спінові стани мають не однакову енергію. Причина в тому, що спіновий магнітний момент електрона $\vec{p}_s$ перебуває в магнітному полі $\vec{B}_0$, створюваному орбітальним рухом електрона, і взаємодіє з ним, рис. 6.4*а*. Можливо це буде більш зрозуміло, якщо перейти в систему відліку, пов'язану з електроном (рис. 6.4*б*), в якій протон (ядро атома) рухається навколо електрона так само, як і електрон навколо ядра, тільки в протилежний бік. Проте й знак заряду протона протилежний до знаку заряду електрона. Тому поле, створюване ядром атома в системі відліку електрона за величиною та напрямком співпадає з полем $\vec{B}_0$, яке створюється орбітальним рухом електрона в системі відліку ядра.



а) *б*)

Рис. 6.4

Взаємодія спінового моменту електрона $\vec{p}_s$ з орбітальним полем $\vec{B}_0$ називається ***спін-орбітальною взаємодією***. За рахунок неї електрон, згідно з (6.1) і (6.11), залежно від орієнтації спіну отримує додаткову енергію

$$W_{so} = -\vec{p}_s \vec{B}_0 = \pm \mu_Б B_0. \qquad (6.12)$$

Тому при заданому головному квантовому числі n енергія електрона складає

$$E_{\text{п}} = E_{\text{п}0} \pm \mu_{\text{Б}} B_0,$$

де $E_{\text{п}0}$ – енергія n-го рівня без урахування спіну, котра визначається формулою (5.3). Отже, спін-орбітальна взаємодія в атомі водню призводить до того, що кожен енергетичний рівень (окрім основного) виявляється розщепленим на два підрівні (рис. 6.5), відстань між якими

$$\Delta E_{\text{so}} = 2\mu_{\text{Б}} B_0. \qquad (6.13)$$

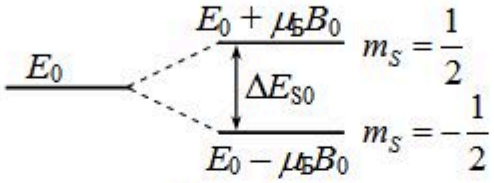


Рис. 6.5

Відповідно, двократно розщепленими (дублетними) виявляються й усі спектральні лінії атома водню. Величина розщеплення $\Delta\omega_{\text{so}} = \Delta E_{\text{so}}/\hbar$, згідно з (5.16), дорівнює:

$$\Delta\omega_{\text{so}} = g_{\text{s}} B_0, \qquad (6.13a)$$

де g_{s} – спінове гіромагнітне співвідношення (6.6).

Спін-орбітальна взаємодія пояснює і дублетність спектральних ліній атомів лужних металів, які мають в основному стані один s-електрон і тим схожі на атом водню. Аналогічно пояснюється тонка структура спектрів й у складніших атомах. Але через наявність декількох валентних електронів картина спін-орбітального розщеплення рівнів у таких атомах є набагато складнішою.

Через велику швидкість орбітального руху та дуже малі розміри електронних «орбіт» орбітальні магнітні поля в атомах є дуже сильними і складають десятки Тл. Тому спін-орбітальне розщеплення спектральних ліній може бути навіть сильнішим, аніж у зовнішніх полях (ефект Зеємана), де воно має порядок 10^{-11} м. Приміром, для найяскравішої лінії в спектрі атомів натрію (жовтий дублет) спін-орбітальне розщеплення дорівнює $6 \cdot 10^{-10}$ м.

4. Принцип Паулі. Оболонкова структура атомів

Принцип Паулі. Одним із проявів спіну є не однакова колективна поведінка частинок у залежності від величини спіну. Це відображає сформульоване Паулі на основі аналізу дослідних фактів і пізніше доведене теоретично положення, що називається ***принципом*** або ***забороною*** Паулі:

> ***у системі частинок із напівцілим спіном у кожному квантовому стані не може одночасно перебувати більше, ніж одна частинка.***

Стосовно електрона в атомі це означає, що в кожному стані із заданим набором квантових чисел n, l, m_l, може перебувати не більше двох електронів із протилежно орієнтованими спінами. Для частинок із цілочисловим спіном таких обмежень немає.

Указані розбіжності зумовлюють не однаковий розподіл по енергіях у макроскопічних системах частинок із різним типом спіну. Для частинок із напівцілим спіном він є підпорядкований закону розподілу Фермі-Дірака, а для частинок із цілочисловим спіном діє закон розподілу Бозе-Ейнштейна. Відповідно, частинки з напівцілим спіном звуться ***ферміонами***, а з цілочисловим – ***бозонами.***

Оболонкова структура атома. Те, що електрони є ферміонами, суттєво відображається на властивостях електронних систем, зокрема таких, як метали та атоми. Приміром, на кожному енергетичному рівні в металі може бути тільки по два електрони. Тому навіть при температурі абсолютного нуля вільні електрони займають кількість рівнів порядку кількості атомів у зразку металу і мають дуже велику середню кінетичну енергію. Це суттєво понижує для них ефективну висоту потенціального бар'єра на межі з вакуумом, що уможливлює зовнішній фотоефект та тунельні явища, про які говорилося в попередніх лекціях.

В атомі електрони теж не можуть усі перебувати на найнижчому енергетичному рівні навіть у основному стані, коли атом має найменшу можливу енергію. Вони розподіляються по найнижчих рівнях так, аби в кожному квантовому стані був лише один електрон. Тому максимальна кількість електронів $N_{\max}$ у станах із заданим головним квантовим числом n дорівнює (Див. Лекція 6.5)

$$N_{\max} = 2n^2. \qquad (6.14)$$

Отже, на першому рівні може бути максимум 2 електрони, на другому – 8, третьому – 18, тощо.

Квантове число n є головним чинником, який визначає не лише енергію електрона, а й середню відстань електрона від ядра. Тому всі електрони, що перебувають на заданому рівні n, є певною мірою просторово відокремленими від інших і утворюють відповідну *електронну оболонку* атома. Оболонки прийнято позначати великими латинськими літерами, починаючи від K: $K, L, M, \ldots$. При цьому K-оболонка містить рівень $1s$, L –оболонка – рівні $2s, 2p$, M-оболонка – рівні $3s, 3p$ і $3d$, тощо. У межах оболонки також спостерігається розшарування електронів залежно від геометричної конфігурації орбіталей, яка визначається квантовим числом l. Через це всі електрони, що мають задане значення l, утворюють відповідну ***підоболонку*** даної оболонки. Максимальна кількість електронів у підоболонці дорівнює кількості квантових станів електрона із заданим значенням l і, з урахуванням спіну, складає

$$N'_{\max} = 2(2l+1). \qquad (6.14a)$$

Зауважимо, що інколи оболонки називають «електронними шарами», а підоболонки – «оболонками».

Залежно від фактичної кількості електронів N в оболонці, вона називається *замкненою* (якщо $N = N_{\max}$), або *незамкненою* (якщо $N < N_{\max}$), так само й для підоболонок. Відмітимо, що, згідно з (6.14) і (6.14*a*), замкнена оболонка чи підоболонка завжди містить парну кількість електронів.

Властивості оболонок і періодична система елементів. Теорія і дослід свідчать про те, що властивості оболонок і підоболонок суттєво залежать від ступеня їх заповненості («заселеності») електронами. Зокрема, замкнені оболонки та підоболонки не мають електричного та магнітного моментів і не взаємодіють із зовнішніми полями. Про це опосередковано свідчить той факт, що, сума всіх значень m_l і m_s у замкненій оболонці чи підоболонці дорівнює нулю. Електрони в замкненій оболонці чи підоболонці сильно зв'язані з ядром та між собою. Тому для їхнього вивільнення (внутрішньої іонізації атома) потрібна велика енергія ($>10^3$ еВ), у той час як для виривання електрона з незамкненої оболонки треба всього порядку 10 еВ. Всередині замкненої оболонки електрон теж не може змінити свій стан, бо всі дозволені стани зайняті іншими електронами цієї оболонки. Через указані особливості замкнені оболонки практично не беруть участі у взаємодії атомів з іншими частинками та зовнішніми полями, і всі основні фізико-хімічні властивості хімічних елементів визначаються тільки електронами незамкнених (валентних) оболонок.

Властивості незамкнених оболонок теж залежать від кількості електронів у них. Зокрема, атоми із майже заповненими валентними оболонками легко забирають недостатні електрони в інших атомів і утворюють міцні негативні іони із замкненою електронною структурою. Яскравим представником таких хімічних елементів є галогени: F, Cl, Br, ... З іншого боку, коли валентна оболонка атома далека від заповнення, він легко позбавляється «зайвих» електронів, утворюючи міцний позитивний іон із замкненою електронною структурою. Найкраще така властивість виражена в лужних металів: Li, Na, K, ... Ці речовини є дуже хімічно активними, зокрема, з'єднуючись із галогенами, вони утворюють міцні молекули лужно-галоїдних сполук – солей, до яких належить і харчова сіль $NaCl$. Показовим прикладом є також інертні гази He, Ne, Ar, ... Валентні електрони в атомах цих елементів утворюють замкнені оболонки або підоболонки. Тому інертні гази за звичайних умов існують тільки в атомарному стані й не утворюють хімічних сполук з іншими елементами.

Електронна структура будь-якого атома, тобто кількість і тип його електронних оболонок і підоболонок, визначається загальною кількістю електронів та умовою мінімуму енергії атома в основному стані. Не зупиняючись на цьому детально, вкажемо лише, що позаяк кількість електронів у оболонках є обмеженою, їхня структура в міру збільшення загальної кількості електронів у атомі періодично повторюється. Відповідно повторюються і характерні властивості хімічних елементів, що відображає періодична система хімічних елементів Менделєєва.

Принагідно зауважимо, що періодична система була відкрита на основі емпіричних даних, але вичерпне теоретичне обґрунтування і завершеність вона отримала лише після створення квантової механіки.

5. Характеристичні рентгенівські спектри

Енергетичний спектр багатоелектронних атомів. Багатоелектронні атоми відрізняються від Гідроґену більшим зарядом ядра $q = Ze$, де Z – порядковий номер елемента в таблиці Менделєєва. Тому, на перший погляд, енергетичні рівні мали би визначатися виразом (6.6*a*) і розміщуватися глибше, ніж у Гідроґені, причому тим глибше, чим далі розміщений елемент у періодичній системі. Але насправді енергетичні і оптичні спектри багатоелектронних атомів набагато складніші, позаяк в них квантові стани визначаються взаємодією електронів не лише з ядром, а й між собою.

Одним із наслідків електрон-електронної взаємодії є утворення електронних оболонок, через що різні групи електронів у багатоелектронному атомі певною мірою відокремлені від інших і знаходяться на суттєво різній середній відстані від ядра. Тому в першому наближенні можна вважати, що для електронів даної оболонки електричне поле ядра частково послаблене (екрановане) електронами, що знаходяться ближче до нього. Через це найвіддаленіші від ядра електрони зовнішніх (валентних) оболонок усіх атомів, окрім гелію, зв'язані з ядром навіть слабше, ніж у Гідроґені. На це вказують, зокрема, відносно невеликі енергії іонізації атомів та величини частот у спектрах випромінювання багатоелектронних атомів. Але електрони внутрішніх оболонок зв'язані з ядром дуже сильно, тому їхні енергетичні рівні залягають дуже глибоко. Відстань між рівнями різних оболонок теж є дуже великою і для атомів із великим зарядом ядра складає $\sim 10^4$ еВ.

Ще одна особливість спектрів багатоелектронних атомів полягає в тому, що в них енергія квантових станів залежить не лише від головного n, а і від орбітального квантового числа l. Це пояснюється тим, що навіть у межах однієї оболонки хвильові функції (орбіталі) електронів із різних підоболонок мають різні розміри та просторову конфігурацію, і тому для них ступінь екранування ядра іншими електронами є не однаковим. З цієї причини енергетичні та оптичні спектри багатоелектронних атомів є набагато складнішими, ніж у Гідроґену.

Характеристичні рентгенівські спектри. За звичайних умов переходи електронів між рівнями замкнених внутрішніх оболонок атома є неможливими через відсутність на них «вільних місць». Але такі переходи стають можливими при внутрішній іонізації атома – вириванні електрона із внутрішньої оболонки внаслідок опромінення жорсткими рентгенівськими променями або бомбардування швидкими електронами чи іншими частинками. В такому разі на відповідному рівні з'являється «дірка» – не зайнятий квантовий стан, у який може перейти електрон із більш високої оболонки. Відтак уже в ній з'явиться дірка, в яку перейде електрон із наступної вищої оболонки, і т. д. У результаті таких каскадних переходів випромінюються фотони з упорядкованими частотами, що утворюють спектральні серії в рентгенівському діапазоні, як це схематично показано на рис. 6.6.

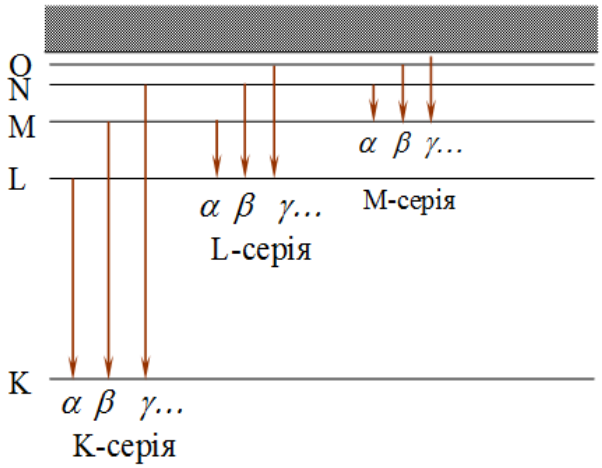


Рис. 6.6

Рентгенівські серії позначають літерою оболонки, на яку здійснюються переходи, а лінії в серії – нижнім індексом $\alpha, \beta, \gamma, \ldots$. Енергії переходів і частоти переходів є індивідуальними для кожного хімічного елемента, тому вказане рентгенівське випромінювання атомів називається ***характеристичним*** випромінюванням.

Закон Мозлі. Характеристичні рентгенівські спектри нагадують спектр атома Гідроґену, але є простішими. Це пояснюється тим, що навіть наймасивніші атоми мають тільки п'ять електронних оболонок. Тому кількість ліній в серії (кількість можливих переходів на дану оболонку) не перевищує чотирьох. Правда, через залежність енергії електронів у заданій оболонці не тільки від головного, а й від орбітального квантового числа, кожен рівень і кожна лінія рентгенівського спектра складається з відповідної кількості близько розташованих компонент. Але через дуже великі енергії і частоти переходів між оболонками така тонка структура рентгенівських спектрів не є суттєвою. Тому характеристичні рентгенівські спектри різних елементів вирізняються одноманітністю й невеликою кількістю ліній.

Ще одна особливість рентгенівських спектрів полягає в тому, що утворення хімічних сполук не впливає на характеристичні спектри атомів у складі молекули. Це прямо свідчить про те, що внутрішні оболонки атомів у молекулі не взаємодіють між собою, тобто що за хімічний зв'язок відповідають тільки валентні електрони. Енергії внутрішніх оболонок головним чином визначаються взаємодією електронів з ядром і залежать від його заряду. Тому існує закономірний зв'язок між частотами ліній характеристичного випромінювання даного хімічного елемента та його порядковим номером Z у періодичній системі, який задає заряд ядра $q = Ze$. Цей зв'язок виражається ***законом Мозлі***, що визначає частоту ω ***даної лінії в даній серії*** характеристичного спектра кожного елемента:

$$\sqrt{\omega} = C(Z - \sigma), \tag{6.15}$$

де Z – порядковий номер елемента в періодичній системі, а σ і C – константи, що залежать від серії та лінії в серії.

Слід одразу зауважити, що закон Мозлі не є точним. Він добре виконується для легких елементів, але для важких спостерігається деяке відхилення від лінійної залежності в (6.15), а також значні відміни констант σ і C для однієї й тієї ж лінії в одній і тій самій серії різних елементів.

Для розуміння закону Мозлі суттєво, що, як свідчить дослід, у випадку легких елементів для K_α-ліній $\sigma = 1$, а константу C можна подати як

$$C = \sqrt{R\left((1/1^2) - (1/2^2)\right)},$$

де R – стала Ридберга (6.7). Відтак частота лінії K_α

$$\omega = R(Z-1)^2\left(\frac{1}{1^2} - \frac{1}{2^2}\right). \tag{6.16}$$

Прикметно, що, згідно з (5.5*a*), цей вираз відповідає першій лінії першої серії в оптичному спектрі водневоподібного іона із зарядом ядра $q = (Z-1)e$. Це можна пояснити тим, що K_α-лінія зумовлена переходом електрона L-оболонки на K – оболонку, в якій лишився тільки один електрон. Відтак у процесі переходу L-електрон рухається в полі ядра, що послаблене (екрановане) одним K-електроном так, якби заряд ядра дорівнював $q_{\text{åô}} = (Z-1)e$. Отже, величина σ має зміст константи екранування. Те саме можна сказати й про L_α-лінії, для яких

$$C = \sqrt{R\left((1/2^2) - (1/3^2)\right)}$$

і

$$\omega = R(Z-\sigma)^2\left(\frac{1}{2^2} - \frac{1}{3^2}\right). \tag{6.17}$$

Але в цьому випадку дослід дає $\sigma \approx 5{,}5$, хоча за «арифметичною» логікою мало би бути $\sigma = 9$, оскільки для M-електрона при його переході на звільнене місце в L-оболонці ядро екранують 2 електрони K-оболонки та 7 електронів L-оболонки. Цю розбіжність можна пояснити тим, що, на відміну від компактної і близької до ядра K-оболонки, вищі оболонки частково перекриваються між собою. Тому для M-електрона, що здійснює перехід, поле ядра екранується не всім, а лише частиною заряду електронів L-оболонки. Подібне пояснення, ймовірно, мають і відмічені вище відхилення від простого закону (6.15) у важких елементах, які мають більшу кількість внутрішніх оболонок і більший ступінь їх перекриття.

На завершення зазначимо, що закон Мозлі відіграв велику роль в обґрунтуванні та становленні періодичної системи хімічних елементів Менделєєва. Цей закон показав, що місце елемента в періодичній системі визначається не масою атома, як вважалося, а зарядом ядра. Це дозволило остаточно впорядкувати таблицю хімічних елементів. Зокрема, Мозлі показав, що між Гідроґеном та Ураном має існувати рівно 92 елементи та визначив кількість і місце в таблиці невідомих на той час і відкритих пізніше елементів.

6. Контрольні запитання

1. Чи має, згідно з механікою Шрьодінґера, залежати кількість компонент, на які розщеплюється енергетичний рівень атома водню в зовнішньому магнітному полі від індукції поля? Від номера рівня?

2. Чи має, згідно з механікою Шрьодінґера, залежати відстань ΔE між компонентами розщепленого в зовнішньому магнітному полі енергетичного рівня атома водню від індукції поля? Від номера рівня?

3. Чи залежить кількість компонент, на які розщеплюється спектральна лінія атома Гідроґену в нормальному ефекті Зеємана від індукції поля? Від номерів рівнів, між якими відбувається електронний перехід?

4. Чи залежить величина розщеплення $\Delta\omega$ спектральної лінії атома в нормальному ефекті Зеємана від індукції поля? Від номерів рівнів, між якими відбувається електронний перехід?

5. Відомо, що нормальний ефект Зеємана спостерігається тільки для атомів із парною кількістю електронів. Чому?

6. Як пояснити, що в дослідах Штерна і Герлаха спостерігалося розщеплення пучків атомів, котрі не мають орбітального магнітного моменту?

7. Як пояснити, що спінове квантове число електрона s, на відміну від l, має єдине можливе значення?

8. Які електрони в атомі утворюють оболонку? підоболонку? Чому дорівнює максимальна можлива кількість електронів у оболонці? підоболонці?

9. Що таке валентна оболонка атома? Чому взаємодія між атомами за звичайних умов визначається тільки валентними електронами?

10. Як виникає характеристичне рентгенівське випромінювання?

11. Чому серії в характеристичному рентгенівському спектрі складаються з невеликої кількості ліній?

12. Спочатку елемент кобальт Co в періодичній таблиці розміщували за нікелем Ni в порядку зростання атомної маси (відповідно, 58,7 і 58,9). Але за допомогою закону Мозлі з'ясували, що вказану послідовність треба змінити. Поясніть, як це взнали.